Annual Review of Neuroscience

Toward An Integrative Theory of Thalamic Function

Rajeev V. Rikhye,[1,2] Ralf D. Wimmer,[1,3] and Michael M. Halassa[1,2,3]

[1]Department of Brain and Cognitive Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA; email: mhalassa@mit.edu

[2]McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

[3]Stanley Center for Psychiatric Genetics, Broad Institute, Cambridge, Massachusetts 02139, USA

Annu. Rev. Neurosci. 2018. 41:163–83

The *Annual Review of Neuroscience* is online at neuro.annualreviews.org

https://doi.org/10.1146/annurev-neuro-080317-062144

Keywords

cognition, thalamus, neocortex, brain development, cortical microcircuits, predictive coding

Abstract

The thalamus has long been suspected to have an important role in cognition, yet recent theories have favored a more corticocentric view. According to this view, the thalamus is an excitatory feedforward relay to or between cortical regions, and cognitively relevant computations are exclusively cortical. Here, we review anatomical, physiological, and behavioral studies along evolutionary and theoretical dimensions, arguing for essential and unique thalamic computations in cognition. Considering their architectural features as well as their ability to initiate, sustain, and switch cortical activity, thalamic circuits appear uniquely suited for computing contextual signals that rapidly reconfigure task-relevant cortical representations. We introduce a framework that formalizes this notion, show its consistency with several findings, and discuss its prediction of thalamic roles in perceptual inference and behavioral flexibility. Overall, our framework emphasizes an expanded view of the thalamus in cognitive computations and provides a roadmap to test several of its theoretical and experimental predictions.

Contents

INTRODUCTION

Among the many marvels of evolution, the human brain stands out as an exceptional triumph. With a mass three times that of our closest primate relative, it has given rise to Sun Tzu's *Art of War*, Isaac Newton's *Principia Mathematica*, Wolfgang Amadeus Mozart's *Requiem*, and Steve Jobs's iPhone. When one contemplates the forces that drive human civilization, it becomes clear that our cognitive capacity, rather than sensory acuity or motor strength, is the secret to our species' success. As such, understanding which brain attributes are most relevant for cognition has been a major fascination of humanity, with pioneering investigations documented by Galen of Pergamum nearly two thousand years ago (Hellier 2014). Galen contended that the forebrain is associated with cognition, whereas the peripheral nerves are tied to lower-level motor output (Hellier 2014). He coined the term *thalamus* to refer to the brain's inner chamber that connected the animal spirit with the cerebral nerves. His thalamus, however, is most likely what we would call the third ventricle today, rather than the group of diencephalic gray matter nuclei we collectively refer to by that name (Jones 2007).

Despite his error, Galen's intuition about the forebrain being tied most closely to cognitive capacity has withstood the test of time. Here, we operationally define cognition as the flexible control of sensorimotor transformations, which determine the capacity and range of adaptive behavioral output (Miller & Cohen 2001). The relationship between forebrain evolution and cognitive capacity is evident when comparing how skillfully a frog can catch a fly, with its inflexibility in modifying that behavior depending on the kind of fly it is catching or where the hunt takes place (Matesz et al. 2014, Ramsay et al. 2013). This type of behavioral flexibility, however, is evident in birds and mammals (Emery & Clayton 2004), which unlike their common reptilian ancestor (Striedter 2006) have an expanded forebrain. In fact, the intrinsic architecture of the midbrain and hindbrain is conserved across the vertebrate lineage, with the progressive expansion of the forebrain being the biggest differentiator (Striedter 2006) (**Figure 1*a***), suggesting a relationship between forebrain size and cognitive capacity.

But what is it about the forebrain that gives rise to cognitive capacity? The mammalian forebrain is composed of the telencephalon (cortex and basal ganglia) and diencephalon (thalamus and hypothalamus). The hypothalamus is conserved across the vertebrate lineage, with little expansion or variation throughout evolution (Xie & Dorsky 2017). Conversely, the cortex, basal ganglia, and thalamus show substantial expansion and variation (Butler & Hodos 2005) (**Figure 1*b***); therefore, they are thought to be more closely tied to cognitive capacity (Lefebvre et al. 2004). From the perspective of circuits in the brainstem and spinal cord, forebrain output would adaptively modify their sensorimotor transformations, providing a basis for behavioral flexibility.

The contribution of the forebrain to behavioral flexibility is nonuniform, dictated by distinct regional architectures that impose local computational constraints. Specifically, both parts of the

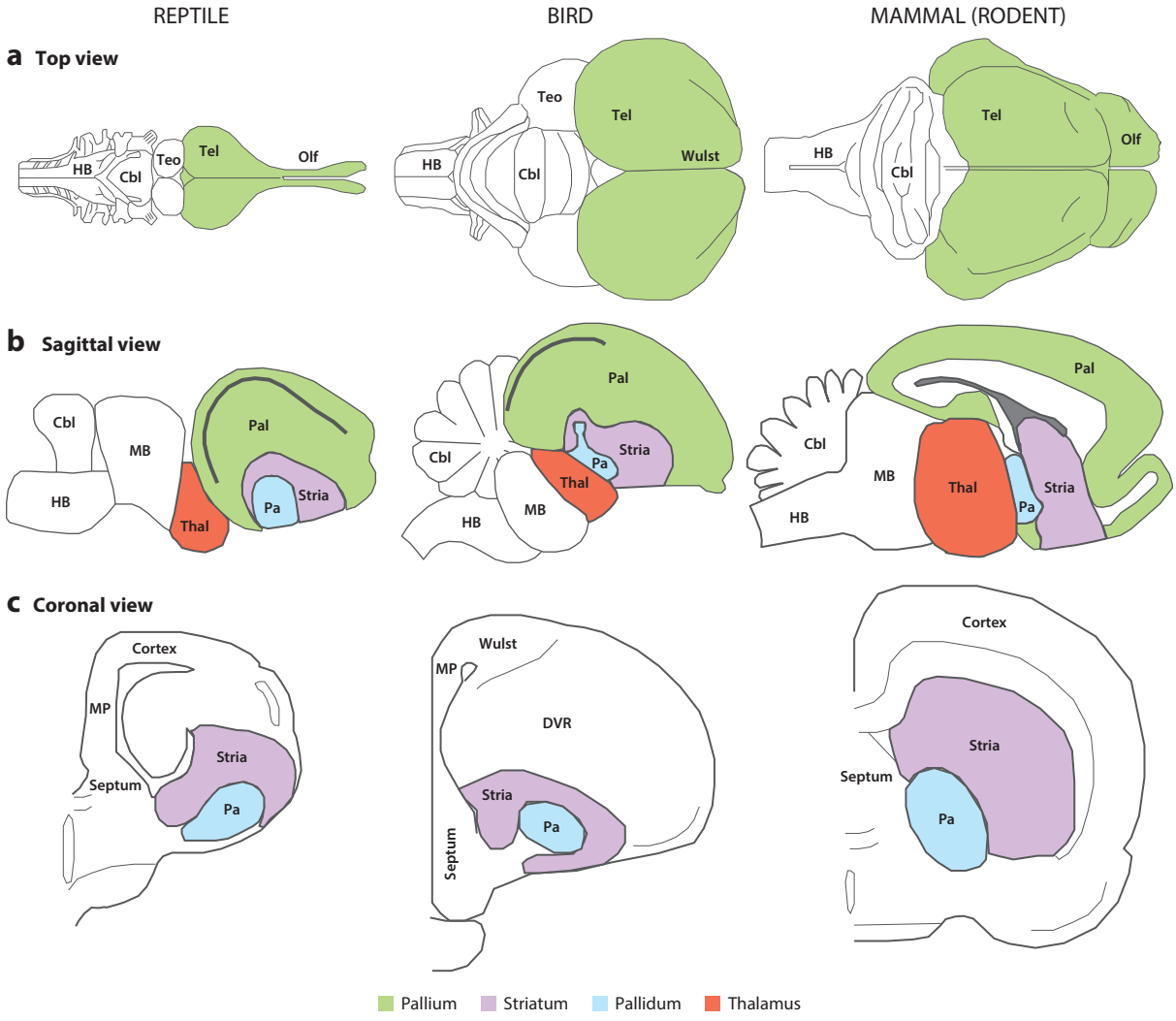


Figure 1

Comparable organization of adult telencephalon of reptiles, birds, and mammals (rodent). (*a*) Top view, (*b*) sagittal view, and (*c*) coronal view. Note the expansion in the size of the pallium (Pal), striatum (Stria), and thalamus (Thal) from reptile to mammal, which correlates with an increase in cognitive capability. Adapted with permission from Jarvis et al. (2005) and Kalman (2009). Abbreviations: Cbl, cerebellum; DVR; dorsal ventricular ridge; HB, habenula; MB, midbrain; MP, medial pallium; Olf, olfactory bulb; Pa, pallidum; Tel, telencephalon; Teo, optic tectum. Wulst and hyperpallium are synonymous.

telencephalon—the cortex and the basal ganglia—exhibit local recurrent connections, with the former dominated by excitatory connections and the latter by inhibitory ones (Bolam & Bevan 2001, Douglas & Martin 2004). By contrast, the thalamus is composed primarily of excitatory neurons that show no local recurrent connections (Jones 2007). These distinct anatomical features lead to differential capacities to generate and maintain local internal states over time. By coupling these states to motor outputs in the brainstem and spinal cord, the telencephalon can adaptively modulate behavior (Bastos et al. 2012, Harris & Shepherd 2015).

So, does the thalamus contribute to cognition? The answer is an unequivocal yes, and here we review experimental data that support this notion. Our view substantially diverges from the dominant view of the thalamus as a relay of sensory information to the cerebral cortex (Guillery & Sherman 2002). As we explain below, this view is derived from how a highly specialized thalamic circuit, the lateral geniculate nucleus (LGN), appears to operate in visual processing. Within this feedforward model of vision, the LGN serves as a relay of retinal inputs to the striate cortex, and behaviorally relevant object features are extracted through a hierarchically organized set of computations implemented by extrastriate areas (DiCarlo et al. 2012, Felleman & Van Essen 1991). This process is ultimately thought to result in highly specialized cortical responses to particular objects (e.g., faces, cars, and landscapes) (Orban 2008, Yamins & DiCarlo 2016). As we argue in subsequent sections, this purely feedforward cortico-cortical model is unable to account for the flexibility by which objects are categorized or how an object is inferred when sensory inputs are ambiguous (den Ouden et al. 2012, Lee & Mumford 2003, Rohe & Noppeney 2015). In addition, the diversity of input-output patterns across the mammalian thalamus makes many of its circuits unsuitable for simply relaying information. Therefore, the LGN may be the exception rather than the rule, with the rest of the thalamus exhibiting diverse functions relevant to cognition.

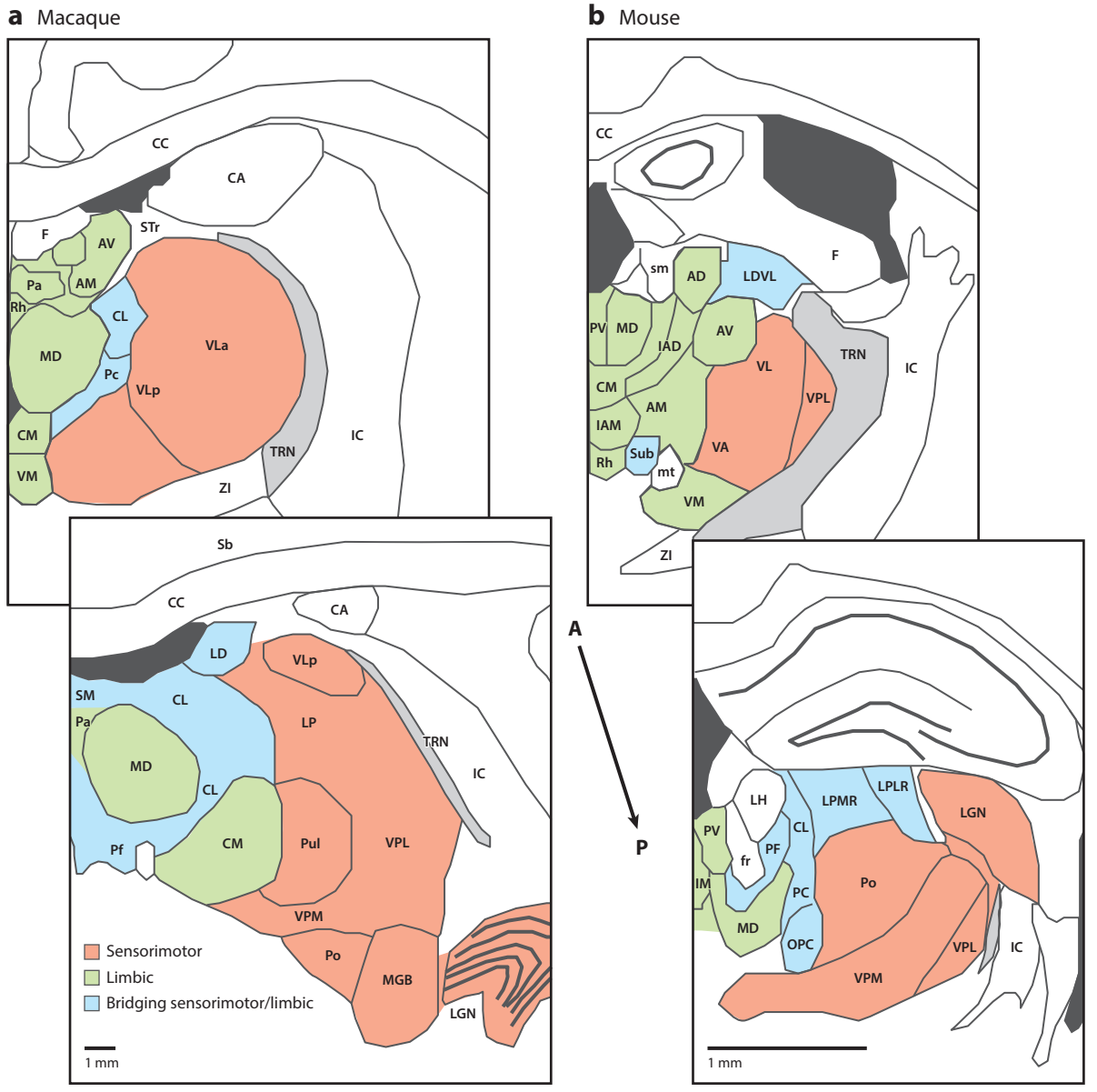
STRUCTURAL DETERMINANTS OF THALAMIC COMPUTATIONS: AN EVOLUTIONARY PERSPECTIVE

In mammals, the thalamus is grossly divided into four sections: an anterior group (anteromedial, anterolateral, anterodorsal, and laterodorsal nuclei); a medial division composed of the midline group (paratenial, paraventricular, and centromedian), an intralaminar group (centrolateral, centromedial, and parafascicular), and the mediodorsal nucleus (MD); a lateral division containing the ventroanterior/ventrolateral (VL) group, the ventrobasal (VB) complex, and the ventromedial nucleus; and a posterior group containing the posteromedial nucleus, the lateral posterior nucleus, the pulvinar, and the medial geniculate body (MGB) and lateral geniculate nuclei (LGN) (**Figure 2**). These divisions are based on a combination of gross anatomical features such as their relation to the internal medullary lamina and their cytoarchitectural staining (Bold et al. 1984). Dividing the thalamus in this manner may not generate corresponding functional divisions; the lack of thalamic lateral connectivity coupled with the diversity of inputs and outputs seen within each nucleus suggests that function is nonhomogeneous along a nuclear dimension. Other classification schemes have grouped nuclei on the basis of a predominant type of input (sensory, motor, or limbic) (Vertes et al. 2015), relative input origin and strength (first order versus higher order), or output type (core versus matrix) (Jones 1998, 2002).

In addition to the thalamus's unique structural features, its lack of single-cell connectivity data makes it difficult to find dimensions along which thalamic functions are clustered (Phillips et al. 2017). Until more experimental data are available, we propose that a reasonable way to identify functional groups is to focus on individual thalamic microcircuits, each of which is composed of a group of neurons that share common input features (origin, strength, and degree of convergence), output features (destination, strength, and degree of divergence), and inhibition features (origin and strength). Therefore, individual nuclei may contain multiple types of thalamic microcircuits, with functional parallels potentially arising in a distributed manner across several nuclei (see **Figure 3**).

So, what could the most functionally relevant anatomical features be? We think that an evolutionary perspective is useful here because it provides functional constraints. In reptiles, the thalamus does not receive substantial descending projections from the cortex (Pritz 1995). Instead, it provides two types of ascending telencephalic inputs: one from the roof of the midbrain

(collothalamic nuclei) mainly to the striatum, and another from the retina (LGN homolog), trigeminal nucleus, and lateral spinothalamic tract (called lemnothalamic nuclei; mainly the nucleus dorsolateralis) to the dorsal cortex (Butler 2008, 2009; Molnár & Butler 2002) (**Figure 3*a***). Parallels with mammalian thalamic organization can be broadly observed. For example, the LP/pulvinar and the LGN have collothalamic and lemnothalamic origins, respectively (Grant et al. 2012). Interestingly, the MGB and LGN, which are generally classified together as core, sensory, or first-order nuclei, do not appear to share a common evolutionary origin.



(*Caption appears on following page*)

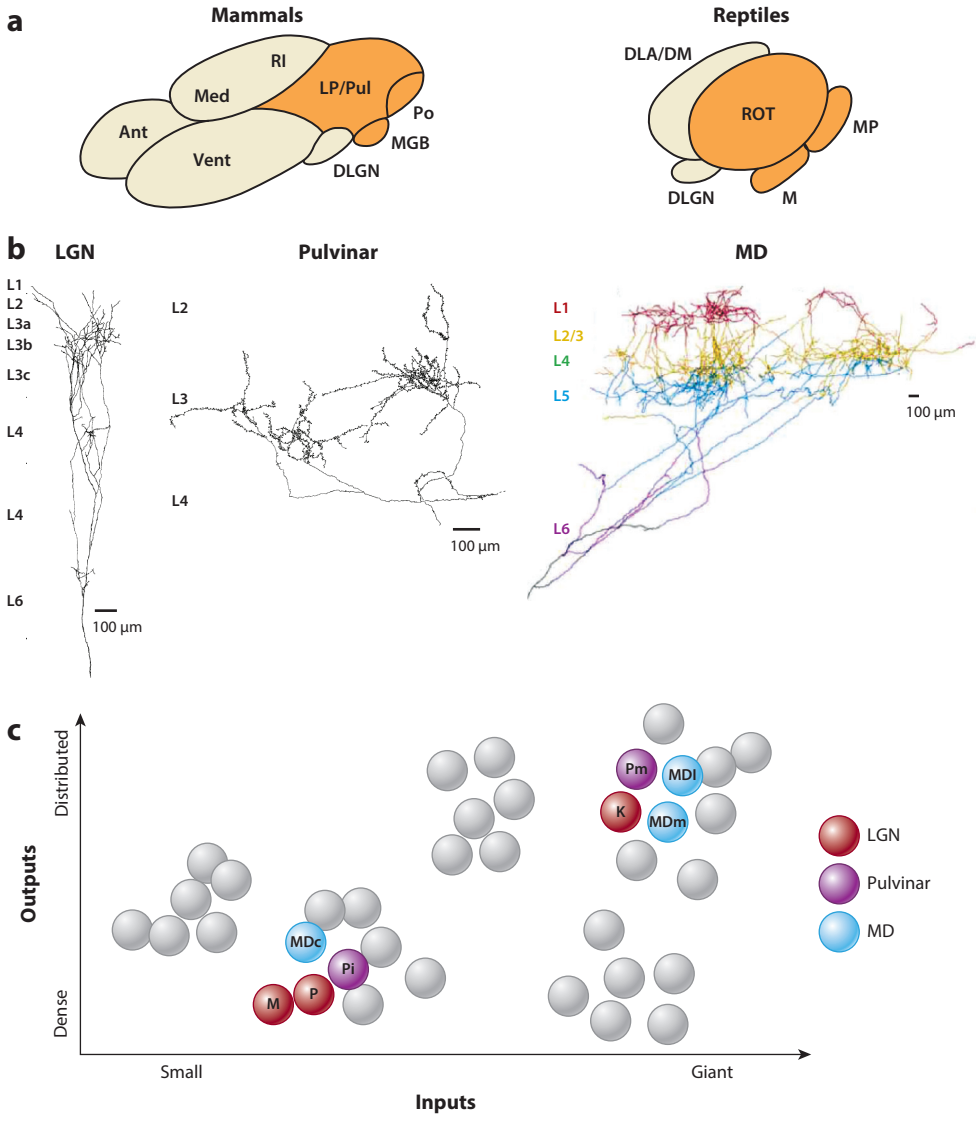
Figure 2 (*Figure appears on preceding page*)

Organization of the thalamic areas of the (*a*) macaque (adapted from **http://brainmaps.org/index.php**) and (*b*) mouse [redrawn from Paxinos & Franklin (2004) with permission from Academic Press]. Areas are color coded according to their connections with the sensory, motor, and limbic systems (Vertes et al. 2015). Abbreviations: A, anterior; AD, anterodorsal; AM, anteromedial; AV, anteroventral; CA, caudate nucleus; CC, corpus callosum; CL, centrolateral; CM, centromedial; F, fornix of hippocampus; fr, fasciculus retroflexus; IAD, interanterodorsal thalamic nucleus; IAM, interanteromedial thalamic nucleus; IC, internal capsule; LD, laterodorsal; LDVL, ventromedial part of the laterodorsal nucleus; LGN, lateral geniculate nucleus; LH, lateral habenula; LP, lateroposterior; LPLR, laterorostral part of the lateral posterior nucleus; LPMR, mediorostral part of the lateral posterior nucleus; M, medial; MD, mediodorsal nucleus; MGB, medial geniculate body; mt, mammillothalamic tract; OPC, oval paracentral thalamic nucleus; P, posterior; Pa, paraventricular nucleus; Pc, paracentral nucleus; PD, posterodorsal nucleus of medial geniculate complex; Pf, parafascicular thalamic nucleus; Po, posteromedial; Pul, pulvinar; Rh, rhomboid nucleus; sm, paraventricular nucleus of the thalamus; STr, strial terminalis; Sub, submedius thalamic nucleus; TRN, thalamic reticular nucleus; VA, ventroanterior; VL, ventrolateral; VLa, anterior ventrolateral; VLp, posterior ventrolateral; VM, ventromedial; VPL, ventral posterolateral nucleus; ZI, zona incerta.

An important change that occurred during the transition from reptiles to mammals is the rise of not only a six-layered cortex but also functional specialization and topographical maps across the cortex (Montagnini & Treves 2003). The reptilian medial, dorsal, and lateral cortices are multimodal areas (Naumann et al. 2015) with weak responses to visual inputs that are devoid of retinotopy (Aboitiz et al. 2002). Therefore, the notion that the reptilian thalamus relays sensory inputs to its cortex, as the mammalian geniculate does to the primary visual cortex (V1), is unlikely to be accurate. Instead, it appears that the reptilian thalamus may provide a modulatory cortical input informative of broad changes in a sensorimotor context. On the basis of the structural diversity of mammalian thalamic axonal outputs (**Figure 3*b***), similar thalamic effects on cortical function in the mammalian brain are likely to be widespread.

Another evolutionary adaptation that accompanied the appearance of a specialized mammalian sensory cortex is a change in the density of thalamic terminals in the cortex. These terminals can range from being highly restricted to a patch in layer 4 (L4) (e.g., LGN axons in V1) (Usrey et al. 1992) to being highly diffuse, covering large swaths of multiple cortical areas (e.g., MD axons across L2/3) (Kuramoto et al. 2017). Multiple intermediate varieties, which provide an individual thalamic neuron the ability to exert combinatorial effects on a single or multiple cortical targets, have also been observed. Thalamostriatal projections have similar diversity (Unzai et al. 2017). The dense and focal projections characteristic of certain thalamic neurons (e.g., magno- and parvocellular neurons of the primate LGN) (Blasdel & Lund 1983, Livingstone & Hubel 1988) are likely more recent evolutionary adaptations that function to convey specialized sensory information, whereas diffuse thalamocortical terminals convey more abstract contextual information. Additionally, another critical evolutionary specialization in mammals is the invasion of the thalamus by descending cortical input, a feature absent in reptiles (Pritz 1995). These terminals arise from corticothalamic neurons located within the two major infragranular layers (L5 and L6), which exhibit divergent features at multiple scales. For instance, whereas all thalamic territories receive L6 inputs, only a subset receives L5 inputs (Rouiller & Welker 2000). In fact, the group of nuclei collectively called primary sensory (LGN, MGB, VB) is entirely devoid of L5 inputs. Although inputs from L6 are numerically abundant, they utilize small terminals (<1 μm) to impinge on the distal dendrites of thalamic neurons, predominantly through single synaptic contacts (Crandall et al. 2015). As a consequence, L6 inputs onto thalamic neurons, which have been studied mostly in sensory systems, have minimal roles in shaping thalamic neuronal receptive fields (RFs). In contrast, L5 inputs utilize terminals of varying sizes, some are similar to the >4-μm giant subcortical inputs (e.g., retinal inputs to LGN) with multiple synaptic contacts, while others approach the size of L6 terminals. These smaller L5 terminals exhibit a higher number of mitochondria and synaptic vesicles than do L6 terminals, suggesting that they are more effective at supporting synaptic transmission onto their thalamic

targets (Rovo et al. 2012). We hypothesize that the diversity of L5 inputs enables thalamic RF diversity; by exhibiting little convergence, giant L5 inputs would impart cortical-like RFs on some neurons, as experimentally observed in areas of the posteromedial and lateral pulvinar (Chalupa & Werner 2003, Curry 1972, Zhou et al. 2016) where somatotopy and retinotopy are preserved. In contrast, smaller sized terminals exhibiting higher convergence would endow the thalamic neurons with unique RFs. This difference in convergence may explain the observation that pulvinar neural spike rates reflect confidence by which a perceptual judgement is made (Komura et al. 2013) rather than the perceptual category reflected by their parietal cortical targets (Kiani & Shadlen 2009). In addition, this difference would explain the contextual signals observed in some pulvinar circuits (the context would be a thalamic computation that involves convolving multiple cortical signals



(*Caption appears on following page*)

Figure 3 (*Figure appears on preceding page*)

Input-output diversity of thalamic neurons gives rise to multiple thalamic circuits. (*a*) Schematic comparison of lemnothalamic (*yellow*) and collothalamic (*orange*) regions in mammals and reptiles. Note the massive expansion of the lemnothalamic regions in the mammalian brain. Adapted from Butler (2009) with permission from Springer. (*b*) Thalamic neurons have diverse patterns of terminal axonal arborization in the cortex. Within the visual cortex, LGN neurons have dense, focal terminals in the upper layers of the visual cortex (Usrey et al. 1992). By contrast, pulvinar and MD neurons have multiple focal terminals [adapted from Rockland et al. (1999) with permission from the authors] and diffuse axonal terminals throughout the cortex [adapted from Kuramoto et al. (2017) with permission from Wiley]. (*c*) Thalamic circuits can be constructed on the basis of the nature of their inputs (e.g., corticothalamic terminal sizes), the pattern of their axonal terminals in either the cortex or the striatum, and the source of the inhibition they receive. In this way, the classical nuclear definition of the thalamus (*colored circles*) gives way to a more diverse classification in which each nucleus can have circuits that exert different computations on the cortex. We highlight the LGN [koniocellular (K), parvocellular (P), and magnocellular (M) pathways], pulvinar [medial (Pm) and inferior (Pi)], and MD [central (MDc), medial (MDm), and lateral (MDl)] neurons, whose input-output anatomy have been well characterized. Abbreviations: Ant, anterior nuclear group; DLA, dorsolateral anterior nuclei; DLGN, dorsal lateral geniculate nucleus; DM, dorsal medial nuclei; LGN, lateral geniculate nucleus; LP, lateroposterior nucleus; MD, mediodorsal nucleus; Med, medial nuclear group; MGB, medial geniculate body; MP, nucleus medialis posterior; Po, posteromedial; Pul, pulvinar; RI, rostral intralaminar nuclei; Rot, nucleus rotundus; Vent, ventral nuclear group.

impinging on an individual neuron) (Roth et al. 2012). Convergence may also explain why MD neurons temporally tile delay periods of working memory and could be the reason why such representation lacks categorical tuning found in their prefrontal cortex (PFC) targets (Schmitt et al. 2017).

Inhibitory inputs to the thalamus are almost as diverse as excitation and may form another axis of defining mammalian thalamic function. All thalamic territories receive inputs from the thalamic reticular nucleus (TRN). A subset of these territories also receive inhibition from an extrathalamic system composed of a group of inhibitory forebrain and midbrain nuclei (including basal ganglia, ventral pallidum, zona incerta, anterior pretectum, and pontine reticular formation) (reviewed in more detail by Halassa & Acsady 2016). Unlike reticular inputs that control overall rate, these extrathalamic nuclei exert a powerful inhibitory impact that likely controls thalamic spike timing.

Altogether these studies suggest that different thalamic nuclei receive anatomically different sources of excitatory or inhibitory inputs depending on the computations they perform. We propose that thalamic circuits can be constructed with different combinations of the anatomical features discussed above (**Figure 3*c***). For example, the primary sensory thalamus receives neither cortical L5 nor extrathalamic system inputs. The combinatorial nature of these features makes the thalamus a mosaic-like structure that may endow it with highly diverse functional properties that generate a variety of computational state variables capable of affecting cortical function. Therefore, we stress that a functionally meaningful classification of thalamic circuits is unlikely based on a nuclear organization but rather on clusters along the structural features (**Figure 3*c***). Efforts to map thalamic connectivity at the single-cell resolution hopefully aid in this quest (Lerner et al. 2016, Paul et al. 2017, Phillips et al. 2017).

PHYSIOLOGICAL DETERMINANTS OF THALAMIC COMPUTATIONS: SPIKING AND SYNAPTIC PROPERTIES

In addition to the anatomical features outlined in the previous section, thalamic neurons have physiological characteristics that can be classified into spiking and synaptic features. The best-known spiking feature of thalamic neurons is their ability to produce high-frequency (>100 Hz) sodium spikes (burst discharges) driven by a slow depolarization caused by calcium influx from

low-voltage-activated T-type calcium channels (Jahnsen & Llinás 1984). These T-type calcium channels inactivate rapidly at more depolarized membrane potentials, permitting neurons to encode inputs as either bursts or tonic spikes depending on their membrane potential. In particular, inputs arriving when the membrane is hyperpolarized trigger a burst of sodium spikes. In contrast, inputs arriving at more depolarized potentials, where T-type calcium channels are inactive, generate tonic sodium spikes (Zhan et al. 2000).

In the sensory thalamus, burst spiking was originally detected in low arousal states, such as sleep and anesthesia (Steriade et al. 1993), but numerous studies across a variety of species have shown that burst spiking could be an efficient way to relay information during alert wakefulness (Ramcharan et al. 2000, Sherman 2001). In fact, the frequency of burst spiking within the LGN varies with statistics of the visual stimulus (Lesica & Stanley 2004). Furthermore, the timing of the burst onset, the number of spikes in a burst, and the duration of the burst can each carry distinct sensory information (Butts et al. 2010, Gaudry & Reinagel 2008, Lesica & Stanley 2004, Reinagel & Reid 2000, Reinagel et al. 1999). Hence, burst-tonic spiking may be an effective method for thalamic neurons to multiplex and transmit information to the cortex. In support of this idea, studies have shown that burst spiking rectifies signals more strongly than tonic spiking does (owing to the nonlinearity of the calcium spike) and results in a higher signal-to-noise ratio, improving stimulus detectability (Guido et al. 1995, Whitmire et al. 2016). By contrast, tonic spiking preserves response linearity and permits linear summation of inputs, which in the LGN, is crucial for preserving RF structure from retinal ganglion cells. Taken together, these studies propose that by switching between burst and tonic spiking modes, the sensory thalamus controls not only how much but also what type of information is conveyed to the cortex.

Thalamic bursting is limited not only to the sensory system. A recent study of the motor thalamus, the ventrolateral nucleus (VL), found that inputs from the medial globus pallidus (GPm) of the basal ganglia induced burst spiking in the VL by activating T-type calcium channels (Kim et al. 2017). It has been proposed that projections from the GPm to the VL terminate actions requiring locomotion; therefore, bursts in the VL induced by the GPm are decoded by the motor cortex as a stop signal (Kim et al. 2017). Similarly, bursting in the MD suppresses the extinction of expressed fear in mice (Lee et al. 2011). This finding is consistent with the notion that bursts encode information in an all-or-none manner and can have a stronger impact on their targets than tonic spiking can (Sherman 2001).

If the different thalamic firing modes do indeed multiplex different forms of information, then the cortex must be able to decode and separate this information (Mease et al. 2017). This ability must depend on the biophysical properties of the thalamocortical synapse. In particular, these synapses can exhibit either short-term synaptic facilitation or short-term depression, which enhance or diminish the postsynaptic effect of thalamic bursting, respectively (Fuhrmann et al. 2002, Tsodyks et al. 1998). Therefore, cortical decoders, which take advantage of these different forms of plasticity, should be able to demultiplex information contained in thalamic bursts (Naud & Sprekeler 2017). Another possibility is that the cortex decodes thalamic inputs through the recruitment of different inhibitory mechanisms (Berger et al. 2010). For example, parvalbumin-expressing cells receive depressing synaptic inputs (Karnani et al. 2016), whereas somatostatin-expressing Martinotti cells receive facilitating inputs (Fino & Yuste 2011, Lee & Huguenard 2011). As a result, thalamic bursts may recruit Martinotti cells and tonic spikes may recruit parvalbumin-expressing cells, which would in turn have different effects on cortical computations (Tremblay et al. 2016).

Altogether these anatomical and biophysical properties suggest that the thalamus could be a modular building block in which specific thalamic computations arise as a result of three factors: the type of input that thalamic neurons receive, how these inputs are translated into burst and tonic spiking, and how spikes from these thalamic neurons are then decoded in the cortex by different

thalamocortical synapses. This modularity is especially beneficial for cognition, as it confers the thalamus with the ability to perform flexible computations that adjust to changing behavioral and contextual demands.

A PATH TOWARD A THEORY OF THALAMIC FUNCTION

Our goal for a theory of thalamic function is to explain the numerous micro- and macroscopic neural observations, both within the thalamus and across its interactions with other regions. This theory should also intuitively answer the question, What does the thalamus do? Such theoretical goals are distinct because they take different perspectives, one neural and one cognitive. For example, explaining the meaning of incoming thalamic spikes from the perspective of a cortical neuron does not necessarily provide a satisfactory answer for their role in recognizing a face or inferring another person's intention. Achieving both goals requires a bridge between neural hardware and cognitive outputs (thoughts and behaviors), which can be achieved through an intermediate description—the neural algorithm (Peebles & Cooper 2015). Therefore, we think elucidating the neural algorithms implemented by thalamic circuits is of utmost importance toward our understanding of thalamic functions.

So, what is an example of a neural algorithm? One well-studied example is visual object recognition. Primates can recognize objects regardless of color, size, luminance, background, or viewing angle. This remarkable capacity for identification is posited to be an untangling of a specific object manifold in a high-dimensional representational space (DiCarlo & Cox 2007), which at lower levels of processing (e.g., the retina) would be completely entangled with many other objects represented as raw pixel information (**Figure 4*a***). This computationally challenging problem can be efficiently solved by making the simple assumption that complex nonlinear transformations can arise from simple computations applied in a series (DiCarlo et al. 2012) (**Figure 4*b***). In support of such a model, the Gabor wavelet RFs of V1 neurons can be formed from pooled center-surround LGN RFs spanning a particular orientation (De Valois & De Valois 1980, Hubel & Wiesel 1962). Therefore, this simple function is thought to underlie the ability to detect edges in a luminance-invariant manner. With the use of similar transformations, and based on the notion that the ventral visual stream is hierarchically organized (Felleman & Van Essen 1991), object recognition can arise through a series of convolution operations followed by output nonlinearity, such as divisive normalization, that render visual representations increasingly tolerant to changes in viewing angle and therefore linearly decodable by object identity classifiers (Yamins & DiCarlo 2016) (**Figure 4*b***). Variations of these hierarchical convolutional neural networks (HCNNs) have been effective at predicting neural responses not only in the ventral stream (Cadieu et al. 2014) but also in the auditory system (Yamins & DiCarlo 2016). In fact, this algorithmic structure has been fundamental for computational vision, and recent advances in coupling artificial HCNNs with more efficient learning algorithms have given rise to the revolution of machines that are almost on par with humans in their ability to recognize objects (Hassabis et al. 2017, LeCun et al. 2015).

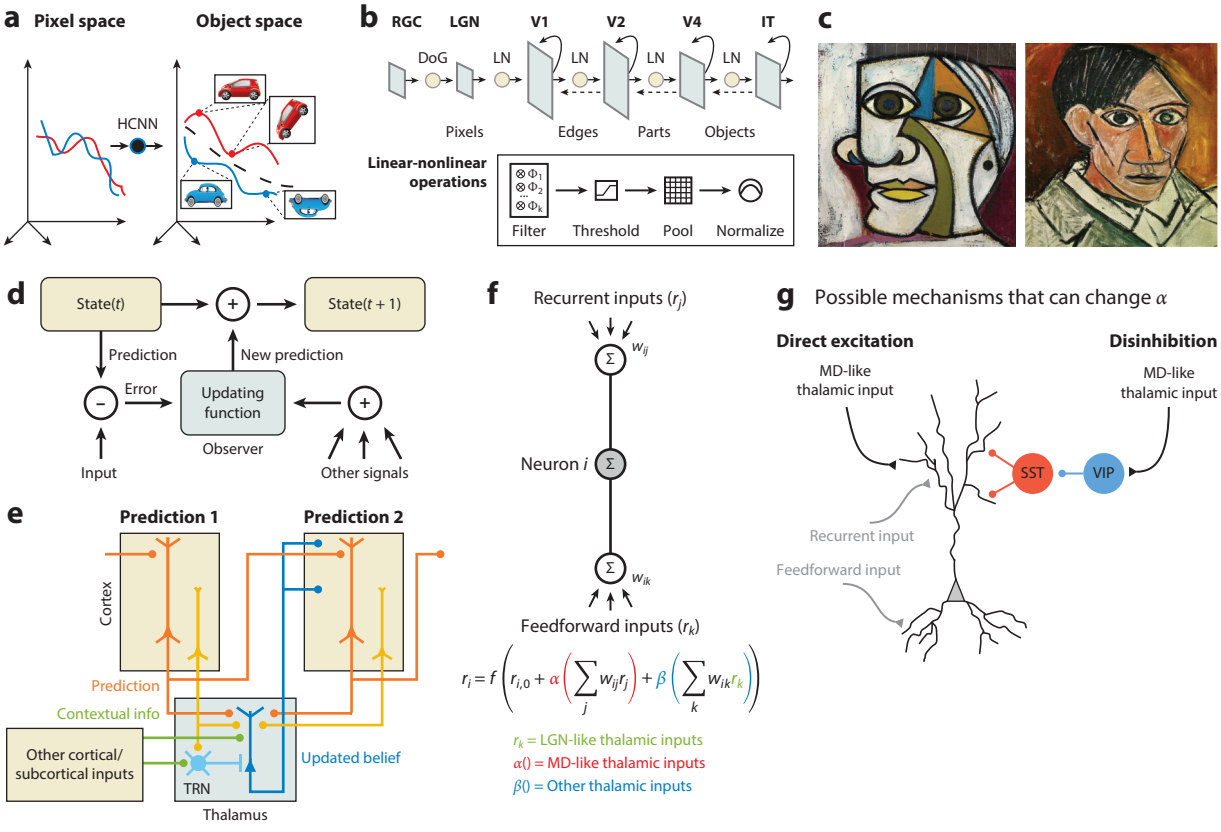
Despite such success, these deep HCNNs perform poorly in conditions of perceptual ambiguity, where meaning has to be inferred from either a broader visual context or associated nonvisual cues that depend on memory (Kok et al. 2013). An increasingly influential framework of how the brain deals with and adjusts to these conditions of ambiguity is predictive coding, which proposes that the brain actively tries to predict or infer what its sensory inputs are rather than passively registering them (Friston & Kiebel 2009). The ability to perform this active inference is a hallmark of intelligence (Griffiths & Tenenbaum 2006). Predictive coding relies on an internal model, which endows the brain with the ability to simulate the physical (or social) world, describing the dynamic relationship between behaviorally relevant objects (or events). Such an ability allows

the brain to generalize known action-outcome relationships to ones that had not been previously learned (Johnson-Laird 2010, Lee 2015, Shadlen & Shohamy 2016, Tolman 1948).

In essence, the internal model provides a mapping of how causes (x) generate sensations (A), and by simply inverting this model, via Bayes's theorem (Equation 1), the brain is able to infer the most likely causes of its sensory inputs (known as the likelihood function).

$$\underbrace{p(A|x)}_{\text{posterior probability}} = \frac{p(x|A)p(A)}{p(x)} \propto \underbrace{p(A)}_{\text{prior probability}} \underbrace{L(x;A)}_{\text{Likelihood function}} \quad 1.$$

Within this Bayesian framework, neural activity encodes beliefs (or posterior probability distributions) over states of the environment that cause sensation, and predictive coding provides a framework to reduce the uncertainty inherent in the sensory information that the brain samples. In particular, when information is ambiguous, existing belief states have to be tracked and the internal model dynamically updated with newly accumulated evidence (Haefner et al. 2016). This can be achieved with a Bayesian observer model, or a predictive Kalman filter, which estimates a new belief state given the sensory or environmental evidence and the current state (Jordan & Rumelhart 1992, Wolpert & Miall 1996). These new state estimates can then be used to generate new predictions, and this process continues recursively to produce better predictions with each new piece of sensory evidence. A central component of a Bayesian observer is its ability to create an estimate of the stimulus distribution (Mlynarski & Hermundstad 2017), which it derives from



(Caption appears on following page)

Figure 4 (*Figure appears on preceding page*)

Neuronal algorithms and computational architectures. (*a*) Object recognition can be thought of as untangling two object manifolds (*red* and *blue lines*, respectively). Here, in pixel space (i.e., raw pixel statistics), both red car and blue car manifolds are highly intertwined and cannot be easily separated. Processing through a cascade of LN transformations separates these object manifolds such that a classifier (*dashed black line*) can easily discriminate between the two objects in a viewing-angle-invariant manner. (*b*) Hierarchical organization of different visual processing stages. Processing along the ventral visual stream can be described as a cascade of LN transformations, starting with raw pixel information in RGCs, with each cortical area acting as an independent layer. As information ascends the hierarchy, increasingly more complex features, such as edges, contours, parts of objects, and eventually the object itself, are extracted from the image. Information not only flows forward but is also fed back (*dashed arrow*) to previous stages. These LN transformations that are applied between stages are relatively stereotyped and involve filtering with a linear filter (of increasing complexity) followed by thresholding, pooling, and normalization across neurons in that layer. Adapted from Yamins & DiCarlo (2016) with permission from Springer Nature. (*c*) A Picasso face (*left*), when taken out of the context of a Pablo Picasso painting, is highly irregular compared with what most people consider a normal-looking face (*right*). (*d*) An observer model (Kalman filter) updates state variables using prediction errors. (*e*) A biological implementation of a Kalman filter involving feedback interactions between the cortex, the TRN, and the thalamus. (*f*) Schematic of a typical cortical neuron whose firing rate (r_i) can be expressed as a weighted sum of inputs arriving at two different compartments. α and β are nonlinear gain functions that can selectively alter the integration of each set of inputs (e.g., recurrent versus feedforward inputs in this case), while the function f represents the nonlinear spike generation function. We propose that unlike LGN-like inputs, which can serve as a rate parameter (r_k) that drives and specifies the RF of this neuron under certain conditions, MD-like thalamic inputs selectively alter the excitability of an entire compartment that receives intracortical recurrent connections (the α term in the weighted sum equation). Other thalamic inputs may function in a manner similar to the MD-like inputs but on a different compartment, thereby changing the selectivity of the neuron from being driven by locally recurrent inputs to others originating elsewhere (the β term). In this way, different thalamic inputs can have contrasting effects on which inputs are used by a cortical neuron to construct its own RF. For our formulation to fully make sense, the α and β parameters would have to be linked in some manner; however, this is beyond the scope of this review. (*g*) Schematic illustrating the potential mechanisms by which MD-like thalamic inputs alter the gain of recurrent inputs. First, these thalamic inputs can alter dendritic integration by providing branch-specific excitation or by altering plasticity rules in a compartment-specific manner (Losonczy et al. 2008). Second, these thalamic inputs could also recruit different subtypes of inhibitory neurons, such as vasoactive intestinal peptide–expressing (VIP), resulting in branch-specific disinhibition via somatostatin-expressing (SST) neurons (Muñoz et al. 2017). Abbreviations: DoG, difference-of-Gaussians operation; HCNN, hierarchical convolutional neural network; IT, inferotemporal area; L, lateral; LGN, lateral geniculate nucleus; LN, linear–nonlinear operation; MD, mediodorsal nucleus; RGC, retinal ganglion cell; TRN, thalamic reticular nucleus; V1–4, visual cortical areas 1–4.

the greater context in which that stimulus occurs (see sidebar titled Thalamic Neurons Perform Forward Inference to Update Belief Functions for details). Within a typical context, elements (objects or episodes) can be related by specific spatial or temporal relationships (Bar 2004). For example, if one is viewing real human faces at a coffee shop, the visual context would make it highly surprising to find an ear between two eyes, resulting in large prediction errors. However, such an observation would be less surprising when viewing paintings by Pablo Picasso (**Figure 4*c***). In this framework, visual context (or gist, as it is more commonly called) defines a likelihood function that relates the spatial relationship between the typical subparts of a face, which varies depending on whether a natural face or a face painted by Picasso is viewed (Biederman et al. 1982). In contrast, when crossing the street in the United States, one automatically looks to their left, whereas in the United Kingdom one instinctively looks to their right. Here, the context is based on explicitly taught rules that depend on social norms rather than on the statistical structure of the physical world. Therefore, by having explicit representation of such contextual frames, a Bayesian observer can reduce uncertainty in the predictions of sensory inputs by recursively updating the beliefs the brain has of these inputs (Bar & Ullman 1996, Oliva & Torralba 2007).

How can we relate these processes to the role of the thalamus in cognitive function? The Bayesian observer represents a biologically plausible scheme for updating cortical beliefs (or expectations) about the environment with sensory samples. In particular, systems that encode or

THALAMIC NEURONS PERFORM FORWARD INFERENCE TO UPDATE BELIEF FUNCTIONS

Assume that at some time t and in some environment or context (θ_t), a thalamic neuron gets input from a cortical neuron (firing rate: y_t) in response to some stimulus (x_t). Assume also that this neuron has access to history of the past estimates of the environment ($\hat{\theta}_t$), which it uses to construct its belief of the current state of the environment. In a dynamic environment, for example, θ_t changes stochastically over time; hence, past estimates ($\hat{\theta}_t$) provide an estimate of how these changes occur. Using this information, the thalamic neuron can then compute the likelihood of this response conditioned on the current stimulus by marginalizing over past stimuli ($\hat{x}$) as follows:

$$\underbrace{p(y_t|x_t,\hat{\theta}_t)}_{\text{Likelihood}} = \int \underbrace{p(y_t|x_t,\hat{x},\hat{\theta}_t)}_{\text{Stimulus}\rightarrow\text{Response mapping}} \underbrace{p(\hat{x}|\hat{\theta}_t)}_{\text{History}} d\hat{x}. \quad 1.$$

Using Bayes's rule, one can infer the identity of this stimulus from the response and past estimates as follows:

$$\underbrace{p(x_t|y_t,\ \hat{\theta}_t)}_{\text{inferred stimulus}} = \frac{\overbrace{p(y_t|x_t,\hat{\theta}_t)}^{\text{Likelihood}}\ \overbrace{p(x_t|\hat{\theta}_t)}^{\text{Stimulus prior}}}{p(y_t|\hat{\theta}_t)}. \quad 2.$$

Working as an optimal Bayesian observer, a thalamic neuron can then use this inferred stimulus identity and accumulated cortical neural responses ($\hat{y}_{t-1}$) to compute its belief about the current state of the environment. This can be written as the following posterior probability:

$$\underbrace{p(\theta_t\,|\widehat{y}_{t-1},\widehat{\theta_t})}_{\substack{\text{Posterior of environment}\\ \text{(estimated state)}}} \sim \int \underbrace{p(x_t\,|y_t,\widehat{\theta_t})}_{\text{Inferred stimulus}}\ \underbrace{p(x_t|\theta_t)}_{\text{Stimulus probabilty}}\ \underbrace{p(\theta_t|\widehat{y}_{t-1},\widehat{\theta}_{t-1})dx_t}_{\text{Previous posterior}}, \quad 3.$$

$$\underbrace{p(\theta_{t+1}|\hat{y}_t,\hat{\theta}_t)}_{\substack{\text{Updated posterior}\\ \text{(new state)}}} = \int \underbrace{p(\theta_{t+1}|\theta_t)}_{\text{Environmental dynamics}}\ \underbrace{p(\theta_t|\hat{y}_t,\hat{\theta}_t)}_{\substack{\text{Previous posterior}\\ \text{(estimated state)}}}\ d\theta_t. \quad 4.$$

Therefore, with access to the posterior distribution and a model of environmental dynamics, thalamic neurons can update this posterior distribution to form predictions of the future state of the environment. This information, when propagated back to the cortex, will be crucial for predictive coding, in particular the computation of prediction errors between expected and unexpected stimuli.

mediate a contextual updating of belief states must receive convergent projections from large regions of the cortex, specifically from areas involved in encoding prediction errors (e.g., the frontal areas), reciprocate divergent projections back to these areas, and mediate some form of gain control over these areas (Kanai et al. 2015). As we have described above, the thalamus, particularly the pulvinar and the MD, meets all these requirements. By integrating inputs from many different cortical areas (Mitchell & Chakraborty 2013, Shipp 2004), these thalamic circuits can compute likelihood functions on the basis of a broader cortical context (**Figure 4*d*,*e***). In turn, the outputs from these thalamic nuclei can update cortical posterior probabilities depending on the regions they innervate by altering either synaptic gain (**Figure 4*f***) or functional connectivity between neurons in that region (O'Reilly et al. 2017, Purushothaman et al. 2012, Schmitt et al. 2017).

Therefore, as a Bayesian observer, the thalamus would have a central role in predictive coding by encoding expectations and thereby reducing the uncertainty of cortical predictions.

Recent neurophysiological studies indicate that the pulvinar indeed fulfills its role as a Bayesian observer. In contrast to the LGN, whose RFs are the building blocks for V1 RFs, pulvinar neurons have broad RFs that are sensitive not only to luminance and contrast changes but also to task-relevant variables (such as attention or decision confidence) (Komura et al. 2013). Pulvinar neurons project diffusely to many higher-order visual areas and prefrontal cortical areas, and LGN projections are highly retinotopically organized (Seabrook et al. 2017). Owing to this diffuse connectivity and large RFs, the pulvinar likely integrates over a much larger visual space than the LGN does and therefore is optimized for conveying coarse contextual information rather than structural visual information (McCotter et al. 2005, Oliva & Torralba 2006). Consistent with this notion, the rodent analog of the pulvinar, the lateral posterior (LP) nucleus, conveys information about running speed, specifically mismatches between optic flow and motor speed (Roth et al. 2016). These mismatches are akin to updated belief states about the flow of visual information. In this way, the rodent LP could provide a common sensorimotor context in which stimuli appear to V1, higher visual areas, and the cingulate cortex, priming them to detect unexpected stimuli (Fiser et al. 2016). Furthermore, the result that pulvinar neurons are sensitive to the reliability of perceptual decisions (Komura et al. 2013) suggests that they are specialized in encoding the expected precision or confidence in the information used for perceptual decisions.

An important point is that unlike V1 RFs, which are optimized for feature extraction, neurons within higher visual areas that receive pulvinar inputs have dynamic RFs that are optimized for task-relevant functions other than sensory identification. For example, neurons within the lateral intraparietal cortex (LIP) or middle temporal (MT) neurons have RFs that remap depending on attentional demands (Marino & Mazer 2016). Therefore, circuits within the pulvinar and elsewhere across the thalamus likely contribute to either computing or updating posterior distributions (about both the stimulus and the environment) rather than to computing observations themselves.

Additional evidence for the role of the thalamus in cortical state updating can be seen in planning goal-directed saccadic eye movements. Saccadic eye movements engage a broad network of both cortical and subcortical structures (Schall & Thompson 1999) and have a critical role in visuospatial attention (Bisley & Goldberg 2003, Ibos et al. 2013). The superior colliculus (SC), located in the midbrain, has a critical role in generating saccadic eye movements. In addition to the motor areas of the midbrain and the pons, which drive extraocular muscles, the SC also projects to the MD neurons and the pulvinar (Sommer & Wurtz 2008). These SC-recipient MD neurons in turn project to the frontal eye fields (FEF), a region of the frontal cortex involved in visuospatial attention and eye movement planning. Using a combination of electrophysiology and pharmacological inactivation, Sommer & Wurtz (2004a,b) demonstrated that SC-recipient MD neurons carry an efferent copy of the motor signal, known as the corollary discharge. This finding suggests that the MD contains an internal representation of the planned saccade direction, which the FEF then uses to move its RFs. Formalizing this according to our framework discussed above and in the sidebar titled Thalamic Neurons Perform Forward Inference to Update Belief Functions, if θ_{t+1} is the position of the eye after the planned saccade, then MD neurons can use the history of the previous saccade ($p(\theta_{t+1}|\theta_t)$) to update its belief of the location of the visual field (the posterior distribution; see Equation 4 in the sidebar). The idea is that the MD moves FEF RFs preemptively to mitigate the sudden shift in the visual field that would be caused by a saccade, and in doing so ensures stable vision (Sommer & Wurtz 2008). In support of this hypothesis, blocking corollary discharge signals from reaching the FEF by inactivating the MD prevented the FEF RFs from remapping to the location of the new saccade (Shin & Sommer 2012). Furthermore, successful saccade planning depends on the behavioral context in which these saccade targets

occur (Wyder et al. 2004). Similar to the pulvinar, the MD is highly sensitive to context and thus perfectly positioned to multiplex motor commands from the SC with contextual information about behavioral goals (Mitchell & Chakraborty 2013). How does this predictive remapping occur in the FEF? On the basis of recent results (Schmitt et al. 2017), MD neurons might alter the effective functional connectivity between specific ensembles of FEF neurons, which in turn would allow them to remap their RFs by integrating corollary discharge with visual input (Rao et al. 2016). As we explain in more detail below, this change in functional connectivity could be achieved by a selectivity gating mechanism, where MD-like inputs enhance the effect of certain recurrent inputs (**Figure 4*f*,*g***).

Contextual modulation of cortical processing applies to nonmotor cognitive functions as well. In particular, mapping a cue to the correct rule (e.g., looking to the left or to the right at a pedestrian crossing) depends on the context in which the cue occurs (United States versus United Kingdom). Unlike sensory context, this is a more abstract context that is taught and develops with experience. Evidence that the thalamus is involved in this type of contextual control can be seen in the interactions between the MD and PFC of the mouse during working memory. For example, when holding an item, such as a cue that informs the mouse of a rule, in working memory, PFC neurons show persistent activity at the level of the population, where each neuron exhibits enhanced spiking at a brief moment in time during this working memory delay (i.e., a temporal RF). The sequential progression of these temporal RFs across the population reflects the algorithmic requirement of keeping in mind a particular task-relevant categorical representation. Because individual PFC neurons categorically tuned to the same item would form a categorical sequence, one neuron's RF is determined by the spiking of another PFC neuron with an RF that is temporally ahead in the sequence. As such, the mechanistic finding that local effective PFC connectivity is enhanced during working memory is not surprising. What is surprising, however, is that the source of this enhancement is the MD thalamus, and that individual MD neurons of the mouse are also characterized by the lack of categorical tuning to the item. MD neurons are instead sensitive to the task context and show firing rate modulation that scales with whether an animal is engaged in the task or is in a different environment (Schmitt et al. 2017). Artificially activating MD is sufficient for algorithmically engaging PFC neurons outside the task but does not influence their task-relevant categorical preferences. In this scenario, MD neurons could update the belief state of tuned PFC neurons, increasing their confidence of the associating a cue with a particular rule.

Model-based inference relies on precisely tuned interactions between sensory information and prior information. In particular, feedforward computations extract features and feedback computations synthesize expectations based on an internal model (Bastos et al. 2012, Rao & Ballard 1999). The anatomy and laminar specificity of reciprocal connections between the thalamus and the cortex, which we have described in the preceding section, fit in well with the computational architecture implied by predictive coding. In particular, L5 neurons, which integrate inputs from many cortical sources (Harris & Mrsic-Flogel 2013), could provide a prediction signal to the pulvinar or the MD, which it then uses to update the posterior distribution as described in the sidebar titled Thalamic Neurons Perform Forward Inference to Update Belief Functions. Furthermore, by acting through the TRN, input from other corticothalamic cells could either enhance or suppress these thalamic predictions depending on behavioral demand or brain state (**Figure 4*d***). Thalamic signals to the cortex could act in either an LGN-like manner by providing inputs that directly drive cortical neurons to spike or an MD-like manner by enhancing the impact of other inputs (e.g., locally recurrent ones) via selective gating mechanisms, such as compartment-specific excitation or disinhibition (**Figure 4*f*,*g***). Disinhibition may be achieved by recruiting vasoactive intestinal peptide–expressing neurons that inhibit somatostatin-expressing neurons, which in turn

would enhance the excitability of pyramidal neurons in a compartment-specific manner (Muñoz et al. 2017) (**Figure 4*f*,*g***). More generally, this compartment-specific gain control mechanism could allow thalamic inputs to gate specific inputs to pyramidal neurons, allowing them to change their circuit-wide algorithmic engagement depending on dynamically changing behavioral goals. For example, pulvinar inputs to supragranular striate cortex appear to enhance the responsiveness of these neurons to their preferred sensory inputs, suggesting that it may control the gain on feedforward excitatory drive (Purushothaman et al. 2012) (**Figure 4*g***). [Note the precise biophysical interpretation of this parameter may not be straightforward, as the LGN-like input (r_k) we have defined in **Figure 4*f*** may be provided by neighboring cortical neurons that share similar tuning (Ko et al. 2011).] In summary, by recruiting subnetworks of cortical interneurons, different thalamic inputs can define task-specific ensembles within and across cortical microcircuits and, in doing so, can configure the computational state of the cortex in a manner relevant to the behavioral goal.

Taken together, the studies that we have reviewed support the notion that both the pulvinar and the MD are specialized in using contextual information to reconfigure cortical responses in a task-dependent manner. Given the input-output diversity of each thalamic nucleus, each thalamic nucleus possibly contains circuits dedicated to encoding contextual information, in addition to circuits dedicated to other distinct computations (e.g., relaying information). Within this modular framework, a common set of thalamocortical circuit operations (e.g., forward model, information relay) whose deployment varies according to behavioral demands can form the basis for attention, decision-making, and working memory.

CONCLUSIONS AND FUTURE DIRECTIONS

Most of our understanding of thalamic computations is derived from investigating LGN function, leading to the commonly held notion that the thalamus simply relays information from the primary sensory organs to the neocortex. Although this is likely the case for many thalamic circuits—even in the pulvinar and MD—we argue along evolutionary, developmental, biophysical, behavioral, and theoretical lines that the thalamus is well suited to perform much richer computations. In particular, we propose that the thalamus functions as a Bayesian observer to derive forward predictive models on the basis of contextual information. Through these models, the thalamus can then inform the cortex about expected outcomes or changes in key task-relevant variables. Thus, we propose a revised model of the thalamus in which individual thalamic nuclei may contain various microcircuits, each performing distinct computations. These microcircuits are defined by the inputs they receive, their biophysical properties, and the cortical areas they target.

As highlighted above, we have only a rudimentary understanding of the role of the thalamus in cognitive function. We conclude this review by proposing two key challenges to be addressed going forward. The first is to identify other thalamic circuits and the second is to decipher the computations they perform. Addressing these challenges is an endeavor that requires novel, well-controlled behavioral paradigms, genetically accessible species that can perform highly cognitive tasks, and the development of genetic tools that allow causal manipulation of neural circuit function. With their enhanced cognitive capacity over that of rodents, nonhuman primates may be a tractable model system to identify new thalamic computations. Recent advances have permitted the development of transgenic nonhuman primates (Sasaki 2015) and the successful application of techniques such as calcium imaging (Seidemann et al. 2016) and optical perturbation of defined cell types (Stauffer et al. 2016); however, current applications are limited.

Meanwhile, the genetic tractability of mice can still be leveraged by designing cognitively challenging sensory attention tasks (Schmitt et al. 2017, Wimmer et al. 2015) or by training them to make decisions from accumulated sensory evidence (Hanks et al. 2015). Together these

techniques will be instrumental in broadening our understanding of the computations performed by thalamic circuits.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We would like to thank Karl Friston and Xiao-Jing Wang for insightful discussions on our computational formulation and László Acsády for advice on thalamocortical anatomy. We also gratefully acknowledge funding support to M.M.H. from the National Institutes of Health, the National Science Foundation, the Human Frontiers Science Program, the Simons Foundation, the Klingenstein Foundation, the Sloan Foundation, the Feldstein Foundation, the Brain & Behavior Research Foundation, and the Pew Foundation.

LITERATURE CITED

Aboitiz F, Montiel J, Morales D, Concha M. 2002. Evolutionary divergence of the reptilian and the mammalian brains: considerations on connectivity and development. *Brain Res. Brain Res. Rev.* 39:141–53

Bar M. 2004. Visual objects in context. *Nat. Rev. Neurosci.* 5:617–29

Bar M, Ullman S. 1996. Spatial context in recognition. *Perception* 25:343–52

Bastos AM, Usrey WM, Adams RA, Mangun GR, Fries P, Friston KJ. 2012. Canonical microcircuits for predictive coding. *Neuron* 76:695–711

Berger TK, Silberberg G, Perin R, Markram H. 2010. Brief bursts self-inhibit and correlate the pyramidal network. *PLOS Biol.* 8:e1000473

Biederman I, Mezzanotte RJ, Rabinowitz JC. 1982. Scene perception: detecting and judging objects undergoing relational violations. *Cogn. Psychol.* 14:143–77

Bisley JW, Goldberg ME. 2003. The role of the parietal cortex in the neural processing of saccadic eye movements. *Adv. Neurol.* 93:141–57

Blasdel GG, Lund JS. 1983. Termination of afferent axons in macaque striate cortex. *J. Neurosci.* 3:1389–413

Bolam JP, Bevan MD. 2001. Microcircuits of the striatum. In *Basal Ganglia and Thalamus in Health and Movement Disorders*, ed. K Kultas-Ilinsky, IA Ilinsky, pp. 29–39. Boston: Springer

Bold EL, Castro AJ, Neafsey EJ. 1984. Cytoarchitecture of the dorsal thalamus of the rat. *Brain Res. Bull.* 12:521–27

Butler AB. 2008. Evolution of brains, cognition, and consciousness. *Brain Res. Bull.* 75:442–49

Butler AB. 2009. Evolution of the dorsal thalamus. In *Encyclopedia of Neuroscience*, ed. MD Binder, N Hirokawa, U Windhorst, pp. 1346–51. Berlin: Springer

Butler AB, Hodos W. 2005. *Comparative Vertebrate Neuroanatomy: Evolution and Adaptation*. Hoboken, NJ: Wiley. 2nd ed.

Butts DA, Desbordes G, Weng C, Jin J, Alonso J-M, Stanley GB. 2010. The episodic nature of spike trains in the early visual pathway. *J. Neurophysiol.* 104:3371–87

Cadieu CF, Hong H, Yamins DLK, Pinto N, Ardila D, et al. 2014. Deep neural networks rival the representation of primate IT cortex for core visual object recognition. *PLOS Comput. Biol.* 10:e1003963

Chalupa LM, Werner JS. 2003. *The Visual Neurosciences*. Cambridge, MA: MIT Press

Crandall SR, Cruikshank SJ, Connors BW. 2015. A corticothalamic switch: controlling the thalamus with dynamic synapses. *Neuron* 86:768–82

Curry MJ. 1972. The exteroceptive properties of neurones in the somatic part of the posterior group (PO). *Brain Res.* 44:439–62

De Valois RL, De Valois KK. 1980. Spatial vision. *Annu. Rev. Psychol.* 31:309–41

den Ouden HEM, Kok P, de Lange FP. 2012. How prediction errors shape perception, attention, and motivation. *Front. Psychol.* 3:548

DiCarlo JJ, Cox DD. 2007. Untangling invariant object recognition. *Trends Cogn. Sci.* 11:333–41

DiCarlo JJ, Zoccolan D, Rust NC. 2012. How does the brain solve visual object recognition? *Neuron* 73:415–34

Douglas RJ, Martin KAC. 2004. Neuronal circuits of the neocortex. *Annu. Rev. Neurosci.* 27:419–51

Emery NJ, Clayton NS. 2004. The mentality of crows: convergent evolution of intelligence in corvids and apes. *Science* 306:1903–7

Felleman DJ, Van Essen DC. 1991. Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex.* 1:1–47

Fino E, Yuste R. 2011. Dense inhibitory connectivity in neocortex. *Neuron* 69:1188–203

Fiser A, Mahringer D, Oyibo HK, Petersen AV, Leinweber M, Keller GB. 2016. Experience-dependent spatial expectations in mouse visual cortex. *Nat. Neurosci.* 19:1658–64

Friston K, Kiebel S. 2009. Predictive coding under the free-energy principle. *Philos. Trans. R. Soc. B* 364:1211–21

Fuhrmann G, Segev I, Markram H, Tsodyks M. 2002. Coding of temporal information by activity-dependent synapses. *J. Neurophysiol.* 87:140–48

Gaudry KS, Reinagel P. 2008. Information measure for analyzing specific spiking patterns and applications to LGN bursts. *Network* 19:69–94

Grant E, Hoerder-Suabedissen A, Molnár Z. 2012. Development of the corticothalamic projections. *Front. Neurosci.* 6:53

Griffiths TL, Tenenbaum JB. 2006. Optimal predictions in everyday cognition. *Psychol. Sci.* 17:767–73

Guido W, Lu SM, Vaughan JW, Godwin DW, Sherman SM. 1995. Receiver operating characteristic (ROC) analysis of neurons in the cat's lateral geniculate nucleus during tonic and burst response mode. *Vis. Neurosci.* 12:723–41

Guillery RW, Sherman SM. 2002. Thalamic relay functions and their role in corticocortical communication: generalizations from the visual system. *Neuron* 33:163–75

Haefner RM, Berkes P, Fiser J. 2016. Perceptual decision-making as probabilistic inference by neural sampling. *Neuron* 90:649–60

Halassa MM, Acsady L. 2016. Thalamic inhibition: diverse sources, diverse scales. *Trends Neurosci.* 39:680–93

Hanks TD, Kopec CD, Brunton BW, Duan CA, Erlich JC, Brody CD. 2015. Distinct relationships of parietal and prefrontal cortices to evidence accumulation. *Nature* 520:220–23

Harris KD, Mrsic-Flogel TD. 2013. Cortical connectivity and sensory coding. *Nature* 503:51–58

Harris KD, Shepherd GMG. 2015. The neocortical circuit: themes and variations. *Nat. Neurosci.* 18:170–81

Hassabis D, Kumaran D, Summerfield C, Botvinick M. 2017. Neuroscience-inspired artificial intelligence. *Neuron* 95:245–58

Hellier JL. 2014. *The Brain, the Nervous System, and Their Diseases, Volume 1: A–F*. Santa Barbara, CA: ABC-CLIO

Hubel DH, Wiesel TN. 1962. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.* 160:106–54

Ibos G, Duhamel J-R, Ben Hamed S. 2013. A functional hierarchy within the parietofrontal network in stimulus selection and attention control. *J. Neurosci.* 33:8359–69

Jahnsen H, Llinás R. 1984. Ionic basis for the electro-responsiveness and oscillatory properties of guinea-pig thalamic neurones in vitro. *J. Physiol.* 349:227–47

Jarvis ED, Güntürkün O, Bruce L, Csillag A, Karten H, et al. 2005. Avian brains and a new understanding of vertebrate brain evolution. *Nat. Rev. Neurosci.* 6:151–59

Johnson-Laird PN. 2010. Mental models and human reasoning. *PNAS* 107:18243–50

Jones EG. 1998. Viewpoint: the core and matrix of thalamic organization. *Neuroscience* 85:331–45

Jones EG. 2002. Thalamic circuitry and thalamocortical synchrony. *Philos. Trans. R. Soc. B* 357:1659–73

Jones EG. 2007. *The Thalamus*. Cambridge, MA: Cambridge Univ. Press. 2nd ed.

Jordan MI, Rumelhart DE. 1992. Forward models: supervised learning with a distal teacher. *Cogn. Sci.* 16:307–54

Kalman M. 2009. Evolution of the brain: at the reptile-bird transition. In *Encyclopedia of Neuroscience*, ed. MD Binder, N Hirokawa, U Windhorst, pp. 1305–12. Berlin: Springer

Kanai R, Komura Y, Shipp S, Friston K. 2015. Cerebral hierarchies: predictive processing, precision and the pulvinar. *Philos. Trans. R. Soc. B* 370:20140169

Karnani MM, Jackson J, Ayzenshtat I, Tucciarone J, Manoocheri K, et al. 2016. Cooperative subnetworks of molecularly similar interneurons in mouse neocortex. *Neuron* 90:86–100

Kiani R, Shadlen MN. 2009. Representation of confidence associated with a decision by neurons in the parietal cortex. *Science* 324:759–64

Kim J, Kim Y, Nakajima R, Shin A, Jeong M, et al. 2017. Inhibitory basal ganglia inputs induce excitatory motor signals in the thalamus. *Neuron* 95:1181–96

Ko H, Hofer SB, Pichler B, Buchanan KA, Sjostrom PJ, Mrsic-Flogel TD. 2011. Functional specificity of local synaptic connections in neocortical networks. *Nature* 473:87–91

Kok P, Brouwer GJ, van Gerven MAJ, de Lange FP. 2013. Prior expectations bias sensory representations in visual cortex. *J. Neurosci.* 33:16275–84

Komura Y, Nikkuni A, Hirashima N, Uetake T, Miyamoto A. 2013. Responses of pulvinar neurons reflect a subject's confidence in visual categorization. *Nat. Neurosci.* 16:749–55

Kuramoto E, Pan S, Furuta T, Tanaka YR, Iwai H, et al. 2017. Individual mediodorsal thalamic neurons project to multiple areas of the rat prefrontal cortex: a single neuron-tracing study using virus vectors. *J. Comp. Neurol.* 525:166–85

LeCun Y, Bengio Y, Hinton G. 2015. Deep learning. *Nature* 521:436–44

Lee CK, Huguenard JR. 2011. Martinotti cells: community organizers. *Neuron* 69:1042–45

Lee S, Ahmed T, Lee S, Kim H, Choi S, et al. 2011. Bidirectional modulation of fear extinction by mediodorsal thalamic firing in mice. *Nat. Neurosci.* 15:308–14

Lee TS. 2015. The visual system's internal model of the world. *Proc. IEEE. Inst. Electr. Electron. Eng.* 103:1359–78

Lee TS, Mumford D. 2003. Hierarchical Bayesian inference in the visual cortex. *J. Opt. Soc. Am. A. Opt. Image Sci. Vis.* 20:1434–48

Lefebvre L, Reader SM, Sol D. 2004. Brains, innovations and evolution in birds and primates. *Brain Behav. Evol.* 63:233–46

Lerner TN, Ye L, Deisseroth K. 2016. Communication in neural circuits: tools, opportunities, and challenges. *Cell* 164:1136–50

Lesica NA, Stanley GB. 2004. Encoding of natural scene movies by tonic and burst spikes in the lateral geniculate nucleus. *J. Neurosci.* 24:10731–40

Livingstone MS, Hubel DH. 1988. Do the relative mapping densities of the magno- and parvocellular systems vary with eccentricity? *J. Neurosci.* 8:4334–39

Losonczy A, Makara JK, Magee JC. 2008. Compartmentalized dendritic plasticity and input feature storage in neurons. *Nature* 452:436–41

Marino AC, Mazer JA. 2016. Perisaccadic updating of visual representations and attentional states: linking behavior and neurophysiology. *Front. Syst. Neurosci.* 10:3

Matesz K, Kecskes S, Bácskai T, Rácz É, Birinyi A. 2014. Brainstem circuits underlying the prey-catching behavior of the frog. *Brain Behav. Evol.* 83:104–11

McCotter M, Gosselin F, Sowden P, Schyns P. 2005. The use of visual information in natural scenes. *Vis. Cogn.* 12:938–53

Mease RA, Kuner T, Fairhall AL, Groh A. 2017. Multiplexed spike coding and adaptation in the thalamus. *Cell Rep*. 19:1130–40

Miller EK, Cohen JD. 2001. An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24:167–202

Mitchell AS, Chakraborty S. 2013. What does the mediodorsal thalamus do? *Front. Syst. Neurosci.* 7:37

Mlynarski W, Hermundstad AM. 2017. Adaptive coding for dynamic sensory inference. *bioRxiv* **https://doi.org/10.1101/189506**

Molnár Z, Butler AB. 2002. The corticostriatal junction: a crucial region for forebrain development and evolution. *Bioessays* 24:530–41

Montagnini A, Treves A. 2003. The evolution of mammalian cortex, from lamination to arealization. *Brain Res. Bull.* 60:387–93

Muñoz W, Tremblay R, Levenstein D, Rudy B. 2017. Layer-specific modulation of neocortical dendritic inhibition during active wakefulness. *Science* 355:954–59

Naud R, Sprekeler H. 2017. Burst ensemble multiplexing: a neural code connecting dendritic spikes with microcircuits. *bioRxiv* **https://doi.org/10.1101/143636**

Naumann RK, Ondracek JM, Reiter S, Shein-Idelson M, Tosches MA, et al. 2015. The reptilian brain. *Curr. Biol.* 25:R317–21

Oliva A, Torralba A. 2006. Building the gist of a scene: the role of global image features in recognition. *Prog. Brain Res.* 155:23–36

Oliva A, Torralba A. 2007. The role of context in object recognition. *Trends Cogn. Sci.* 11:520–27

Orban GA. 2008. Higher order visual processing in macaque extrastriate cortex. *Physiol. Rev.* 88:59–89

O'Reilly RC, Wyatte DR, Rohrlich J. 2017. Deep predictive learning: a comprehensive model of three visual streams. Cornell University Library. **https://arxiv.org/abs/1709.04654**

Paul A, Crow M, Raudales R, He M, Gillis J, Huang ZJ. 2017. Transcriptional architecture of synaptic communication delineates GABAergic neuron identity. *Cell*. 171:522–39.e20

Paxinos G, Franklin KBJ. 2004. *The Mouse Brain in Stereotaxic Coordinates*. Houston: Gulf Professional Publishing

Peebles D, Cooper RP. 2015. Thirty years after Marr's vision: levels of analysis in cognitive science. *Top. Cogn. Sci.* 7:187–90

Phillips JW, Schulmann A, Hara E, Liu C, Shields B, et al. 2017. A topographic axis of transcriptional identity in thalamus. *bioRxiv* doi: 10.1101/241315

Pritz MB. 1995. The thalamus of reptiles and mammals: similarities and differences. *Brain Behav. Evol.* 46:197–208

Purushothaman G, Marion R, Li K, Casagrande VA. 2012. Gating and control of primary visual cortex by pulvinar. *Nat. Neurosci.* 15:905–12

Ramcharan EJ, Gnadt JW, Sherman SM. 2000. Burst and tonic firing in thalamic cells of unanesthetized, behaving monkeys. *Vis. Neurosci.* 17:55–62

Ramsay ZJ, Ikura J, Laberge F. 2013. Modification of a prey catching response and the development of behavioral persistence in the fire-bellied toad (*Bombina orientalis*). *J. Comp. Psychol.* 127:399–411

Rao HM, Mayo JP, Sommer MA. 2016. Circuits for presaccadic visual remapping. *J. Neurophysiol.* 116:2624–36

Rao RP, Ballard DH. 1999. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* 2:79–87

Reinagel P, Godwin D, Sherman SM, Koch C. 1999. Encoding of visual information by LGN bursts. *J. Neurophysiol.* 81:2558–69

Reinagel P, Reid RC. 2000. Temporal coding of visual information in the thalamus. *J. Neurosci.* 20:5392–400

Rockland KS, Andresen J, Cowie RJ, Robinson DL. 1999. Single axon analysis of pulvinocortical connections to several visual areas in the macaque. *J. Comp. Neurol.* 406:221–50

Rohe T, Noppeney U. 2015. Cortical hierarchies perform Bayesian causal inference in multisensory perception. *PLOS Biol.* 13:e1002073

Roth MM, Dahmen JC, Muir DR, Imhof F, Martini FJ, Hofer SB. 2016. Thalamic nuclei convey diverse contextual information to layer 1 of visual cortex. *Nat. Neurosci.* 19:299–307

Roth MM, Helmchen F, Kampa BM. 2012. Distinct functional properties of primary and posteromedial visual area of mouse neocortex. *J. Neurosci.* 32:9716–26

Rouiller EM, Welker E. 2000. A comparative analysis of the morphology of corticothalamic projections in mammals. *Brain Res. Bull.* 53:727–41

Rovo Z, Ulbert I, Acsady L. 2012. Drivers of the primate thalamus. *J. Neurosci.* 32:17894–908

Sasaki E. 2015. Prospects for genetically modified non-human primate models, including the common marmoset. *Neurosci. Res.* 93:110–15

Schall JD, Thompson KG. 1999. Neural selection and control of visually guided eye movements. *Annu. Rev. Neurosci.* 22:241–59

Schmitt LI, Wimmer RD, Nakajima M, Happ M, Mofakham S, Halassa MM. 2017. Thalamic amplification of cortical connectivity sustains attentional control. *Nature* 545:219–23

Seabrook TA, Burbridge TJ, Crair MC, Huberman AD. 2017. Architecture, function, and assembly of the mouse visual system. *Annu. Rev. Neurosci.* 40:499–538

Seidemann E, Chen Y, Bai Y, Chen SC, Mehta P, et al. 2016. Calcium imaging with genetically encoded indicators in behaving primates. *eLife* 5:e16178

Shadlen MN, Shohamy D. 2016. Decision making and sequential sampling from memory. *Neuron* 90:927–39

Sherman SM. 2001. Tonic and burst firing: dual modes of thalamocortical relay. *Trends Neurosci.* 24:122–26

Shin S, Sommer MA. 2012. Division of labor in frontal eye field neurons during presaccadic remapping of visual receptive fields. *J. Neurophysiol.* 108:2144–59

Shipp S. 2004. The brain circuitry of attention. *Trends Cogn. Sci.* 8:223–30

Sommer MA, Wurtz RH. 2004a. What the brain stem tells the frontal cortex. I. Oculomotor signals sent from superior colliculus to frontal eye field via mediodorsal thalamus. *J. Neurophysiol.* 91:1381–402

Sommer MA, Wurtz RH. 2004b. What the brain stem tells the frontal cortex. II. Role of the SC-MD-FEF pathway in corollary discharge. *J. Neurophysiol.* 91:1403–23

Sommer MA, Wurtz RH. 2008. Brain circuits for the internal monitoring of movements. *Annu. Rev. Neurosci.* 31:317–38

Stauffer WR, Lak A, Yang A, Borel M, Paulsen O, et al. 2016. Dopamine neuron-specific optogenetic stimulation in rhesus macaques. *Cell* 166:1564–71.e6

Steriade M, McCormick DA, Sejnowski TJ. 1993. Thalamocortical oscillations in the sleeping and aroused brain. *Science* 262:679–85

Striedter GF. 2006. Précis of principles of brain evolution. *Behav. Brain Sci.* 29:1–12

Tolman EC. 1948. Cognitive maps in rats and men. *Psychol. Rev.* 55:189–208

Tremblay R, Lee S, Rudy B. 2016. GABAergic interneurons in the neocortex: from cellular properties to circuits. *Neuron* 91:260–92

Tsodyks M, Pawelzik K, Markram H. 1998. Neural networks with dynamic synapses. *Neural Comput.* 10:821–35

Unzai T, Kuramoto E, Kaneko T, Fujiyama F. 2017. Quantitative analyses of the projection of individual neurons from the midline thalamic nuclei to the striosome and matrix compartments of the rat striatum. *Cereb. Cortex.* 27:1164–81

Usrey WM, Muly EC, Fitzpatrick D. 1992. Lateral geniculate projections to the superficial layers of visual cortex in the tree shrew. *J. Comp. Neurol.* 319:159–71

Vertes RP, Linley SB, Groenewegen HJ, Witter MP, Paxinos G. 2015. Thalamus. In *The Rat Nervous System*, ed. G Paxinos, pp. 335–90. Oxford, UK: Elsevier. 4th ed.

Whitmire CJ, Waiblinger C, Schwarz C, Stanley GB. 2016. Information coding through adaptive gating of synchronized thalamic bursting. *Cell Rep.* 14:795–807

Wimmer RD, Schmitt LI, Davidson TJ, Nakajima M, Deisseroth K, Halassa MM. 2015. Thalamic control of sensory selection in divided attention. *Nature* 526:705–9

Wolpert DM, Miall RC. 1996. Forward models for physiological motor control. *Neural Netw.* 9:1265–79

Wyder MT, Massoglia DP, Stanford TR. 2004. Contextual modulation of central thalamic delay-period activity: representation of visual and saccadic goals. *J. Neurophysiol.* 91:2628–48

Xie Y, Dorsky RI. 2017. Development of the hypothalamus: conservation, modification and innovation. *Development* 144:1588–99

Yamins DLK, DiCarlo JJ. 2016. Using goal-driven deep learning models to understand sensory cortex. *Nat. Neurosci.* 19:356–65

Zhan XJ, Cox CL, Sherman SM. 2000. Dendritic depolarization efficiently attenuates low-threshold calcium spikes in thalamic relay cells. *J. Neurosci.* 20:3909–14

Zhou H, Schafer RJ, Desimone R. 2016. Pulvinar-cortex interactions in vision and attention. *Neuron* 89:209–20