Cognitive Significance in Neural Systems

Caleb Dewey

Department of Philosophy, University of Arizona

Every explanation of a complex functioning system faces a problem of abstraction: (a) which tiny minority of its processes to include in and (b) which vast majority of its processes to exclude from the explanation of its behaviour. Privileging certain levels of explanation is a popular solution to the problem of abstraction, but it fails to account for inter-level processes. Thus, there remains a need for an abstraction strategy that’s appropriate for multi-level explanations. A promising way to address this need is to identify the abstraction strategies that are implicit in the cognitive neuroscientific explanations of behaviour in fully-mapped behavioural circuits. In this paper, I apply this strategy to the behavioural circuit that the barn owl (Tyto alba) uses to localize sound sources in the frontal azimuth (i.e., the left-to-right plane). I argue that the best explanation of behavioural performance by this circuit includes only computations that are cognitively significant, which only includes any computation that is (a) a causal relation between two states that information about the task state and (b) that creates or satisfies functional needs for the cognitive system that are satisfied or created, respectively, by some other, distally-related component of the same system. I argue that cognitive significance supports a novel kind of empirical analysis into the relation between behaviour and anatomy. Yet I insist that cognitive significance is distinct from cognition: it is either (a) one kind of cognition among others or (b) a precursor to cognition. I conclude by suggesting that the results of this case study generalize to the explanations of behaviour in any complex functioning system.

Keywords: cognition, neuroscience, abstraction, level, information, structure-function relationship

There are many conceptions of levels in the philosophy of science but they are unified by the roles that they typically serve in abstraction. That is, scientists typically use levels to distinguish (a) the minority of information that should be included in their explanations from (b) the majority of information that should be excluded from their explanations. The simplest way to do this is to specify that one level of a system is privileged for explanation: (a) all information that should be included in an explanation belongs to that level of the system and (b) all information at other levels of the system should be excluded from the explanation. For example, reductionists argue that the lowest level of the system should be privileged for explanation to maximize precision (e.g., Bickle, 2003) whereas emergentists argue that the highest level of the system should be privileged for explanation to maximize generality (e.g., Fodor, 1974).

While these approaches exclude a lot of information, they are now thought to exclude too much information: scientific explanation is rarely limited to a single privileged level. Thus, they “over-
solve” the problem of abstraction. This has prompted two responses (Potochnik & de Oliveira, 2019). The first is explanatory pluralism: each level of a system is privileged by a unique kind of explanation. This accounts for the fact that different disciplines often privilege different levels, but it neglects inter-level explanations. The second is explanatory integration: explanations should integrate across levels without privileging any level. This accounts for inter-level explanations, but it restores the problem of abstraction that levels were supposed to solve. If we integrate information across levels into our explanations, then what information should we include and exclude from our explanation? If levels aren’t serving our need for abstraction, what is? Without an answer, explanatory integration “under-solves” the problem of abstraction.

Thus, there is a dilemma between solving the problem of abstraction and pursuing an inter-level explanation. For a few cognitive systems, known as behavioural circuits, cognitive neuroscience has managed to break this dilemma: it’s developed inter-level explanations that solve the problem of abstraction. These circuits generate simple behaviours: e.g., reflexive, rhythmic, and saccading movements. Some dismiss the relevance of these simple circuits, preferring instead to consider more complex systems that are recognizably “cognitive”. However, these simple circuits are the only systems that we understand nearly well enough that we can provide explanations across levels that solve the problem of abstraction (without privileging levels for explanation). For better or for worse, then, behavioural circuits are the only case studies that can help us solve the problem of abstraction for inter-level explanations.

In this paper, I’ll develop a solution to the problem of abstraction that’s suitable for inter-level explanations with the help of a case study into an 11-layer, 8-level behavioural circuit in the barn owl (Tyto alba). In §1, I’ll motivate the problem of abstraction. In §2, I’ll review this behavioural circuit. In §3, I’ll argue that the problem of abstraction is solved by including only computations that are cognitively significant—i.e., computations of states that (a) encode task information and (b) create and/or satisfy a functional need for the cognitive system that is satisfied and/or created by another, distally-related component of the cognitive system. In §4, I’ll argue that levels are indispensable for the empirical analysis of the flow of information through the discrete network of cognitively significant computations. In §5, I’ll argue that cognition doesn’t reduce to cognitive significance because cognitive significance is task-specific whereas certain cognitive kinds (e.g., capacities) are task-general. In §6, I’ll conclude by considering some methodological implications for all sciences of functioning systems.

§1. The Problem of Abstraction

As aforementioned, the problem of abstraction is to decide which computations should be included in and excluded from the best explanation of behaviour. For cognitive neuroscience, though, this problem is especially severe: there are easily millions of neurons that respond to even the simplest tasks to collectively compute even the simplest behaviour. And neurons are only one anatomical unit of computation: they are composed by smaller units (e.g., biochemical cascades) and they themselves compose even larger units (e.g., neural populations, whole-brain networks). Clearly, there is too much computation that mediates between any stimulus input and behavioural output.

---

2 Saccades are fast movements between points of sensory fixation: most often in the eyes.
to include all of it in the best explanation of behaviour. For explanation to be possible for cognitive neuroscientists with finite resources, there must be a good solution to this problem of abstraction.

Cognitive neuroscientists often privilege certain anatomical units when they explain behaviour: they use anatomical criteria to individuate levels (e.g., the cell level, the cell population level) and then they include only units from that anatomical level in their explanation of behaviour. When they do this, they’re committed to the claim that all computations that should be included in the best explanation of behaviour (or the best explanation of a certain kind, per explanatory pluralism) belong to an anatomical kind. For example, Barack & Krakauer (2021) distinguish Sherringtonian neuroscientists, who believe that the best explanations of all behaviour include only computation at the microcircuit (cell connections) level, from Hopfieldian neuroscientists, who tend to believe that the best explanations of certain behaviours include only computation at the brain region level.

These are very useful simplifying assumptions for experimental science: different techniques are suitable for investigating different anatomical structures, so privileging an anatomical level limits the number of techniques required to develop the best explanation of behaviour. If Sherringtonian neuroscientists are correct, then we must undertake the gargantuan task of mapping out the microcircuits of the cortex with cell-level techniques to achieve the best explanation of sophisticated behaviour. This is a weak solution to the problem of abstraction. But if Hopfieldian neuroscientists are correct, then we don’t need to do all that: we can more quickly map out the “representational space” of any given brain region using region-level techniques to achieve the best explanation of behaviour. This is a much stronger solution to the problem of abstraction.

I mentioned in the introduction that solving the problem of abstraction by privileging certain levels for explanation (or even certain types of explanations) entails that (a) intra-level computations may be included in, but (b) inter-level computations must be excluded from the best explanation of behaviour. The deeper problem with this solution, though, is that it stipulates that there is a perfect coincidence between the anatomical organization of the brain and the functional organization of cognition, such that all computations that should be included in the best explanation of behaviour are conveniently located at one anatomical level (i.e., belong to one anatomical kind). But surely, this relationship should be the object of empirical analysis—not the object of stipulation for a convenient solution to the problem of abstraction. And moreover, the empirical evidence indicates that this structure-function relationship is much more complicated than perfect coincidence (e.g., Honey et al., 2010; Poldrack et al., 2011; Cao, 2012, 2014; Klien, 2012; Anderson, 2014).

So, the anatomical solution for the problem of abstraction fails, yet cognitive neuroscientists have still managed to develop complete (and almost complete) explanations of certain behaviours. What else are they doing when they do manage to solve the problem of abstraction? To answer this question, let’s (a) consider a nearly complete explanation of behaviour, (b) identify the distinctive properties of computations that are included in the explanation, and then (c) speculatively infer that having these properties is necessary and sufficient for computations to be included in the best explanation of any behaviour. The result should be a much more plausible solution to the problem of abstraction. Then we can double back and consider the empirical implications of this abstract explanation of behaviour for the structure-function relationship.

---

3 We’ll rationally reconstruct their tacit, anatomical conception of levels in §4.
Figure 1. Cartoon of the complete FASL system for reference. Figures 2–5 depict different sections of this cartoon.
§2. Frontal Azimuthal Sound Localization

I propose that we review one of the most impressive behavioural circuits in cognitive neuroscience: the frontal azimuthal sound localization (FASL) system in the barn owl (*Tyto alba*). The input for the basic FASL system consists of a speaker playing a sound at any position from the left to the right (i.e., the frontal azimuth) of the barn owl as they face forward (Knudsen et al., 1979). The correct output consists of the barn owl saccading (or orienting) their head to the frontal azimuth of the speaker. The FASL system is an effective example for an abstract explanation because (a) it is a feedforward network, (b) it only uses simple arithmetic computations, (c) it has relatively few layers (11), and yet (d) it crosses relatively many levels (8) for behavioural circuit with so few layers (Fig. 1). In this section, I’ll present the explanation in one smooth, detailed pass.

The input to the FASL system is task information about the location of the sound source in the frontal azimuth. Unfortunately, this information doesn’t come encoded in a stimulus format, so the Layer 1 of the FASL system re-encodes this information into a stimulus format: into an interaural (between-ears) time delay (ITD) between the left and right ears. For example, if a sound originates

---

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).
5 I’ll clearly distinguish layers and levels in §4: roughly, layers are just the nodes of the feedforward network and the levels are the anatomical kinds that the nodes fall under.
6 Uninterested readers may skip to §§3–4, where I focus on individuating the layers and levels.
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 barn owl’s head and reaches their left ear 66 μsec later. In other words, then, the geometry of the barn owl’s head will convert the sound source information into an ITD format (Fig. 2). Since the ITD format can be encoded by the ears, it counts as a stimulus format.

Most sound waves are composed of multiple frequency components (Fig. 3A) and the FASL system 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 in the left and right ears at an ITD. Each basilar papilla is tapered, such that its apical end is narrower but thicker (and stiffer) and its basal end is broader but thinner (and less stiff). Since higher-frequency sound waves have more energy than lower-energy ones, the higher-frequency components concentrate at the apical end (where their higher energy quickly dissipates) and lower-frequency components concentrate at the basal ends. Thus, a frequency map forms along the length of each basilar papilla at an ITD during stimulation from the sound wave.

Since neurons communicate with electrical signals, the FASL system requires task information to be transduced: to be encoded in an electrical format. Layer 3 satisfies this need for transduction by lining hair cells along the length of each basilar papilla. Each peak in a sound wave causes the hair cells on each basilar papilla to collide into the adjacent tectorial membrane and each collision

---

**Figure 3.** (A) Every sound wave is composed of single-frequency components. (B) The soundwave enters each ear and strikes each basilar papilla. Each basilar papilla separates out the two tones and hair cells transduce the signals in separate frequency channels along the length of the basilar papillae. These frequency channels are preserved by the auditory nerve fibers, which use spikes to encode the timing of peaks in the membrane potentials of hair cells.
causes hair-like fibers (stereocilia) on the hair cells to bend against the membrane, opening ion channels and depolarizing the hair cells. This encodes phase information from each sound wave in the continuous fluctuations of the membrane potential (hair cells don’t generate action potentials). Moreover, hair cells transduce unique positions on the frequency map, such that the frequency map of the basilar papillae is copied over by the population of hair cells (Fig. 3).

Since digital processing (the use of discrete action potentials) is much more energy-efficient than analogue processing (the use of continuous membrane potentials), the FASL system requires task information to be digitized: encoded with action potentials. Layer 4 satisfies this need by having hair cells 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 hair cells, creating “frequency channels” that remain until Layer 7 (hence, I won’t mention them further till our discussion of Layer 7).
Auditory nerve fibers 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 system doesn’t need amplitude information, such that encoding amplitude information with the rate code prevents it from encoding task-relevant information. So, the FASL system requires this information to be discarded. Layer 4 is known to satisfy this need but how exactly this is achieved isn’t fully understood yet. What we do know that it is achieved between auditory nerve fibers (ANFs) and the nucleus magnocellularis (NM) (Sullivan & Konishi, 1984). Meanwhile, the NM also re-encode phase information from the ANFs (Köppl, 1997a).

Encoding and re-encoding task information in an ITD format requires the temporal resolution of a few microseconds, which is costly for the FASL system and thereby creates the need for a cheaper, non-temporal format (see footnote 4). Layer 6 solves this problem using one of the most impressive circuits known to neuroscience. Each NM cells sends axonal projections to each 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 right NL, makes synapses on several NL cells as it passes straight through the right NL, and ends at the dorsal surface of the NL (Fig. 4; Takahashi & Konishi, 1988).

The axonal projections from each 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 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 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. Thus, the NL will encode two crossing events: one which corresponds to the real ITD (the true position in the frontal azimuth) and one which corresponds to a “phantom” ITD (a false position 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 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 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 system 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). Thus, the strongest signal is the true signal and any weaker signals are phantom signals. Then the ICX uses inhibitory interneurons to enhance the true signal across frequency channels (primary peak facilitation) and diminish 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

The 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. Layer 8 satisfies this functional need by

7 Unless the sound is a tone (which contains only one frequency component), the mode signal across frequency channels will always be the true signal. Saberi et al. (1999) confirmed that the FASL systems cannot resolve phase ambiguities in tones: the barn owls saccade to the hypothetical sources of the phantom IPDs.
using the optic tectum (OT), which copies the frontal azimuthal map from the ICX, to send convergent projections through the medial efferent pathway to synapse onto premotor neurons in several 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) multisensory map into a set of four coordinates (up, down, right, left), which are place-coded and so interpreted as a head-centered motor map (Masino & Knudsen, 1990, 1993). The values of these coordinates are rate-coded to build excitatory drive: larger values are encoded by higher spiking rates in these nuclei (Masino & Knudsen, 1992, 1993).

Next, these spherical coordinates in the mTeg need to be translated into the specific groups of muscles that can drive the appropriate saccading behaviour. Layer 9 satisfies this need by using projections from the midbrain tegmentum (mTeg) down the tectospinal pathway, into the spine, and onto motor nuclei in the cervical spinal cord (Masino & Knudsen, 1992). These junctions translate the mTeg's place code into a new, motor-based place code: areas in the mTeg that encode the up, down, right, and left coordinates project to the complex arrays of motor nuclei that encode motor primitives for upward, downward, rightward, and leftward head movement, respectively (Thoroughman & Shadmehr, 2000; Flash & Hochner, 2005; Alessandro et al., 2013; Giszter, 2015). The mTeg's rate code is preserved—ensuring that muscles contract enough to drive the direction of gaze far enough, into the appropriate position in the frontal azimuth.8

As the signal exits the nervous system, it’s encoded in both a place code (by the configuration of active motor nuclei), which is an analogue and spatial format, and a rate code (by the amount of activity in those motor nuclei), which is a digital and electric format. Since saccading behaviour is a continuous motion in space, the analogue, spatial format is appropriate for saccading behaviour but the digital, electric format isn’t. Thus, the FASL system needs to (a) reverse-digitize the format, by converting the signal from a digital format back into an analogue one, and (b) to reverse-transduce the format, by converting the signal from an electric format into a mechanical one. Layer 10 satisfies both needs at once: it encodes the signal in the analogue, mechanical contractions of up to 31 pairs of cervical muscles (Masino & Knudsen, 1990).

Finally, the FASL system needs to convert the muscular contractions into movement. Layer 11 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 each layer achieves its function (and if the sound contains more than one frequency component), then the FASL system will achieve its function by causing the barn owl’s head to successfully saccade to the frontal azimuth of the sound source.

§3. Cognitive Significance

Our explanation of the basic FASL system is abstract—i.e., it excludes a tremendous amount of computation involved in producing a behavioural response to the task. Moreover, it appears to do so as it should: the best explanation of behaviour in the FASL system hasn’t changed since the early 1990s. So, I propose that we extract its tacit solution to the problem of abstraction and treat

---

8 After all, further azimuthal positions (e.g. +30° vs. +10°) require more force to saccade the head to, so encoding the furthest azimuthal positions with the highest spiking rates allows for them to deliver more excitatory drive to the muscles, which makes it possible for the muscles to generate enough force to move the head to those positions.
that solution as provisionally normative: this is our best candidate for a correct solution to the problem of abstraction. We’ll find that it solves the problem of abstraction by identifying a rare compound property of computations that is necessary and sufficient for them to count as one of the 11 layers of the FASL system. I propose that we refer to this compound property as cognitive significance. We’ll distinguish it from other kinds of cognition in §5. The goal of this section is to identify what cognitive significance is and hence, to explicate the tacit rule that our explanation used to solve the problem of abstraction.

First, the most salient feature of our explanation is that each layer that we individuated in the FASL system functions to encode information about the frontal azimuth of the sound source. In other words: each layer functions to enter a state that is maximally correlated with the sound source being in the frontal azimuth that it is actually in (Dretske, 1981; Skyrms, 2010; Shea, 2018). This was true from input to output for the FASL system: even the barn owl’s behavioural response of saccading to +30° functions to encode the task information because it functions to increase the conditional probability that the sound wave is located at +30°. I propose that this generalizes:

**Cognitive significance (1):** if a computation C is cognitively significant for a task T and a cognitive system S, then C is a causal relation between two states of two components of S that encode information about the state of T.⁹

This encoding relation has been the main focus of cognitive neuroscientists: they look for (a) significant correlations between spiking activity in various units of the brain and task states and (b) significant second-order correlations between those first-order correlations (i.e., the amount of task information in particular units of the brain) and task performance. These inform them about the correlational and functional properties of this encoding relation, respectively.¹⁰

This first criterion shouldn’t be confused with informational teleosemantics, which is a theory about the nature of representation: that it reduces to informational functions (e.g., Neander, 2017; Shea, 2018; c.f. Dretske, 1981). I’m advocating a theory about the nature of processes that warrants their inclusion in explanations of behaviour: they must have informational functions, regardless of whether representations reduce to informational functions. Of course, friends of informational teleosemantics usually accept this further claim too because it explains how representations are relevant to explaining behaviour: e.g., Shea (2018) defends his informational teleosemantics by arguing that identifying the informational functions of cognitive processes is indispensable for generally explaining the flexible behavioural performance of task functions. But that’s orthogonal to my argument here. More on that in §5.

If a process having the function to encode task information were sufficient for it to be relevant to explaining behaviour, there would be no solution to the problem of abstraction. After all, there is a continuous flow of task information through any cognitive system: there is a continuous flow of causation through infinitely many states (from input to output) that all function to encode task information.

---

⁹ Shea (2018) proposes some modifications to this basic formula: that information must be exploitable, explanatory, and unmediated in its explanatory power. I’m happy to add those modifications.

¹⁰ Although this work is important, it has led some to believe that cognitive neuroscience is irrelevant to understanding a richer notion of cognition than information encoding and so that psychology is autonomous from neuroscience (e.g., Putnam, 1967; Fodor, 1974; Pylyshyn, 1980; Lycan, 1981; Cummins, 1983).
information. Thus, there are infinitely many possible ways to model any cognitive system as a discrete network of finitely many senders and receivers. So, I propose that identifying the coding of task information is necessary but insufficient for a process to warrant inclusion in behavioural explanations. A complete solution to the problem of abstraction will require us to identify another property of layers that is more distinctive than just having informational functions.

Second, another salient feature of our explanation in §1 is that each layer that we individuated in the FASL system functions to re-encode task information in a novel format that creates and/or satisfies a functional need for the FASL system that is created and/or satisfied by another, distally-related layer. For example, the basilar papillae (Layer 2) function to encode task information in a frequency-separated format and their doing so satisfies the need of the external nucleus of the inferior colliculus (ICX; Layer 7), which is five layers downstream, for a frequency-separated format so that the ICX can achieve its function to resolve phase ambiguity by finding the mode signal across frequency channels. I propose that this generalizes:

**Cognitive significance (2):** if a computation C is cognitively significant for a task T and a cognitive system S, then C is a causal relation between two states of two components of S that (a) encode information about the state of T and (b) create and/or satisfy functional needs for S that are satisfied and/or created by other, distally-related components of S.\(^\text{11}\)

This criterion includes other layers. For example, consider the distal functional relation between the head (Layer 1) and the cervical muscle contractions (Layer 10). The head converts the task signal from a spatial/movement format (the sound wave moving through space from its source) to a temporal format (an ITD) for the neural system. But that creates the need for the FASL system to convert the task signal back into a spatial/movement format: into a head saccade (a movement format) to the frontal azimuth (a spatial format) of the sound source. This need is satisfied ten layers downstream, by the cervical muscle contractions.

Since this criterion must advance our solution to the horizontal problem of abstraction, it is critical that it rules out the vast majority of computations that function to re-encode task information. As it turns out, the vast majority of such computations do not seem to create and/or satisfy functional needs for distally-related layers of the FASL system. For example, consider the ossicles (three small bones) in the barn owl’s middle ears. Their function is impedance matching: they maintain the energy of sound waves as they pass from the low-impedance air of the outer ear to the high-impedance fluid of the inner ear by using a hammer structure to concentrate the energy of the sound waves. In the context of the FASL system, the ossicles in the barn owl’s ears function to convert task information from an air-based, frequency-combined ITD format to a fluid-based, frequency-combined ITD format that is equally energetic (due to impedance matching).

The computation that ossicles function to perform is excluded from explanations of behavioural performance. The reason cannot be our first criterion: they do function to compute over input and output states that function to encode FASL information. The reason for this exclusion is our second criterion: they only stand in functional relations with proximal components of the FASL system.

\(^{11}\) There are multiple ways to quantify the conception of distance that’s involved: (a) the time that it takes a signal to reach one process from another, (b) the number of intermediate processes, or (c) the number of subsystems that they both belong to (e.g., whether they both belong to the sensory system).
That is, they satisfy the need of the basilar papillae (Layer 2) for a frequency-combined ITD format with sufficient energy (despite being immersed in the fluid of the inner ear) so that they can achieve their function to convert to a frequency-separated ITD format. However, the basilar papillae are proximally related to the ossicles: the ossicles deliver the soundwaves directly to the fluid that immerses the basilar papillae. As far as I can tell, the ossicles don’t create and/or satisfy functional needs that are satisfied and/or created by any distally-related components. If they were found to do so, though, then my proposal would predict that the ossicles should then be included in our best explanations of behaviour. Otherwise, my proposal would be false.

Our review in §2 suggests that all and only the layers in our explanation encode task information while standing in functional relations with distal layers. However, I should encourage the reader to question this seriously. After all, information about the FASL system is distributed across the literature: one cannot find an explanation in that literature as systematic and abstract as the one that I presented in §2, not even in review articles. So, I had to develop a solution to the problem of abstraction by trial-and-error to develop a unique discretization of the FASL system that includes all and only the information that seems relevant to explaining behaviour—given the way that cognitive neuroscientists talk about the FASL system. However, this admittedly involved a lot of judgment on my part. So, I encourage the critical reader to engage directly with the literature on the FASL system and question whether my explanation of the FASL system does seem to provide the most useful approach to abstraction over all that information. If they find that it doesn’t, my proposal should be considered for modification or rejection.

This second criterion shouldn’t be confused with teleosemantics either. Producer- and consumer-based teleosemantics are theories about the nature of representation: that representational content is constituted by certain functional need relations between senders (i.e., producers) and receivers (i.e., consumers) (Millikan, 1984; Neander, 1991; Papineau, 1991). Hence, they require cognitive systems to be partitioned into a unique network of producers and consumers prior to individuating representational content. Critics have argued that it probably isn’t possible to individuate a unique discrete network of finitely many producers and consumers prior to individuating informational functions (Cao, 2012, 2014; Shea, 2018). I agree: the first step in our solution to the problem of abstraction is to narrow down the computations that could feature in a behavioural explanation to those that have the function to encode task information (following Shea, 2018). Hence, my solution to the problem of abstraction contradicts the claim that informational functions are constituted by certain functional need relations between senders and receivers.

Recall again that the goal of my theory is not to reduce or ground representational content but to rule out computations that aren’t relevant to explaining behaviour. I’ve proposed that individuating informational functions allows us to rule out computations over states that don’t function to encode task information. This sidesteps Cao’s (2012, 2014) indeterminacy problems for those who would try to solve the problem before individuating informational functions. I’ve also argued that this won’t be enough: we’ll have to look at all the functional need relations between the remaining computations, sort the remaining computations into proximal and distal categories according to the functional need relations that they stand in, and then finally, build our explanation of behaviour.

---

12 Cognitive neuroscientists have not systematically identified the network of distal functional relations in the FASL system, so I predict that a layer or two may need to be added (or maybe even removed) from my discretization of this example (assuming my proposed solution to the horizontal problem is still correct).
by excluding computations from the proximal category and including computations from the distal category. My proposal is that what remains will be one discrete network model of senders and receivers that is uniquely relevant to explaining behaviour.

Therefore, third, I propose that these two criteria—(a) encoding functionally-required information and (b) standing in distal functional relations—are sufficient to individuate the computations that are cognitively significant:

**Cognitive significance (3):** a computation $C$ is cognitively significant for a task $T$ and a cognitive system $S$ if and only if (and because) $C$ is a causal relation between two states of two components of $S$ that (a) encode information about the state of $T$ and (b) create and/or satisfy functional needs for $S$ that are satisfied and/or created by other, distally-related components of $S$.

This is a bold commitment: for it to be true, there must exist a discrete network of distal functional relations among a relatively small set of processes in any cognitive system. Till now, cognitive neuroscience hasn’t looked for anything like such networks. Even when most of the information to reconstruct this network is available, as in the FASL system, cognitive neuroscientists haven’t tried to actually, systematically reconstruct it—not even to the incomplete extent that I did in §2. Future work should explore this.

Thus, I should also encourage the critical reader to question this commitment: to question whether the features of my explanation of the FASL system generalize, in any of three ways. First, they could show that there is no such network of distal functional relations, such that my proposal should be totally rejected. This would be the most surprising result. Second, they could show that the distinction between proximal and distal relations is so vague that there are multiple non-arbitrary ways to individuate networks of distal functional relations. Thus, there would be no unique solution to the problem of abstraction and my proposal would require further conditions to identify one. This would be the least surprising result. Third, they could show that the network of distal functional relations itself is so large that it raises a new problem of abstraction. Then my proposal would require further conditions to solve this new problem.

**§4. Structure-Function Relationship**

No conception of level serves any role in our solution to the problem of abstraction in §3. Given the tight relationship that is thought to hold between levels and abstraction, should we dispense with the concept of levels then? No, certainly not. As soon as we solve the problem of abstraction without privileging levels for explanation, we free up the concept of levels to extract patterns in the trajectory that task information is functionally required to take across the levels of a cognitive system. In other words, since our solution to the problem of abstraction takes the methodological role of abstraction away from levels, it frees them up so that they can serve a much more valuable empirical role instead. In particular, I’ll argue that levels are indispensable to the empirical analysis of task information flow through the cognitively significant computations of any cognitive system.

There are many conceptions of levels, but Craver (2007) notes that each faces a dilemma. Older conceptions entail heavyweight ontological claims beyond the claim that levels have meaningful
identities (e.g., Oppenheim & Putnam, 1956; Churchland & Sejnowski, 1992; Wimsatt, 1994). Newer conceptions maintain ontological neutrality yet entail that levels lack meaningful identities (e.g., Craver, 2007; Bechtel, 2008). So, I’ll develop a novel conception of level that maintain ontological neutrality yet entails that levels have meaningful identities. I’ll show that we need both for our empirical analysis but I’ll leave it to future work to defend this conception of levels against others. Disclaimers aside, I propose that we define levels as partitions into instances of anatomical kinds:

**Level identity:** L is the “level of K” if and only if (and because) L is a partition of some entity into parts that are all instances of the natural (often anatomical) kind K.

This conception of level justifies the common practice of naming levels in cognitive neuroscience: e.g., whenever cognitive neuroscientists talk about the level of brain regions, we can interpret them as identifying the partition of the brain (or some part of it that can be partitioned into brain regions) into parts that are all brain regions (an anatomical kind).

Moreover, it entails that any two anatomical-kind parts of a cognitive system can stand in three possible relations with each other. The first possibility is that they can belong to the same level:

**Same levels:** if two entities belong to the same level, then this is so because both are members of the partition of a third entity into instances of some natural kind.

For example, the NL and the mTeg belong to the same level (the level of unilateral nuclei) because they are both members of the partition of the nervous system into instances of unilateral nuclei. The second possibility is that two parts can belong to higher or lower levels:

**Higher level:** if an entity E₁ belongs to a higher level than another entity E₂, then this is so because (a) E₂ is a member of the partition of an entity E (which could be E₁) into instances of some natural kind and (b) E and E₁ are both members of the partition of another entity E₃ into instances of some other natural kind.

For example, the left NL and the bilateral NM belong to different levels because (a) the left NL is a member of the partition of the bilateral NL into instances of unilateral nuclei (i.e., the left NL and the right NL) and (b) the bilateral NL and the bilateral NM both belong to the partition of the nervous system into regions. The third possibility is that two parts can belong to adjacent levels:

**Adjacent levels:** if two entities belong to adjacent levels, then this is so because they don’t belong to the same level and neither belongs to a higher level than the other.

It just so happens that there are no examples of adjacent levels in the FASL system. More complex systems often have adjacent levels, though. For example, the spatial navigation system requires both spatial cues in the environment and place cells in the hippocampus to encode information about the task state. The spatial cues will be members of the partition of the environment into instances of some natural kind whereas the place cells will be members of the partition of the hippocampus into instances of some other kind. The result will be two adjacent levels: one that spans the environment and the other that spans the hippocampus (or the entire nervous system).
We can use this partition conception of level to develop a novel kind of empirical analysis: a *level plot*. First, we identify all the natural kinds that the cognitively significant processes in a cognitive system belong to. The input and output layers are states of the entire *extended system*: (a) the sound source in the environment emanates sounds at a certain position relative to the barn owl and (b) the barn owl saccades in response to a certain position relative to the sound source. Layers 1 and 11 compute states of the *entire organism*: (a) the geometry of the entire head (the result of its skeletal, muscular, and other organ systems) creates an interaural (between the ears) time delay and (b) the entire body (the result of its skeletal, muscular, and other organ systems) creates the saccading movement by resisting against the contraction of cervical muscles.

Layer 10 computes a state of an *entire organ system*: the contraction of cervical muscles. Layer 2 computes a state of an entire bilateral pair of organs: the basilar papillae (BP). Layer 9 computes a state of an *entire array of related bilateral nuclei*: the motor nuclei (MN). Layers 3, 4, and 5 compute states of *entire bilateral pairs of nuclei*: the hair cells (HC), auditory nerve fibers (ANFs), and nuclei magnocellularis (NM). Layers 6 and 8 compute states of an *entire unilateral nucleus*: the nucleus laminaris (NL) and the midbrain tegmentum (mTeg). Layer 7 computes states of an *entire unilateral nuclear locus*: loci in the external nucleus of the inferior colliculus (ICX). Unlike every other natural kind of part in the FASL system, unilateral nuclear loci aren’t anatomically distinct at all: they’re only functionally distinct vis-à-vis their connectivity with upstream sources and downstream targets.

So, we count 8 natural kinds of states that are computed by cognitively significant computations in the FASL system: states of (a) the extended system, (b) the organism, (c) the organ systems, (d)
the bilateral organs, (e) the bilateral nuclear arrays, (f) the bilateral nuclei, (g) the unilateral nuclei, and (h) the unilateral nuclear loci. Next, we assemble these natural kinds into levels by relating them to each other via partition relations. The extended system can be partitioned into the organism and their environment. The organism can be partitioned into organ systems. Some organ systems (e.g., the brain) can be partitioned into arrays of related bilateral nuclei. Those can be partitioned into bilateral nuclei, which can be partitioned into unilateral nuclei, which can be partitioned into unilateral nuclear loci. For each partition, we can define a corresponding level (Fig. 6).

Next, we can assemble the network of cognitive significant computations onto the levels, which happen to form a series since there are no adjacent levels. The result is a novel representation of the relation between function and anatomy: the level plot (Fig. 7). The level plot for the FASL system reveals an interesting feature of the relation between function and anatomy: task information gets monotonically deeper after each cognitively significant computation from the task input at the system level until the ICX and gets monotonically shallower until it reaches the behavioural output back at the system level. This triangular trajectory reveals a sharp distinction between sensory (which encodes task information into monotonically sparser formats) and motor (which encodes task information back into monotonically more distributed formats) processing.

Here’s one explanation of that interesting trajectory geometry. The premotor system is a noisy environment insofar as it’s a bottleneck that every signal from the various sensory systems must pass through to get encoded as a behavioural response. This creates a functional need for the rest of the cognitive system to encode the task signal in a format that is least vulnerable to noise. Sparse codes are less vulnerable to noise than distributed codes (e.g., Cayco-Gajic & Silver, 2019), so the sensory layers of the FASL system collectively satisfy this need by re-encoding the signal in a
progressively sparser format. Once the signal is encoded in a maximally sparse code, it’s sent from the sensory systems to the noisy environment of motor processing, where it can be re-encoded in progressively more distributed formats without destructive interference from other signals.

The level plot for the FASL system also reveals another interesting feature of the relation between function and anatomy: the ICX is an intermediate, sensorimotor layer insofar as it’s the completion of the sensory processing of task information into monotonically sparser formats and the initiation of the motor processing of task information back into monotonically more distributed formats. In a sense, the ICX constitutes an inflection point for the FASL system. Now, while the FASL system has a single inflection point, it raises a possibility: do some cognitive systems have more than one inflection point and if so, what would the function of having additional inflection points be? Also, is the unilateral nuclear locus the lowest-level structure that can contain an inflection point for a cognitive system? Or are there cellular or subcellular structures at even lower levels that can also contain inflection points for cognitive systems? These are interesting questions for future research.

Ironically, though, the most interesting thing about inflection points are the cognitively significant computations that are furthest from them. After all, cognitive neuroscientists have a practical bias towards studying states at or near inflection points: they encode task information in the sparsest formats, so their code is the easiest to record and then the easiest to decode. For example, the first region to be decoded in the FASL system was the optic tectum (Knudsen & Konishi, 1978), which encodes task information at the same level as the external nucleus of the inferior colliculus (ICX; Layer 7). Subsequent studies followed inputs to and outputs from the optic tectum upstream and downstream until the entire FASL system was identified. For more complex systems (e.g., spatial navigation systems), though, only the sparsest codes have been measured and decoded. Therefore, a complete understanding of a cognitive system will require us to compensate for this bias and specifically investigate the cognitively significant computations furthest from the inflection points.

One concern that we might have about our empirical analysis of the FASL system is that it won’t generalize to more difficult systems. After all, two contingent properties of the FASL system have significantly facilitated our empirical analysis: (a) it projects onto a series of levels since there are no adjacent levels and (b) it’s a single-stream feedforward network. But level plots generalize to systems without these two features. First, partition levels have the graphical structure of trees, so a series of partition levels is just a tree graph with one branch (e.g., Fig. 7) and a structure with adjacent levels is just a tree graph with multiple branches. Second, any network can be projected onto a graphical structure of trees, so projecting a single-stream feed-forward network onto a single-branch tree of partition levels is the simplest case but projecting a multi-stream feedback network onto a multi-branch tree of partition levels is equally possible.

Ultimately, a level plot represents the relation between behaviour and anatomy. After all, the network of cognitively significant computations is defined with respect to behaviour (§3) and it is projected onto the tree of partition levels, which is defined with respect to anatomy. The level plot

---

13 In fact, the optic tectum directly feeds from the external nucleus of the inferior colliculus (ICX; Layer 7) but it doesn’t count as cognitively significant because it encodes task information in a state that doesn’t create and/or satisfy any functional needs for other distally-related components of the FASL system.

14 There were a few exceptions to this trajectory of discovery: e.g., the encoding of task information in an ITD format (Layer 1) was identified soon after (Knudsen & Konishi, 1979).
reveals that this relation has an interesting structure, even for examples as simple as the basic FASL system. By comparison, those who claim that a single level is privileged for explanation underestimate the complexity of this relationship. Likewise, those who claim that any level can be privileged for explanation underestimate the complexity of this relationship in the opposite way. This counts in favour of our solution to the problem of abstraction then: by including and excluding computations from behavioural explanations independently of their relations to levels, it preserves the complexity of the relation between behaviour and anatomy.

§5. Cognitive Capacities

So, cognitive significance is indispensable for the best solution to the problem of abstraction and explaining the relations between anatomy and behaviour. It may be tempting to infer that cognitive significance is all there is to cognition. But I resist that conclusion: cognition isn’t just cognitive significance. After all, if cognitive neuroscience is just the neuroscience of cognition, then there is a sense in which cognition is just whatever cognitive neuroscience aims to study. If cognition is just cognitive significance, then what counts as cognition would be task-specific because what counts as cognitively significant is task-specific (§3). Then the goal of cognitive neuroscience would be to explain behavioural performance in task-specific ways on (arbitrary) sets of specific tasks. However, that’s false: although cognitive neuroscience does use specific tasks to investigate cognition, it aims to “kick away the ladder” and develop a unified, systematic theory of cognition that is task-general, not task-specific.

Rupert (2009) makes a related point in response to extended and embodied theories of cognition. He grants that states and processes that feature in the best explanation of any cognitive system’s response to a particular task extend out into the body and the environment. I would say the same: cognitive significance is extended (as in the input and output layers of the barn owl) and embodied (as in Layers 1, 2, and 10). However, he notes that cognitive scientists study cognition by swapping environments (as in different tasks and contexts) and even bodies (as in in vitro studies), such that the concept of cognition that is operative in the cognitive sciences must be task-general and even system-general. This affords us a different way to frame the same point: cognition isn’t just cognitive significance because cognitive significance is extended to and embodied in task-specific elements, whereas some kinds of cognition are restricted to organisms and their brains.

The distinction between cognitive significance and cognition provides an important concession towards extended and embodied theories of cognition, though, that Rupert (2009) and other critics overlook. After all, critics often suspect that there is a distinction between a modicum of cognition and as Rupert (2009: 41) puts it, a “pandemic” of computation. But we’ve seen that this distinction is misleading. While there is a pandemic of computation and a modicum of cognition, cognition isn’t the only important and relevant kind of computation to the cognitive sciences. There is a modicum of cognitively significant computations, and they form an important and relevant kind of computation too. This is an important concession to extended and embodied theories of cognition: even if they are inadequate theories of cognition in general, they may still be adequate theories of cognitive significance.

Cognitive significance and cognition are distinct kinds: either cognitive significance is a precursor to cognition or as I suspect, it is a kind of cognition. Either way, cognitive significance constrains
other kinds of cognition. After all, cognitive neuroscientists start by explaining performance on specific tasks and then graduate to explaining performance across tasks. The simplest way is via *abstraction*: we identify the cognitively significant computations that explain performance on a few tasks, we define the intrinsic properties of those computations (i.e., the algorithms that they use to return their outputs in response to their inputs) as cognitive, and we define their extrinsic properties (i.e., their task-specific informational functions and distal functional needs) as non-cognitive. Egan (2010, 2014) defends a view like this—except that she doesn’t recognize the problem of abstraction and so, counts any algorithm (cognitively significant or not) as cognitive.

Within a task, this abstraction approach defines computations as cognitive in a way that neglects what made them interesting in the first place. For example, what’s interesting and distinctive about the ICX is not that it computes the mode over its inputs: many feedback loops between excitatory and inhibitory neurons in the brain can achieve that. What’s interesting and distinctive about the ICX is that—unlike all those other loops—it’s embedded in the FASL system in such a way that its ordinary computation acquires the *cognitive significance* of resolving phase ambiguity across frequency channels. Across tasks, this abstraction approach defines computations as cognitive in the weakest way possible: as any computation that *could* achieve cognitive significance in some task. This does make cognition task-neutral in a sense, but at the cost of removing everything from cognition that makes cognitive significance useful and interesting.

The lesson from the abstraction approach is that we need to define cognition in a way that not only removes the interesting task-specific elements of cognitive significance but also replaces them with interesting elements that aren’t task specific. The way to do this is via *generalization*: we identify the cognitively significant computations that explain performance on a few tasks and we define cognition as a novel category that somehow generalizes over the task-specific elements of cognitive significance. This problem of generalization has been recognized in the literature on neural reuse: e.g., Anderson (2014) influentially argues that while cognitive neuroscience has made significant advances in explaining neural use for specific tasks, conceptual and explanatory innovations may be required to successfully explain neural reuse across tasks.

A few theories of neural reuse have been proposed: e.g., the neural exploitation hypothesis (Gallese & Lakoff, 2005), the shared circuits model (Hurley, 2005, 2008), the neuronal recycling hypothesis (Dehaene, 2005; Dehaene & Cohen, 2007), and Anderson’s (2014) own interactive differentiation and search framework. Philosophical work is needed to determine whether these theories contain a definition of cognition that fulfills its job description—that satisfies the need for generalization over cognitive significance. The notion of cognitive capacities may prove useful here: we may be able to say that cognitive capacities are task-general entities contained within an organism that get exercised in task-specific ways to perform specific tasks (Rupert, 2009). If this is so, then we could look for and possibly find useful philosophical conceptions of cognitive capacities in successful scientific theories of neural reuse.

Clearly, I don’t have the space to plausibly commit to theories of cognition, cognitive capacities, or neural reuse. I’m only claiming that cognitive significance imposes constraints on any adequate general theory of cognition: it must show that certain cognitive kinds (e.g., cognitive capacities) are indispensable for generalizing over cognitively significant computations across specific tasks. While theories of neural reuse have been especially sensitive to this constraint, other approaches
to cognitive ontology might stand to benefit from being more sensitive to them. After all, most approaches to cognitive ontology have emphasized individuating cognitive functions that can generalize over associations between specific brain areas and specific tasks (and their properties) (Price & Friston, 2005, 2007; Poldrack et al., 2011). However, cognitive significance involves more than associations with task properties—it also involves distal functional relations. Hence, adequate approaches to cognitive ontology should take care to generalize over distal functional relations too (see also Khalidi, 2017).

Once a theory of cognition does manage to generalize over cognitively significant computations, we should be able to apply that general explanation to performances on specific tasks. The resulting explanations will be different from explanations of cognitive significance. For example, suppose that cognitive capacities are indispensable for generalizing over cognitively significant processes, such that an adequate theory of cognition claims that cognition consists in capacities. Then the application of this general theory of cognition to performances on specific tasks will explain those performances as exercises of cognitive capacities. This is different from but compatible with our original task-specific explanations of behavioural performance: they will reinterpret task-specific cognitively significant processes as task-specific exercises of task-general cognitive capacities.

In a very real sense, then, cognitively significant processes are cognitive: if a system has cognitive capacities, then they will coincide with the exercises of those cognitive capacities. Still, they fall short of exhausting the domain of cognition: e.g., they are distinct from cognitive capacities. Also, they may extend beyond the domain of cognition: it’s entirely possible that cognitive significance may solve the problem of abstraction for other computational systems (e.g., signalling cascades in cells) that lack the flexibility to perform across task and hence, lack any kind of cognitive capacity. In such cases, cognitive significance wouldn’t coincide with the exercise of any kind of cognitive capacity. This would be another, very real sense in which cognitive significance isn’t cognitive at all. So, the jury is out on whether cognitive significance is best classified as a kind of cognition or as proto-cognition.15

§6. Conclusion

I’ve argued that there is too much computation in cognitive systems to include in our explanations of behavioural performance: this creates a problem of abstraction. I’ve argued that two criteria are necessary and sufficient to solve this problem of abstraction and suitably decide which processes to include in our behavioural explanations. The first criterion is that the processes must function to encode information about task states. Friends of informational teleosemantics would call these representations (Dretske, 1981; Neander, 2017; Shea, 2018), but our solution is consistent with the denial of that claim. The second criterion is that these encoding states must create and/or satisfy functional needs for the cognitive system that are satisfied and/or created, respectively, by distally-related components in the system. Overall, then, my proposal is that each cognitive system has a discrete network of cognitively significant computations that are sufficiently rare that an accurate model of this network will provide the best abstract explanation of behavioural performance.

15 I expect such outcomes: e.g., that we’ll find cognitive significance as we solve the problem of abstraction in our explanations of cellular behaviour. But I can’t commit to it in this paper: if cognitive significance does extend beyond cognitive systems, then I’d be happy to admit that cognition-neutral terminology might prove more appropriate.
I have mentioned in §4 that this conclusion has several implications for our philosophical theories of cognition, so I’ll conclude by showing that this conclusion also has methodological implications for cognitive neuroscience. That is, it suggests that there are at least four stages of investigation:

1. **Signal transmission**: trace the continuous flow of information through a cognitive system from the task input to the behavioural output (and modify tasks to manipulate this flow);
2. **Distal functional coordination**: discretize the continuous flow of information into a discrete network of cognitively significant computations, which stand in distal functional relations with each other (and modify tasks to manipulate the properties of this network);
3. **Inter-level dynamics**: plot the inter-level trajectory of information through the network of cognitively significant computations that are each nested at a partition level (and modify tasks to manipulate the geometry of this inter-level trajectory);
4. **Cognitive recruitment**: if cognitively significant computations are recruited by multiple different systems, then generalize over the ways that it is recruited to encode information about various tasks and participate in various networks of distal functional relations.

This list isn’t meant to be exhaustive, but it suffices to reveal a problem with research programs in cognitive neuroscience: they stop at the second stage. For example, research on the FASL system has significantly decreased since 1990s, when the functional flow of task information through the FASL system was fully mapped (completing the first stage) and major (anatomical) junctures in that flow (candidates for cognitively significant computations) were intuitively identified (starting the second stage). This is typical for behavioural circuits: cognitive neuroscientists don’t seem to know what to do with them after they understand signal transmission and have a preliminary hypothesis about functional coordination in the second stage. The concept of cognitive significance clarifies that there are at least three stages left before we completely understand cognition in these simple behavioural circuits.

**Acknowledgements**: I thank Sara Aronowitz for her feedback on several versions of this draft, which was invaluable for helping me put this interesting case of the FASL system toward some interesting philosophical conclusions. I thank Lindsey Crown, Daniel Gray, Adam Lester, and especially Sahana Srivathsa for workshopping a draft of this paper, our many hours of discussion about neuroscience (and philosophy) over the past few years, and making me feel at home in the neuroscience program at the University of Arizona. Without them, I might never have endeavoured to write a paper like this one. I thank the editors and anonymous reviewers at [TBA] for the best kind of feedback—the kind that helped me understand my own arguments better. Finally, I thank the audience at the 2021 Meeting for the Society for Philosophy and Psychology for their helpful questions on my presentation of this paper, some of which I hope to answer in future work.
References


