Normalization Principles in Computational Neuroscience `FREE`

Kenway Louie, New York University, Center for Neural Science and Paul W. Glimcher, New York University, Center for Neural Science

Summary

A core question in systems and computational neuroscience is how the brain represents information. Identifying principles of information coding in neural circuits is critical to understanding brain organization and function in sensory, motor, and cognitive neuroscience. This provides a conceptual bridge between the underlying biophysical mechanisms and the ultimate behavioral goals of the organism. Central to this framework is the question of computation: what are the relevant representations of input and output, and what algorithms govern the input-output transformation? Remarkably, evidence suggests that certain canonical computations exist across different circuits, brain regions, and species. Such computations are implemented by different biophysical and network mechanisms, indicating that the unifying target of conservation is the algorithmic form of information processing rather than the specific biological implementation.

A prime candidate to serve as a canonical computation is divisive normalization, which scales the activity of a given neuron by the activity of a larger neuronal pool. This nonlinear transformation introduces an intrinsic contextual modulation into information coding, such that the selective response of a neuron to features of the input is scaled by other input characteristics. This contextual modulation allows the normalization model to capture a wide array of neural and behavioral phenomena not captured by simpler linear models of information processing. The generality and flexibility of the normalization model arises from the normalization pool, which allows different inputs to directly drive and suppress a given neuron, effectively separating information that drives excitation and contextual modulation. Originally proposed to describe responses in early visual cortex, normalization has been widely documented in different brain regions, hierarchical levels, and modalities of sensory processing; furthermore, recent work shows that the normalization extends to cognitive processes such as attention, multisensory integration, and decision making. This ubiquity reinforces the canonical nature of the normalization computation and highlights the importance of an algorithmic framework in linking biological mechanism and behavior.

Keywords: divisive normalization, canonical computation, efficient coding, neurophysiology, sensory encoding, decision making, contextual modulation

Subjects: Computational Neuroscience

Introduction

Neural Coding and Decoding

A central question in neuroscience is how the brain represents behaviorally relevant information (Rieke, Warland, de Ruyter van Steveninck, & Bialek, 1997). Identifying the code by which neural systems represent stimulus or movement information was one of principal questions in the development of neuroscience as a field. Early pioneering studies on peripheral sensory neurons led to the proposal that neurons communicate information via the average frequency of action potentials (Adrian, 1926; Adrian & Zotterman, 1926). While this rate code hypothesis still underlies much current research in neuroscience, ideas about neural encoding have expanded to incorporate more complex characteristics of neural activity such as precise spike timing (Hopfield, 1995; Singer & Gray, 1995; Theunissen & Miller, 1995) and correlated activity (Averbeck, Latham, & Pouget, 2006; Kohn, Coen-Cagli, Kanitscheider, & Pouget, 2016).

The study of neural encoding—how neural responses represent stimulus, movement, or cognitive information—is intricately tied to the study of neural decoding, or how behaviorally relevant information can be reconstructed from neural responses. The question of neural decoding is important for both theoretical and practical reasons; accurately interpreting decoding is critical to understanding neural information processing and cognitive function, but is also a prerequisite in the development of neural prosthetic devices (Andersen, Hwang, & Mulliken, 2010; Hochberg et al., 2006; Schwartz, Cui, Weber, & Moran, 2006). Examination of both neural coding and decoding has grown more sophisticated, driven by ongoing technological developments, such as large-scale multineuronal electrophysiological recording and optical imaging techniques, that allow high density measurements of neural activity. Given this increased density of neural data and the inherent stochasticity in neural activity, growing focus has turned to computational and statistical techniques to quantify neural information processing (Paninski, Pillow, & Lewi, 2007; Rao, Olshausen, & Lewicki, 2002).

Algorithmic Approach to Neural Information Processing

A focus on Neural Computation is particularly relevant to systems neuroscience, where potential levels of analysis span from intracellular biophysics to the level of cognition and behavior (Carandini, 2012; Sejnowski, Koch, & Churchland, 1988). An inspiration for this approach is Marr's tri-level hypothesis: a hierarchical framework for thinking about information processing systems, which posits different complementary but functionally distinct levels of analysis (Marr, 1982). In Marr's proposal, the highest, most abstract level of analysis (the *computational* level) focuses on the goal of a system and the intended problems to be solved. At the other end of the spectrum, the lowest level of analysis (the *implementation* level) encompasses the physical realization of the system. Intervening between the two is an intermediate level of analysis (the *algorithmic* level) that describes the relevant representations of input and output, and the algorithms governing input-output transformation. An intuitive description is that these levels represent how (implementation), what (algorithmic), and why (computational) information is

computed; for neural systems, these levels roughly correspond to neurons and neural circuits, Neural Computation and encoding/decoding, and the ultimate goals of the organism (metabolic, behavioral, or evolutionary) (Carandini, 2012). Note that the term "computation" is broadly used in the neuroscience literature, often without explicit definition; most generally, the term refers to the process by which a neuron or population of neurons transforms information between inputs and outputs, which corresponds most closely to Marr's intermediate algorithmic level. Perhaps the most important takeaway from Marr's approach is a distinction between biophysical implementation and information representation: the information processed by a neural system and how it is transformed—and not the specific biological apparatus used to achieve this transformation—defines its functional role and link to behavior.

Modularity and Canonical Computations

The growing focus on Neural Computation within neuroscience has led to the proposal that certain computations are particularly prevalent, perhaps even ubiquitous, in neural systems (Bastos et al., 2012; Carandini, 2012; Carandini & Heeger, 2012; Kouh & Poggio, 2008). Such computations are viewed as canonical in nature, performing analogous information processing functions in different circuits, brain regions, and species. The proposal for canonical computations is motivated, at least in part, by the notion of a canonical microcircuit: a characteristic laminar-based architecture of local circuitry widely repeated across cerebral cortex (Douglas & Martin, 1991, 2004). The nature of canonical microcircuits highlights two points of emphasis relevant to canonical computations: (1) intrinsic connectivity within a circuit— including local inhibition as well as excitation—plays a key role in shaping its information processing function, and (2) neural systems follow a modular organization. The importance of intrinsic connectivity has driven a shift in focus from serial feedforward processing between brain areas (Hubel & Wiesel, 1962; Van Essen, Anderson, & Felleman, 1992) to recurrent excitation and inhibition within brain areas, which defines the computations performed within circuit modules. The modular anatomy of the brain, as emphasized in the cortical microcircuit proposal, has led to the corresponding search for computations with similar hallmarks of modularity; such computations would perform a core operation, occur repeatedly and widely across the brain, be implemented by a variety of circuits and mechanisms, and potentially cascade with other computations.

While there is no standard definition for canonical computations, several prominent operations stand out as potential candidates (Carandini, 2012; Carandini & Heeger, 2012). One example is linear filtering, in which the response of a given neuron can be described as a weighted sum of inputs. The classic example of linear filtering is the receptive field, in which sensory neuron activity is driven as a linear function of spatiotemporal properties of the stimulus (Hartline, 1940; Kuffler, 1953). In addition to characterizing receptive field-driven sensory responses in vision (DeAngelis, Ohzawa, & Freeman, 1995; Hubel & Wiesel, 1962), audition (Aertsen & Johannesma, 1981; deCharms, Blake, & Merzenich, 1998), and somatosensation (Brecht, Roth, & Sakmann, 2003; DiCarlo & Johnson, 2002), linear filtering—operating on basis functions—may mediate higher-order phenomena such as cue integration and sensorimotor transformations (Deneve, Latham, & Pouget, 2001; Pouget & Sejnowski, 1997; Pouget & Snyder, 2000). A second candidate

canonical computation is the soft thresholding of neural responses. Thresholding is an element of many models of spiking activity, describing the nonlinearity inherent in the conversion of input currents into action potentials. At the single neuron level, this computation is often assumed to be a threshold-linear or power law function (Albrecht & Geisler, 1991; Carandini, 2004; Movshon, Thompson, & Tolhurst, 1978), setting the functional operating point of neural systems (Ringach & Malone, 2007). However, thresholding can also operate at the level of circuits and networks, for example in the recurrent amplification mechanisms underlying action selection and decision making (Cisek, 2006; Hanes & Schall, 1996; Roitman & Shadlen, 2002; Wang, 2008). Consistent with the modular nature of canonical computations, models that cascade computations like linear filtering and soft thresholding together accurately capture neural spike responses (Chichilnisky, 2001; Simoncelli, Paninski, Pillow, & Schwartz, 2004). In addition to linear filtering and soft thresholding, a number of other candidate computations may serve a canonical function including coincidence detection, associative learning, predictive coding, and constrained trajectories in dynamical systems (Carandini, 2012).

This article focuses on an additional neural operation thought to be a prime candidate for a canonical Neural Computation: divisive normalization (Carandini & Heeger, 2012; Louie, Glimcher, & Webb, 2015). In normalization, output neural responses reflect both driving, afferent input and divisive input from a large pool of other neurons (see "The Divisive Normalization Model"). This division implements a form of gain control that allows normalization to capture aspects of neural responses unexplained by linear models. Originally proposed to explain nonlinear response properties in early visual cortex, normalization has since been found to operate widely across different brain regions, sensory modalities, and species (see "Normalization in Neural Responses"). Supporting its proposed canonical nature, normalization extends beyond early sensory coding to higher order processes including attention (Lee & Maunsell, 2009; Reynolds & Heeger, 2009), multisensory integration (Ohshiro, Angelaki, & DeAngelis, 2011, 2017), and valuation and decision making (Khaw, Glimcher, & Louie, 2017; Louie, Grattan, & Glimcher, 2011; Louie, Khaw, & Glimcher, 2013). In addition to its operation in different neural systems, emerging evidence links the normalization process to perceptual and choice behavior (see under "Normalization in Behavior"). While some computational functions of normalization—such as gain modulation and control of saturation—are evident from its formalization, theoretical work has proposed a variety of more complex computational roles such as marginalization, invariant coding, and redundancy reduction (see under "Computational Functions"). Despite similar computational functions across different neural systems, normalization is mediated by a variety of biophysical and circuit mechanisms, consistent with a preservation of computational role rather than circuit organization (see under "Biophysical Implementation"). The final section briefly overviews current active areas of research on normalization, including engineering aspects of implementing normalization-like operations into deep neural networks (see "Future Directions").

The Divisive Normalization Model

Historical Precedent

An early form of divisive normalization was first formally proposed as a computational explanation for non-linear response properties of neurons in primary visual cortex (V1) (Heeger, 1992, 1993). At the time, the predominant models of simple and complex V1 cells were primarily linear and feedforward in nature, performing a half-wave rectified linear filtering of stimulus information (Campbell, Cooper, & Enroth-Cugell, 1969; Hubel & Wiesel, 1962) or a summing of squared outputs of linear filters (Adelson & Bergen, 1985). However, while generally robust in descriptive power and empirically easy to fit, these linear (simple cell) and energy (complex cell) models failed to capture two characteristics of V1 activity: response saturation and nonspecific suppression. First, while visual cortical neurons saturate with increasing stimulus contrast, the linear and energy mechanisms predict monotonically increasing responses across all contrasts. Second, though visual cortical neurons are driven in a stimulus-specific manner, they exhibit nonspecific suppression largely independent of stimulus characteristics (e.g., orientation, spatial frequency, spatial location). This nonspecific suppression cannot be easily explained by the linear and energy models, which only capture the stimulus properties that drive excitatory responses.

To address the inadequacies of existing models to explain V1 data, Heeger formalized a model that combined divisive scaling with the notion of a normalization pool and a form of rectification. Divisive scaling was already known to describe how visual cortical responses depend on stimulus contrast (i.e., the contrast response function). Specifically, V1 contrast response functions were known to be well described by a hyperbolic ratio:

$$R\left(C\right) = R_{max}\frac{C^n}{C^n + C_{50}^n}$$

(1)

where R_{max} denotes the maximal firing rate, n is an exponent controlling the shape of the response, and the semisaturation contrast C_{50} refers to the contrast that produces half-maximal response (Albrecht & Geisler, 1991; Albrecht & Hamilton, 1982; Sclar, Maunsell, & Lennie, 1990). Originally applied to describe neural responses in the retinal (Naka & Rushton, 1966), the hyperbolic ratio equation describes a saturating function analogous to the Michaelis-Menten model for enzyme kinetics (Michaelis, Menten, Johnson, & Goody, 2011). In Heeger's formulation (Heeger, 1993), divisive normalization incorporated a similar hyperbolic form for divisive scaling (shown here for simple cells):

$$R_i = K\frac{A_i}{\sigma^2 + \sum_j A_j}$$

(2)

where R_i is the normalized response of cell i, K is a maximal firing rate, and σ^2 is a semisaturation constant analogous to the hyperbolic ratio C_{50} term. The term A_i represents a linear operator on stimulus information (i.e., a receptive field) followed by a half-squaring and can be viewed as the unnormalized response of neuron i; due to this half squaring, the normalization model is analogous to a hyperbolic ratio with an exponent of 2. Note that for simplicity the linear operator terms A are depicted as static terms, but can be represented as functions of time to introduce dynamics into the model. Similar to other previous approaches (Albrecht & Geisler, 1991), Heeger's model combined a linear stage (weighting of stimulus information) with a static nonlinearity (half-squaring) and divisive gain control nonlinearity.

The key advance of Heeger's normalization model was the contribution of a pool of neurons to the denominator (Heeger, 1992, 1993; Sawada & Petrov, 2017). In Equation 2, this is represented as the summation of A_j over a large number of neurons j (which typically include neuron i, providing for a degree of self-inhibition). This pooling provides a natural mechanism for estimating stimulus contrast, and its contribution in the denominator implements the hyperbolic ratio contrast response function seen in empirical data (Albrecht & Geisler, 1991; Albrecht & Hamilton, 1982; Sclar et al., 1990). More importantly, the addition of a normalization pool allows different inputs to directly drive and suppress a given neuron, providing separate stimulus contributions to excitation and contextual modulation. For example, V1 cells in the normalization model will be directly driven by stimuli at a specific orientation and spatial frequency but receive divisive suppression from neurons driven by all orientations and a broad range of spatial frequencies (Heeger, 1992). This pooled signal allows the normalization model to accurately describe nonspecific suppression, where the response to a preferred stimulus is suppressed by the superposition of additional stimuli (e.g., cross-orientation suppression); this suppression is typically much more nonselective than activation, with broad spatial selectivity, spatial frequency tuning, and dependence on orientation (Blakemore & Tobin, 1972; Bonds, 1989; DeAngelis, Robson, Ohzawa, & Freeman, 1992). Given flexibility in the definition of the normalization pool, the model can also describe additional nonlinear effects in V1 responses including contrast gain control, contrast adaptation, and surround suppression.

General Description

The influence of the divisive normalization model in neuroscience is driven by its ability to capture a wide array of response phenomena in different systems. The fundamental feature of the normalization model—common across different implementations—is a divisive scaling in which different inputs contribute as primary afferent drive (numerator) and modulatory control (denominator). Variants of the normalization model use different parameterizations and terminology (Carandini & Heeger, 2012; Sawada & Petrov, 2017); however, a general equation capturing most of the common implementations can be written (Eqn. 1):

$$R_i = R_{max} \frac{D_i^n + \beta}{\sigma^n + \sum_j \omega_{ij} D_j^n}$$

<div align="right">(3)</div>

where the response R_i of a neuron i is determined by both its direct driving input D_i and the summed inputs to a larger group of neurons denoted by j (often termed the normalization pool, and typically including drive to neuron i itself). In the original model of normalization proposed to explain nonlinearities in primary visual cortex (Heeger, 1992, 1993), this formulation attributes a neuron's stimulus selectivity to summation (a linear filtering stage in the construction of D_i) and its nonlinear response properties to division (the normalization stage). Consistent with this view, the quantities D are often viewed in the sensory literature as properties of the stimulus and have units of stimulus intensity; for example, to model suppressive phenomena in visual cortical neurons D can be expressed in units of contrast (Carandini, Heeger, & Movshon, 1997; Freeman, Durand, Kiper, & Carandini, 2002; Heeger, 1992). Alternatively, D can be expressed in units of neural activity in models implementing normalization in neural circuits (Beck, Latham, & Pouget, 2011; Kouh & Poggio, 2008; Louie, LoFaro, Webb, & Glimcher, 2014; Ohshiro et al., 2017).

In addition to its divisive formulation, the normalization model includes a small number of parameters contributing to its explanatory power. One parameter is a simple scaling parameter, denoted as R_{max} in Eqn. 2, that defines the maximum level of activity; because D_i is typically included in both the numerator and denominator, normalization implements a saturating function of input drive that approaches R_{max}. The semisaturation parameter σ determines how the normalization function responds to driving input, controlling how neural activity approaches saturation and the range of inputs that most effectively drive responses. The general effect of the semisaturation parameter is evident by examining its direct counterpart (e.g., C_{50}) in the predecessor hyperbolic ratio model (Eqn. 1); in that simpler model, the term sets the input level driving half-maximal output (Albrecht & Hamilton, 1982; Naka & Rushton, 1966). Additionally, a nonzero semisaturation term prevents division by zero in the absence of suppressive inputs. The exponent n allows for exponential amplification of the driving input and in most models is assumed to be the same in the numerator and denominator. Theoretically, the exponent represents an expansive nonlinearity in the conversion between input and spiking activity; in the original Heeger formulation for visual cortical neurons, n was set to 2 to implement a half-squaring. Empirically, the exponent parameter is often fit to neural responses, for example yielding values of n between 1.0 and 3.5 with an average value of 2 (with considerable variability across neurons) when fitting single V1 neuron spiking activity (Albrecht & Hamilton, 1982; Busse, Wade, & Carandini, 2009; Sclar et al., 1990). The parameter β controls the effect baseline response of the model; specifically, with no input drive to units in the numerator and denominator, the response in Eqn. 1 will be $R_{max}\beta/\sigma^n$. While not present in many implementations, this parameter enables the model to capture divisive effects on spontaneous levels of activity in the absence of direct input, such as the contextual suppression of visually responsive-neurons in the absence of a stimulus in the receptive field (Louie et al., 2011; Nassi, Avery, Cetin, Roe, & Reynolds, 2015). Finally, the parameters ω_{ij} allow for differential weighting of individual neuron contributions to the normalization pool. These weights provide for a flexible

normalization pool that can be tuned (e.g. to characteristics of the environment) in either in a static (Carandini et al., 1997; Ni, Ray, & Maunsell, 2012; Rust, Mante, Simoncelli, & Movshon, 2006; Schwartz & Simoncelli, 2001) or dynamic (Coen-Cagli, Dayan, & Schwartz, 2012; Coen-Cagli, Kohn, & Schwartz, 2015; Westrick, Heeger, & Landy, 2016) manner.

While most applications of the normalization model are static, with a fixed computation and varying inputs, newer work has formulated dynamical versions of the normalization model. Such models apply differential equations to model the underlying synaptic or firing rate dynamics that underlie the normalization computation, and are conceptually related to earlier dynamic models of the normalization process (Carandini & Heeger, 1994; Carandini et al., 1997; Mikaelian & Simoncelli, 2001; Wilson & Humanski, 1993). For example, recent work implemented a dynamical rate model of normalized value coding, using a set of differential equations to model a simple circuit of excitatory and inhibitory units in posterior parietal cortex (LoFaro, Louie, Webb, & Glimcher, 2014; Louie et al., 2014). This circuit organization implements feedback inhibition via recurrent connectivity, a circuit motif thought to underlie normalization in cortical brain regions (see "Biophysical Implementation"). At steady state, this dynamical model replicates features of context-dependent action value coding observed in monkey parietal neurons and previously described by a static normalization model (Louie et al., 2011) (see "Normalization in Higher Cognitive Processes"). However, the time-varying nature of the dynamical normalization model captures additional novel characteristics of reward-related parietal activity dynamics, including value coding during initial onset transients, time-varying value modulation, and delayed onset of contextual information. More broadly, recent theoretical work shows that normalization models can be derived as the asymptotic solutions to shunting differential equations, which have previously been proposed as fundamental models of Neural Computation (Grossberg, 1988). Using this approach, dynamical normalization models provide a unified account of attentional phenomena related to visual short-term memory including effects on both response time and accuracy (Smith, Sewell, & Lilburn, 2015).

Normalization in Neural Responses

Normalization in Sensory Coding

Originally proposed to describe V1 responses, the normalization model captures a number of contextual and suppressive phenomena in neural activity along the visual hierarchy (Carandini & Heeger, 2012). One of the earliest stages in the visual hierarchy with normalized responses is the retina. Retinal photoreceptors receive a vast range of light intensities, both over time and within individual visual scenes (Rieke & Rudd, 2009); to accurately encode these varying light intensities with a limited dynamic range of neural activity, the retina employs multiple adaptive steps of normalization. The first step is light adaptation, which shifts the input-output function transforming light intensity to photoreceptor activity according to the local average intensity over time (Boynton & Whitten, 1970; Normann & Perlman, 1979; Schneeweis & Schnapf, 1999; Shapley & Enroth-Cugell, 1984). This transformation effectively adjusts the sensitivity of the

input–output function to the predominant background light intensity, producing responses that represent not light intensity but contrast (deviation from mean light intensity). The normalization model describes this adjustment to background light intensity as:

$$R = R_{max} \frac{I^n}{\sigma^n + I_m^n + I^n}$$

(4)

where I represents the light intensity driving a single photoreceptor and I_m the mean background light intensity, with an exponent n equal to 1 (Carandini & Heeger, 2012). The second step is contrast gain control, which produces responses in downstream retinal bipolar and ganglion cells that represent not contrast but contrast relative to the contrast in surrounding spatial locations (Baccus & Meister, 2002; Shapley & Victor, 1978, 1981). Contrast gain control produces a number of suppressive phenomena such as contrast saturation, masking, and size tuning that can be explained by normalization (Bonin, Mante, & Carandini, 2005, 2006). Differing from light adaptation, where the normalization denominator carries a term representing local light intensity, normalization underlying contrast gain control divides by a measure of local contrast (typically the standard deviation of contrast in a region described by the suppressive field). The excitatory and suppressive inputs (D_i and D_j in Equation 3) can be defined as the weighted sum of contrasts in different spatial locations, capturing the spatial profile of empirically measured summation and suppressive fields. Together, light adaptation and contrast gain control highlight two descriptive features of the normalization model. First, normalization implements a relative form of information transmission, with the appropriate selection of denominator terms determining the nature of the contextual representation (e.g., normalization to background light intensity versus local contrast). Second, weighted contributions to both numerator and denominator input terms allow the normalization model to capture a wide range of physiological phenomena. In light adaptation, defining mean intensity as an average over time captures history-dependent effects; in contrast gain control, defining inputs a weighting of spatial locations captures the differing Gaussian spatial profiles of excitatory and suppressive fields.

Beyond the retina, normalization has been observed at multiple cortical areas along the visual hierarchy. As discussed above (see section "The Divisive Normalization Model"), normalization captures multiple nonlinear response phenomena in V1 neurons. One example is contrast saturation, in which firing rates saturate as contrast increases; explaining saturating contrast response functions in part motivated the development of the original normalization model as well as earlier hyperbolic ratio models. In contrast saturation, firing rates reach different asymptotic levels depending on ability of the stimulus to drive the neuron, and thus even at saturation neurons maintain orientation tuning (Albrecht & Geisler, 1991; Albrecht & Hamilton, 1982). Normalization captures this contrast-invariant tuning due to the differential specificities of the driving and suppressive inputs (Carandini et al., 1997; Heeger, 1992): the input in the numerator is orientation-specific and mediates orientation tuning (at all contrasts), while the inputs in the denominator pool over all orientations and drive contrast-dependent suppression

and saturation (regardless of the specific stimulus orientation). In addition to contrast saturation, normalization also captures a variety of suppressive phenomena in which additional stimuli decrease V1 responses to preferred stimuli. In cross-orientation suppression (a specific example of nonspecific suppression), the response to an optimally oriented grating is suppressed by the superposition of an orthogonally oriented grating in the receptive field (Bonds, 1989; DeAngelis et al., 1992; Morrone, Burr, & Maffei, 1982). In surround suppression, responses are suppressed by additional visual stimuli in locations surrounding the receptive field (Blakemore & Tobin, 1972; Cavanaugh, Bair, & Movshon, 2002b; DeAngelis, Freeman, & Ohzawa, 1994). Suppressive stimuli in both cross-orientation and surround suppression decrease responses to preferred stimuli, despite eliciting little activity when presented alone—an example of nonlinear processing. Normalization accounts for such nonlinear suppression because suppressive stimuli contribute selectively to the denominator, due to broader orientation tuning and spatial selectivity in the normalization pool (Carandini & Heeger, 2012; Carandini et al., 1997; Cavanaugh, Bair, & Movshon, 2002a).

Normalization has been documented in neural responses beyond V1 in both the dorsal and ventral visual pathways. Receptive fields are typically larger in these subsequent stages of visual processing, consistent with a serial convergence of feedforward inputs, and demonstrate stimulus interactions consistent with the normalization model. Along the dorsal pathway, neurons in the middle temporal (MT) area are tuned to the direction and speed of motion stimuli. When pairs of moving stimuli are placed within the receptive field, MT neurons exhibit significantly lower firing rates than the sum of individual stimulus responses (Britten & Heuer, 1999; Recanzone, Wurtz, & Schwarz, 1997; Xiao, Niu, Wiesner, & Huang, 2014), a nonlinear output explained by divisive normalization (Heeger, Simoncelli, & Movshon, 1996; Simoncelli & Heeger, 1998). Akin to models in V1, normalization models of MT responses assume that a linear weighting of relevant inputs (e.g., direction-selective inputs from V1) determines stimulus selectivity, whereas a pooled divisive inhibition determines overall response. The linear weighting of relevant V1 inputs allows the model to capture pattern-motion sensitivity (Movshon, Adelson, Gizzi, & Newsome, 1985; Rodman & Albright, 1989; Rust et al., 2006), in which MT responses reflect the overall motion of a stimulus (i.e., a plaid stimulus composed of two gratings) rather than the motion of individual components (i.e., individual gratings). The pooled normalization allows the model to capture suppressive effects such as sublinear additivity to multiple stimuli and suppression by non-preferred motions. Normalization also describes activity in ventral pathway brain regions including V4 and inferotemporal cortex (IT), where neurons respond to complex arrays of visual stimulus features and are thought to be crucial for object recognition. In these brain areas, feedforward pooling from earlier visual areas generate large receptive fields and responses that are generally invariant to stimulus size and position; however, neurons show selectivity for higher order features such as object identity. Consistent with normalization, responses to preferred stimuli are suppressed by non-preferred stimuli, with responses to pairs of objects reflecting the average (rather than the sum) of responses to individual objects alone (Kaliukhovich & Vogels, 2016; Ni et al., 2012; Reynolds & Desimone, 2003; Zoccolan, Cox, & DiCarlo, 2005). Ventral stream visual areas also exhibit significant modulation

by attention, and normalization has been proposed to play a key role in attentional control of neural responses (Boynton, 2009; Lee & Maunsell, 2009; Reynolds & Heeger, 2009); see below for further discussion under "Normalization in Higher Cognitive Processes."

While early work on normalization focused on visual processing, normalization also describes neural responses in other sensory modalities including olfaction, audition, and somatosensation. In olfaction, normalization captures the suppression of odorant-specific neural responses by other odorants and provides a mechanism for concentration-invariant olfactory coding. For example, in the *Drosophila* fruitfly, individual second order projection neurons (PNs) in an antennal lobe glomerulus receive feedforward excitatory inputs from a subset of olfactory receptor neurons, an organization that confers odorant selectivity. PNs also receive additional inhibitory lateral inputs from other glomeruli, providing a circuit mechanism for suppression with broad odor selectivity (Bargmann, 2006; Olsen & Wilson, 2008). Experiments independently manipulating direct and lateral glomerular input confirm that suppression mediates a divisive normalization operation (Olsen, Bhandawat, & Wilson, 2010); specifically, these responses can be described by Equation 4, with I and I_m representing feedforward and lateral activity. Because odorant concentrations can vary over large orders of magnitude, normalization in olfactory coding may be particularly important for maintaining stimulus selectivity and stabilizing perceived odor quality in a concentration-independent manner. In the zebrafish olfactory bulb, mitral cell odor responses are equalized across broad variations in input intensities, an effect driven by a dense network of local interneurons (Zhu, Frank, & Friedrich, 2013). Similar normalization mechanisms may mediate concentration-independent odor representations in the rat olfactory bulb, where both concentration-invariant odor identity and perceptual similarity are better predicted by normalized rather than raw measures of bulbar activity (Cleland, Johnson, Leon, & Linster, 2007). Thus, normalization appears to be a common feature of olfactory coding in both invertebrate and vertebrate systems, with direct drive generating stimulus selectivity and a broadly tuned suppression mediating contextual gain control. Analogous forms of normalization exist in auditory and tactile processing, as demonstrated by nonlinear forms of spectrotemporal contrast gain control in primary auditory cortex (Rabinowitz, Willmore, Schnupp, & King, 2011) and cross-digit suppression in human somatosensory cortex hemodynamic responses (Brouwer et al., 2015). In addition to the processing of information in individual sensory modalities, normalization describes key features of neural responses in multisensory integration (Ohshiro et al., 2011, 2017), suggesting an integral role at both early and late stages of sensory coding.

While most of the experiments demonstrating normalization have employed neurophysiological recording, functional neuroimaging studies provide additional evidence for normalization in human subjects. Because the normalization model is defined by the driving input and normalization pool specific to individual neurons (or neuron subpopulations), using functional magnetic resonance imaging (fMRI)—which measures the activity of a large number of neurons in a given voxel—poses technical challenges to testing the normalization model. One study addressed these challenges by examining intra-ocular interactions, taking advantage of the broad anatomical organization of monocular and binocular inputs in early visual cortex (Moradi & Heeger, 2009). This study found subadditive responses to binocular stimuli, consistent with an inter-ocular suppresion driven by divisive normalization. To examine normalization phenomena

occuring at a finer grain, later studies used a forward modeling approach that use voxel-wise biases to transform activity into predicted channel responses (Brouwer & Heeger, 2009; Kay, Naselaris, Prenger, & Gallant, 2008). These studies show that the normalization model captures cross-orientation suppression in human visual cortex (Brouwer & Heeger, 2011) and cross-digit suppression in human somatosensory cortex (Brouwer et al., 2015). Together, these results reinforce the hypothesis that normalization is a widely prevalent computation in sensory processing.

Normalization in Higher Cognitive Processes

While most examples of normalization occur in early sensory processing, 21st-century work suggests that normalization extends to higher order cognitive processes (Lee & Maunsell, 2009; Louie & Glimcher, 2012; Louie et al., 2011; Ohshiro et al., 2011, 2017; Reynolds & Heeger, 2009). One example is multisensory integration, evident in brain regions including the superior colliculus and dorsal medial superior temporal area (MSTd), which can enhance the perceptual detection and discrimination of environmental events (Stein, Stanford, & Rowland, 2014). In multisensory integration, neural responses to inputs from multiple sensory modalities display characteristic nonlinearities in how that information is combined. Many of these nonlinearities arise naturally from a divisive normalization model of multisensory integration (Ohshiro et al., 2011). In this normalization model, individual neuron responses depend on a weighted sum of unisensory inputs (feedforward drive, in the numerator) and a summation over a large population of multisensory neurons (pooled normalization signal, in the denominator). These (possibly asymmetric) dominance weights are fixed for a given neuron but vary across the population of neurons, producing for each neuron a modality-specific excitation and a modality-general suppression. This model reproduces several characteristic empirical principles, including stronger multisensory enhancement with weak versus strong inputs (principle of inverse effectiveness) and the requirement for spatial and temporal congruence of sensory signals to generate enhancement versus suppression (spatial/temporal principle). Furthermore, the normalization model predicts that a non-preferred sensory input from one modality, which is excitatory when presented alone, should suppress the response to a preferred input from another modality when cues are combined. This form of cross-modal suppression was subsequently verified in monkey MSTd, where neurons integrate visual and vestibular information for self-motion (Ohshiro et al., 2017).

Normalization has also been proposed to explain the attentional modulation of neural responses in visual brain areas (Boynton, 2009; Ghose & Maunsell, 2008; Lee & Maunsell, 2009; Reynolds, Chelazzi, & Desimone, 1999; Reynolds & Heeger, 2009). Attention produces a variety of effects on neural responses such as changes in contrast gain (lateral shifts in contrast response functions), changes in response gain (multiplicative gain changes in contrast response functions), and changes in feature tuning; these various experimental findings have led to multiple, alternative theories of attention (Desimone & Duncan, 1995; Martinez-Trujillo & Treue, 2004; McAdams & Maunsell, 1999; Moran & Desimone, 1985; Reynolds, Pasternak, & Desimone, 2000; Treue & Martinez Trujillo, 1999). Many of these diverse experimental results and theories can be explained by normalization-based models of attention. In the most well-known model (Reynolds

& Heeger, 2009), neural responses are determined by three primary components: a stimulation field, a suppressive field, and an attention field. The stimulation field characterizes a neuron's selectivity for stimulus information (e.g., spatial position and orientation), and captures theoretical feedforward driven responses in the absence of attentional and suppressive effects. Similar to earlier normalization models of V1, the suppressive field characterizes the typically broadly tuned stimuli that drive divisive suppression and effectively normalizes the response of one neuron by the activity of a large pool of neurons. The attention field characterizes the effect of attention as a function of spatial position and feature information (e.g., orientation) and operates multiplicatively on the stimulus drive before normalization. Critically, because attention affects both the stimulus drive to a given neuron and the drive to neurons that constitute the suppressive pool, attentional modulation influences both the numerator and denominator in the normalization model. Thus, in this normalization model, attention reshapes the distribution of activity across the population of visual neurons by controlling the relative levels of excitation and suppression.

The normalization model of attention derives much of its explanatory power from the relative balance of its three components (stimulus drive, suppressive drive, and attentional modulation), and their dependence on the size of the visual stimulus and the extent of attention (Reynolds & Heeger, 2009). For example, as in simpler normalization models of visual responses without attention (Carandini et al., 1997; Cavanaugh et al., 2002a; Heeger, 1992), stimulus size controls the relative balance between excitation and suppression: large stimuli encompass a neuron's stimulation field and the larger suppressive field, and drive both excitation and suppression, while small stimuli drive strong excitation but relatively weak suppression. In addition, because the attention field is also defined as a function of spatial position and visual features (e.g. orientation), modulatory effects on neural responses will depend on the spatial extent and featural selectivity of attention. This flexibility allows the normalization model to capture different experimental results of attention. When the stimulus is small and the attention field is large, attention affects the normalization numerator and denominator equally, and the model predicts changes in contrast gain. When the stimulus is large and the attention field is small, attention primarily affects stimulus drive, and the model predicts changes in response gain. When the normalization model is adjusted for specifics of experimental implementation, it explains changes in contrast gain (Martinez–Trujillo & Treue, 2002; Reynolds et al., 2000), response gain (McAdams & Maunsell, 1999; Treue & Martinez Trujillo, 1999), and results intermediate between the two (Williford & Maunsell, 2006). Furthermore, variability in the effect of attention across different neurons may be related to variability in individual neuron normalization, in particular the strength of tuned normalization (Lee & Maunsell, 2009; Ni & Maunsell, 2017; Ni et al., 2012; Ray, Ni, & Maunsell, 2013; Verhoef & Maunsell, 2017). In addition to capturing the effects of attention within brain areas, normalization can also explain attentional effects across brain areas, such as the attention-driven increase in correlated variability between V1 and MT (Ruff, Alberts, & Cohen, 2016; Ruff & Cohen, 2017).

In addition to higher order functions in sensory processing, normalization also plays a key role in decision-related neural activity (Louie & Glimcher, 2012; Louie et al., 2015; Rangel & Clithero, 2012). In particular, normalization appears to define the form of action value information

represented in motor planning circuits, introducing a contextual value coding dependent on the composition of the choice set. For example, in the monkey posterior parietal cortex, neural activity in the lateral intraparietal area (LIP) specifies individual saccadic eye movements (Gnadt & Andersen, 1988; Snyder, Batista, & Andersen, 1997) and covaries with the values associated with those actions (Platt & Glimcher, 1999). This value coding is robust to behavioral requirements and task design, occurring whether value is determined by sensory information, reward magnitude, delay discounting, foraging requirements, social information, or strategic interactions (Churchland, Kiani, & Shadlen, 2008; Dorris & Glimcher, 2004; Klein, Deaner, & Platt, 2008; Louie & Glimcher, 2010; Platt & Glimcher, 1999; Roitman & Shadlen, 2002; Shadlen & Newsome, 2001; Sugrue, Corrado, & Newsome, 2004). Because LIP neurons show little sensory or motor-related spiking for spatial locations outside their response field (RF), studies have traditionally viewed LIP decision activity in terms of RF saccadic action value. However, recent work shows that action value coding in LIP depends on both the value of the saccade target in the RF and the value of alternative, extra-RF targets (Louie et al., 2011; Rorie, Gao, McClelland, & Newsome, 2010). Specifically, increases in the value of extra-RF targets suppresses LIP activity elicited by a given RF target; when no target is available in the RF, extra-RF target values reduce LIP activity below baseline levels. Suppression by extra-RF target value, despite these cues eliciting little spiking activity when presented alone, represents a nonlinear contextual modulation analogous to inhibition by nonpreferred stimuli in visual suppressive effects (see "Normalization in Sensory Processes"). This contextual value modulation is well-described by a simple divisive normalization function:

$$R_i = R_{\max} \frac{V_i + \beta}{\sigma + \sum_j V_j}$$

(5)

where the saccade-selective activity of an LIP neuron R_i depends on the value V_i of the RF target, divided by a term including the summed value of all available targets (Louie et al., 2011). As in sensory forms of normalization, the empirical parameters R_{max}, σ, and β govern the maximal level of activity, saturation behavior, and baseline firing rates in model responses. This simple value normalization model outperforms alternative value representations (e.g., absolute value or value difference models) in explaining single neuron and population LIP responses, implementing a context-dependent neural representation of value. Consistent with a role in decision making, this contextual value representation is linked to spatial and temporal context-dependent preferences in monkey and human choice behavior (Khaw et al., 2017; Louie et al., 2013) (see "Evidence for Normalization in Decision Making").

Recent studies suggest that relative value coding, accounting for the context defined by the choice set and implemented via divisive normalization, is likely a general feature of decision-related neural processing. In parietal cortex, LIP activity similarly reflects the relative value of choice options whether value is determined by simple reward outcomes (Louie et al., 2011; Rorie et al., 2010) or more complicated foraging (Sugrue et al., 2004) or game-theoretic (Dorris & Glimcher, 2004) interactions. Similar to parietal activity during saccadic decision making, reach selective

neurons in monkey dorsal premotor cortex encode a relative value signal incorporating the rewards of both preferred and nonpreferred arm movements (Pastor-Bernier & Cisek, 2011). Furthermore, relative value coding is consistent with previously described effects of choice set size and target uncertainty on motor output structures such as the superior colliculus (Basso & Wurtz, 1997, 1998), suggesting that normalized value coding occurs in multiple brain regions associated with action selection. Evidence also suggests that the brain implements relative valuation signals apart from action selection circuits. For example, in the monkey medial orbitofrontal cortex, the relative values of risky and safe options in a lottery choice task are encoded via a divisive normalization representation (Yamada, Louie, Tymula, & Glimcher, 2018); similar normalized value signals during risky choice are also observed in human prefrontal cortex hemodynamic responses (Holper et al., 2017). In contrast to value normalization in action selection circuits, which is driven by the spatial context defined by the choice set, this valuation in frontal brain regions may reflect a relative comparison to the temporal context defined by the recent past (Cox & Kable, 2014; Kobayashi, Pinto de Carvalho, & Schultz, 2010; Padoa-Schioppa, 2009; Tremblay & Schultz, 1999).

Normalization in Behavior

Evidence for Normalization in Perception

While normalization primarily quantifies information coding at the level of neurons and neural populations, a separate question of interest is the influence of the normalization algorithm on behavior. Because much of the empirical electrophysiological literature focuses on normalization in early visual pathways, psychophysical studies of perceptual context effects provide indirect evidence for how normalization contributes to visual perception. Such studies generally make predictions about the normalized coding of sensory information, adopt a decision rule to transform normalized information into choice, and compare model predictions to empirical choice behavior. For example, normalization models of surround effects in V1 neurons explain aspects of human perception, including simultaneous contrast effects (Xing & Heeger, 2001), orientation and spatial frequency dependence (Chubb, Sperling, & Solomon, 1989; Solomon, Sperling, & Chubb, 1993), and the timing of suppressive effects (Petrov, Carandini, & McKee, 2005). Divisive normalization has also been proposed to describe the intracortical interactions governing visual salience and explain bottom-up perceptual phenomena such as visual popout and visual search asymmetries (Coen-Cagli et al., 2012; Gao & Vasconcelos, 2009; Itti & Koch, 2000; Li, 2002). In addition, normalization models in sensory processing have been proposed to control the competitive interactions that determine visual awareness (Li, Carrasco, & Heeger, 2015; Ling & Blake, 2012); this competition may be regulated by attention, which can itself be described by a normalization mechanism (Boynton, 2009; Ghose & Maunsell, 2008; Lee & Maunsell, 2009; Reynolds et al., 1999; Reynolds & Heeger, 2009). Behavior consistent with normalized sensory coding also occurs in other sensory modalities. In the *Drosophila* fruitfly, olfactory behavior is well described by a decoding model involving normalization at the level of individual glomeruli followed by linear summation over glomerular channels; in addition to normal olfactory behavior, the normalization model accurately predicts behavioral responses to

the silencing of specific glomerular channels (Badel, Ohta, Tsuchimoto, & Kazama, 2016). While these changes in olfactory preference arise from normalization early in sensory processing, they suggest that changes in choice behavior may reflect multiple normalization computations at both the sensory coding and decision stages.

In addition to behavioral changes linked to normalization in early stages of sensory processing, studies suggest that normalization occurs in more abstract, higher order perceptual representations. One example is that of numerical quantity, which is represented by neural activity in the monkey posterior parietal and prefontal cortices (Nieder, Freedman, & Miller, 2002; Nieder & Miller, 2003, 2004; Roitman, Brannon, & Platt, 2007). While normalization has not been explicitly examined in numerosity coding neural activity, normalization appears to play a role in how monkeys combine number information in behavior (Livingstone et al., 2014). Monkeys trained to combine different symbolically represented magnitudes show a subadditive addition, consistent with a relative evaluation process of number information well described by a normalization model. Human estimates of facial attractiveness also demonstrate a dependence on contextual factors consistent with a normalization process (Furl, 2016). When human subjects are asked to select the most attractive of three presented faces, the relative preference between the two most attractive faces decreases with the attractiveness of the third, least attractive face. This contextual effect of attractiveness is consistent with a divisive mechanism where the attractiveness of each option is scaled by the summed attractiveness of presented faces, suggesting that normalization plays a role in high level perception and social evaluation. Note that these preferences changes are conceptually similar to changes in value based decision making (Louie et al., 2013), and may reflect normalization processes that overlap between higher order perception and decision processes (see the section "Evidence for Normalization in Decision Making").

Evidence for Normalization in Decision Making

For valuation and decision processes, normalized value coding instantiates a comparative form of valuation, in which potential actions are represented relative to other available alternatives. This contextual modulation of value coding, particularly in action selection circuits thought to implement decision making, has implications for theoretical models of choice behavior and context-dependent preferences (Louie & Glimcher, 2012; Louie et al., 2015; Tymula & Plassmann, 2016). Traditional normative theories of rational choice in economics, ecology, and psychology assume that decisions depend solely on the absolute values of individual choice options (Stephens & Krebs, 1986; Von Neumann & Morgenstern, 1944). However, in contrast to these normative models, empirical choice behavior in a wide range of species exhibits context-dependence, where preference between any two options can depend markedly on additional alternatives (Bateson, Healy, & Hurly, 2003; Huber, Payne, & Puto, 1982; Shafir, Waite, & Smith, 2002; Simonson, 1989; Tversky, 1972; Tversky & Simonson, 1993). Cognitive models of context-dependent preferences —which are irrational from the perspective of normative choice models—have been proposed, but the underlying neural mechanisms are unknown.

Because normalization scales a given option value (in the numerator) by a term incorporating the value of other options (in the denominator), normalized value coding naturally generates context-dependent choice behavior. Unlike most existing examples of context-dependent preferences, normalization predicts contextual dependence based on the integrated value of options rather than specific attribute levels and option relationships in a multi-attribute space (Louie et al., 2015; Louie et al., 2013). Specifically, the divisive nature of normalization predicts that increasing the overall (summed) value of the choice set will decrease the neural firing rates representing option values. Given this suppression and the existence of noise in the decision process, contextual modulation will preserve the value-ranked order of options but impact the discriminability between options. At the level of behavior, simple decision models predict that the relative preference between two options in a trinary choice scenario will decrease as the value of a third option increases. This effect, which can occur despite the third option never being selected, violates a common assumption in rational choice theory known as independence from irrelevant alternatives (Luce, 1959). Empirical choice data shows that this form of context-dependent choice behavior occurs in both monkey and human decision-making behavior (Itthipuripat, Cha, Rangsipat, & Serences, 2015; Louie et al., 2013). Such context-dependence, induced by the configuration of the choice set, represents a form of spatial context-dependence analogous to surround suppression phenomena in sensory processing.

In addition, recent evidence suggests that value normalization is linked to temporal context effects in behavior. When human subjects are asked to report their subjective valuation for different food items, individual valuations for the same items depend systematically on the recent history of presented values: valuations are suppressed and enhanced by a history of recent high and low value items, respectively (Khaw et al., 2017). This temporal dependence can be explained by a normalized valuation model incorporating past value information into the denominator. Computationally, the effect of past information can be modeled as an effective change in the semisaturation term, consistent with normalization models of adaptation in sensory neural responses (Heeger, 1992; Sinz & Bethge, 2013; Sit et al., 2009). Normalization-mediated value adaptation may underlie a number of well-known behavioral phenomena, including successive incentive contrast effects (Crespi, 1942; Flaherty, 1982; Zeaman, 1949) and reference-dependent economic choice (Kahneman & Tversky, 1979; Koszegi & Rabin, 2006, 2007).

Further work will be required to establish the generality of the normalization computation in neural valuation and decision-making processes. One open question is the role of normalization in multi-attribute choice: While the value normalization model explains contextual choice phenomena that depend on option values, the majority of empirical context effects involves choices between options that differ along multiple attribute dimensions. Many multi-attribute context effects can be explained by computational models employing forms of normalization at the attribute coding level (Hunt, Dolan, & Behrens, 2014; Soltani, De Martino, & Camerer, 2012). Further research will have to examine whether attribute normalization occurs in neural activity, the anatomical locus of such computations, and how different valuation-related normalization processes are integrated in the choice process. A second question is the relationship between divisive normalization and related models of context-dependent behavior. For example, value-related neural activity adapts to the range of recent rewards in monkey orbitofrontal cortex

(Kobayashi et al., 2010; Padoa-Schioppa, 2009) and related human brain regions (Cox & Kable, 2014), and range normalization has been proposed to explain multi-attribute context effects (Soltani et al., 2012). In addition to divisive models, contextual choice effects can also be explained by subtractive models that implement a precision-weighted prediction error based on normative Bayesian theory (Rigoli, Friston, et al., 2016; Rigoli, Mathys, Friston, & Dolan, 2017; Rigoli, Rutledge, Dayan, & Dolan, 2016). More broadly, a relative comparison of attribute information is an integral component of prominent dynamic models of context-dependent choice behavior (Bogacz, Usher, Zhang, & McClelland, 2007; Roe, Busemeyer, & Townsend, 2001; Trueblood, Brown, & Heathcote, 2014). Whether these alternative models, and the behavior they capture, fit within the normalization framework will require further theoretical and experimental work, particularly research identifying the underlying neural mechanisms governing context-dependent choice.

Normalization and Clinical Disorders

In addition to explaining aspects of perceptual and decision-making behavior, altered divisive normalization and related computational functions are linked to a number of clinical disorders. These studies are part of the larger approach of computational psychiatry, a recent concerted effort to identify computational modeling frameworks for psychiatric disorders (Huys, Maia, & Frank, 2016; Montague, Dolan, Friston, & Dayan, 2012; Wang & Krystal, 2014). Computational deficits linked to aberrant normalization are suggested by altered sensory processing in diseases including epilepsy (Porciatti, Bonanni, Fiorentini, & Guerrini, 2000; Tsai, Norcia, Ales, & Wade, 2011), major depression (Bubl et al., 2010; Golomb et al., 2009; Norton et al., 2016), schizophrenia (Butler, Silverstein, & Dakin, 2008; Butler et al., 2005; Tadin et al., 2006), and autism (Dakin & Frith, 2005; Flevaris & Murray, 2014; Foss-Feig, Tadin, Schauder, & Cascio, 2013; Robertson et al., 2013). In many of these cases, symptomology includes behavioral deficits in contextual phenomena such as surround suppression, suggesting that different genetic-, synaptic-, and circuit-level etiologies may manifest as abnormalities in neural gain control. While these aberrations in gain control only generally imply a relationship to normalization, recent theoretical work has proposed a specific role for normalization in autism; specifically, reduced normalization in neural processing may underlie the sensory, cognitive, and social symptoms in autism pathology (Rosenberg, Patterson, & Angelaki, 2015). This normalization model is driven by the hypothesis that autism involves a disrupted balance of neurophysiological excitatory and inhibitory (E/I) activity (Heeger, Behrmann, & Dinstein, 2017; Rubenstein & Merzenich, 2003; Yizhar et al., 2011). Genetic, biochemical, and animal model studies suggest an increased E/I ratio in autism, which can be modeled as reduced normalization via an increase in excitatory drive (numerator), a decrease in suppressive drive (denominator), or both. Network simulations with reduced normalization replicate a number of findings in autism, including altered perception and statistical inference about the sensory environment; furthermore, reduced normalization is broadly consistent with documented autism consequences involving local versus global processing, multisensory integration, and decision making (Rosenberg et al., 2015). While the reduced normalization is consistent with many symptoms of autism, there is considerable

heterogeneity in underlying etiologies and clinical presentation (Heeger et al., 2017; Mullins, Fishell, & Tsien, 2016), and the prevalence, extent, and causal role of altered normalization in autism are currently unclear.

Computational Functions

Given the different generative mechanisms behind normalization (see section "Biophysical Implementation") and its widespread implementation in different brain areas, cognitive processes, and species (see section "Normalization in Neural Responses"), an important question is whether divisive normalization serves a single unifying function or different, process-specific functions. Theoretical work has proposed a number of functions for normalization, including maximizing sensitivity, invariant coding, discrimination, marginalization, winner-take-all competition, and redundancy reduction (Carandini & Heeger, 2012). Two of these proposed functions—efficient coding and marginalization—are discussed further below, though it is important to note that many of these functions are related.

Efficient Coding

A prominent hypothesis in sensory processing is that neural systems face intrinsic information capacity constraints and compensate with strategies to maximize coding efficiency (Barlow, 1961). For single neurons with a constrained range of activity, efficient coding predicts that the distribution of responses in a given environment should be uniform over that range; in other words, each possible level of neural activity will be used equally (histogram equalization). Normalization mediates a number of sensory phenomena, such as light adaptation and contrast gain control, which adapt neural response functions in this manner to the distribution of sensory inputs. Similar normalization-driven adaptation occurs in neural value coding and choice behavior, suggesting that efficient coding principles extend to valuation and decision making (Louie & Glimcher, 2012; Louie et al., 2015; Rangel & Clithero, 2012). Because the natural sensory environment contains widespread statistical regularities (Geisler, 2008; Simoncelli & Olshausen, 2001), producing redundant information in sensory inputs, sensory systems can further maximize coding efficiency by reducing redundancies in their outputs. For a population of neurons, redundancy reduction requires that neural responses should be as statistically independent as possible. While linear response properties (e.g., receptive field structure) can remove redundancy driven by low order statistics in natural signals (e.g., spatial correlation in intensity), they cannot remove higher order statistical dependencies in neural responses (Bethge, 2006; Simoncelli & Olshausen, 2001). However, these statistical dependencies in natural signals can be significantly reduced by divisive normalization (Lyu, 2011; Schwartz & Simoncelli, 2001). Furthermore, normalization models—with weights (in the divisive denominator) optimized for independence in natural signals—reproduce characteristic response properties of both visual and auditory neurons (Averbeck et al., 2006; Schwartz & Simoncelli, 2001). Similar principles operate in olfaction: normalization in the *Drosophila* antennal lobe decorrelates neural responses to odors, increasing statistical independence (Luo, Axel, & Abbott, 2010; Olsen et al., 2010).

Marginalization

Many different types of neural processes involve a form of probabilistic inference called marginalization, which recovers the probability distribution of a given variable from a joint probability distribution by integrating out (marginalizing over) the other variables. Neural Computations implementing marginalization are thought to be involved in diverse tasks including olfaction (Grabska-Barwinska et al., 2017; Olsen & Wilson, 2008), object recognition (DiCarlo, Zoccolan, & Rust, 2012), coordinate transformation (Pouget & Snyder, 2000), motor control (Wolpert, Ghahramani, & Jordan, 1995), decision making (Beck et al., 2008), and causal reasoning (Blaisdell, Sawa, Leising, & Waldmann, 2006; Griffiths & Tenenbaum, 2009). While marginalization is a conceptually straightforward process, its implementation in neural circuits faces a number of challenges: the need to represent probabilistic information, information coding in neural population activity, and a potentially large number of nuisance variables. However, recent theoretical work shows that biologically plausible circuits with divisive normalization can achieve near-optimal marginalization for processes including coordinate transformations, object tracking, simplified olfaction, and causal reasoning (Beck et al., 2011). Normalization is specifically important when neural activity represents probabilistic information in a logarithmic form (e.g., under probabilistic population codes) (Beck et al., 2008; Ma, Beck, Latham, & Pouget, 2006), since such codes provide for the simple multiplication of probabilities rather than the addition of probabilities required for marginalization; it should be noted that normalization is less advantageous under alternative coding models where probabilities are directly represented (Anastasio, Patton, & Belkacem-Boussaid, 2000; Lee & Mumford, 2003). Normalization-based circuits achieve marginalization while representing the full probability distribution of the encoded variable, a critical requirement for probabilistic inference (Pouget, Beck, Ma, & Latham, 2013). Given its role in a wide range of neural processes, marginalization—and its implementation via normalization—offers a potential reason for the apparent ubiquity of divisive normalization in the brain.

Biophysical Implementation

Apart from computational and coding aspects of normalization, considerable research has focused on potential biophysical and network mechanisms responsible for normalization computations. Given its ubiquity across different brain regions, hierarchical processing level, and species, it is likely that multiple mechanisms can generate normalization-like computations in different neural systems (Carandini & Heeger, 2012). Moreover, even within an individual system, normalization may arise from multiple mechanisms working in concert and across different stages.

One area of focus has been the biophysical mechanism responsible for divisive gain control in normalization. Given the suppressive nature of contextual drive in normalization, many studies have focused on the role of synaptic inhibition. Early work hypothesized a particular role for shunting inhibition (Carandini & Heeger, 1994; Carandini et al., 1997), which affects membrane conductance rather than voltage and has a divisive effect of excitatory potential amplitudes (Silver, 2010). Shunting inhibition is consistent with some forms of normalization the retina and

primary visual cortex but increases in membrane conductance and the strength of normalization are not always coupled in the manner predicted by shunting inhibition (Carandini & Heeger, 2012). More generally, inhibition is clearly linked to normalization in some systems, including olfactory circuits in the *Drosophila* antennal lobe (Olsen et al., 2010; Olsen & Wilson, 2008) and zebrafish olfactory bulb (Zhu et al., 2013). However, some normalization-mediated phenomena— including contrast saturation, cross-orientation suppression, and surround suppression in visual cortex—appear to be unaffected by GABA$_A$ receptor blockade (Katzner, Busse, & Carandini, 2011; Ozeki et al., 2004), suggesting that they do not rely on inhibition. As an alternative to inhibition, other work has suggested that normalization arises from a decrease in excitation. Employing optogenetic activation and intracellular recording, recent work shows that normalization driven by distal inputs to mouse visual cortex is mediated by decreased synaptic excitation (Sato, Haider, Hausser, & Carandini, 2016). These results are consistent with network models with strong local recurrence, which posit that changes in excitation and inhibition may occur together in normalization (Ozeki et al., 2009; Rubin, Van Hooser, & Miller, 2015; Shushruth et al., 2012). In such circuits, excitation and inhibition are thought to be tightly coupled, and inputs that drive normalization (e.g., surround stimuli in surround suppression) act by suppression of local rather than afferent excitatory drive. Notably, different biophysical normalization mechanisms may be linked to differences in circuit architecture: increased inhibition plays a role in simpler circuits without strong recurrent connectivity (Olsen et al., 2010; Olsen & Wilson, 2008; Zhu et al., 2013), whereas reduced excitation (via E/I balance mechanisms) seems more prevalent in dense, recurrent circuits in cortical areas.

Related to the question of biophysical mechanism is the question of the network architectures that can generate divisive normalization. At its core (see section "The Divisive Normalization Model"), normalization involves an interaction between afferent drive (numerator) and suppressive modulatory control (denominator). When driving and suppressive input show similar selectivity and origin (e.g., for the same region of visual space, as in cross-orientation suppression), normalization can be potentially explained by synaptic mechanisms such as synaptic depression (Abbott, Varela, Sen, & Nelson, 1997; Carandini, Heeger, & Senn, 2002). Such forms of normalization may also be implemented by a feedforward inhibition circuit, where driving and suppressive inputs arise from the same upstream brain area. Because both inputs are feedforward, an attractive characteristic of this circuit mechanism is that these signals are not yet normalized, as described in the standard normalization equation (Carandini & Heeger, 2012). However, normalization can also be implemented by a feedback circuit, with suppressive input arriving from lateral or top-down afferents (Carandini & Heeger, 1994; Carandini et al., 1997; Heeger, 1992). A feedback circuit for normalization is generally assumed to play a role in cortical circuits, which display extensive recurrent connectivity; feedback circuits may play a particularly important role when suppressive signals in the normalization pool represent a larger stimulus context than the driving input (e.g., surround suppression). Empirical evidence suggests that both feedback and feedforward circuits can generate divisive normalization, and that the circuit organization is different in different systems. For example, feedforward inhibition via lateral connections generates normalization in the fruitfly olfactory system (Olsen et al., 2010; Olsen & Wilson, 2008). In cortical circuits, such as the mammalian visual cortex, feedback inhibition via lateral or top-down inputs appears to play a larger role in normalization-mediated gain control (Angelucci & Bressloff, 2006; Angelucci, Levitt, & Lund, 2002; Carandini & Heeger, 1994;

Carandini et al., 1997; Nassi, Gomez-Laberge, Kreiman, & Born, 2014; Nassi, Lomber, & Born, 2013; Reynaud, Masson, & Chavane, 2012). Recent optogenetic manipulations studies suggest that local circuitry play a causal role in normalization, consistent with a feedback mechanism (Nassi et al., 2015; Sato, Hausser, & Carandini, 2014). For example, superficial layer somatostatin-expressing inhibitory neurons, likely driven by horizontal inputs, contribute to surround suppression in the rodent visual cortex (Adesnik, Bruns, Taniguchi, Huang, & Scanziani, 2012; Adesnik & Scanziani, 2010); an important open question is whether this circuit mechanism generalizes to other species. Ultimately, the issues of biophysical and circuit mechanisms of normalization are tightly linked and underscore the variety of implementations used by different systems to generate the normalization computation.

Future Directions

Given its ubiquity in neural information coding, the divisive normalization model continues to be applied in a growing number of different directions. One area of active research is extending the standard formalization of the normalization model, with a focus on denominator normalization weights that actively adjust to the environment. In most standard applications of the normalization model, normalization weights are either assumed to be equal for all suppressive inputs or asymmetric but fixed (often fit to empirical data). However, recent work suggests that normalization can be implemented in a flexible manner consistent with a dynamic adjustment of normalization weights. For example, surround suppression in visual cortical neurons can vary substantially depending on the image; this variability is not explained by standard models but can be captured by a normalization model where the strength of divisive suppression depends on the statistical homogeneity in the image (Coen-Cagli et al., 2015). This gating by sensory statistics is consistent with the notion that normalization-mediated processes such as surround suppression are important for efficient coding, removing regularity induced redundancies in sensory information (Coen-Cagli et al., 2012; Schwartz & Simoncelli, 2001; Vinje & Gallant, 2002). Dynamic adjustment of normalization weights have also been proposed to explain adaptation in visual cortical neurons, where extended presentation of oriented stimuli produces suppression for neurons tuned for the adapter and a repulsive shift of tuning curves away from the adapter (Snow, Coen-Cagli, & Schwartz, 2016; Wainwright, Schwartz, & Simoncelli, 2002; Westrick et al., 2016). These characteristic changes are explained by a normalization model with dynamically changing weights, using a simple Hebbian learning rule to adjust the strength of normalization between neurons based on past stimulus-driven response patterns.

Finally, while the focus of this article has been on the normalization computation in neuroscience, there is growing application of divisive normalization principles in software engineering. One ongoing application is image quality assessment: quantifying image distortion in a manner similar to subjective human evaluation is crucial to image compression, rendering, and enhancement. Because normalization captures nonlinearities intrinsic to biological vision, normalization-based algorithms have shown promise as a means to predict subjective image distortion (Laparra, Berardino, Balle, & Simoncelli, 2017; Laparra, Munoz-Mari, & Malo, 2010; Lyu & Simoncelli, 2008; Malo, Epifanio, Navarro, & Simoncelli, 2006; Teo & Heeger, 1994). Another developing application of normalization is in the engineering of deep neural networks—

artificial neural networks with multiple hidden layers between input and output—increasingly used in machine learning and artificial intelligence (Hassabis, Kumaran, Summerfield, & Botvinick, 2017; LeCun, Bengio, & Hinton, 2015). General forms of normalization, such as batch normalization (Ioffe & Szegedy, 2015) and layer normalization (Ba, Kiros, & Hinton, 2016), are widely employed in deep networks to increase network stability, speed learning, and improve performance. More recent approaches have begun to employ forms of divisive normalization, where unit activations are normalized by neighboring activations within a layer (Jarrett, Kavukcuoglu, & LeCun, 2009; Krizhevsky, Sutskever, & Hinton, 2012), and generalizing and improving divisive normalization algorithms in deep networks continues to be an active area of research (Giraldo & Schwartz, 2018; Ren, Liao, Urtasun, Sinz, & Zemel, 2016). While current applications of divisive normalization in deep networks are primarily focused on improving performance, and thus represents the use of biological principles to inform engineering, the examination of regimes in which neural networks can match neurophysiological data and human performance may also inform a deeper understanding of neural organization and function (Yamins & DiCarlo, 2016; Yamins et al., 2014).

Acknowledgments

References

Abbott, L. F., Varela, J. A., Sen, K., & Nelson, S. B. (1997). Synaptic depression and cortical gain control. *Science*, *275*(5297), 220–224.

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, *2*(2), 284–299.

Adesnik, H., Bruns, W., Taniguchi, H., Huang, Z. J., & Scanziani, M. (2012). A neural circuit for spatial summation in visual cortex *<http://dx.doi.org/10.1038/nature11526>*. *Nature*, *490*(7419), 226–231.

Adesnik, H., & Scanziani, M. (2010). Lateral competition for cortical space by layer-specific horizontal circuits *<http://dx.doi.org/10.1038/nature08935>*. *Nature*, *464*(7292), 1155–1160.

Adrian, E. D. (1926). The impulses produced by sensory nerve endings: Part I. *Journal of Physiology*, *61*(1), 49–72.

Adrian, E. D., & Zotterman, Y. (1926). The impulses produced by sensory nerve-endings: Part II. The response of a Single End-Organ. *Journal of Physiology*, *61*(2), 151–171.

Aertsen, A. M., & Johannesma, P. I. (1981). The spectro-temporal receptive field. A functional characteristic of auditory neurons. *Biological Cybernetics*, *42*(2), 133–143.

Albrecht, D. G., & Geisler, W. S. (1991). Motion selectivity and the contrast-response function of simple cells in the visual cortex. *Visual Neuroscience*, *7*(6), 531–546.

Albrecht, D. G., & Hamilton, D. B. (1982). Striate cortex of monkey and cat: contrast response function. *Journal of Neurophysiology*, *48*(1), 217–237.

Anastasio, T. J., Patton, P. E., & Belkacem-Boussaid, K. (2000). Using Bayes' rule to model multisensory enhancement in the superior colliculus. *Neural Computation*, *12*(5), 1165–1187.

Andersen, R. A., Hwang, E. J., & Mulliken, G. H. (2010). Cognitive neural prosthetics *<http://dx.doi.org/10.1146/ annurev.psych.093008.100503>*. *Annual Review of Psychology*, *61*, 169–190, C161–163.

Angelucci, A., & Bressloff, P. C. (2006). Contribution of feedforward, lateral and feedback connections to the classical receptive field center and extra-classical receptive field surround of primate V1 neurons *<http://dx.doi.org/10.1016/ S0079-6123(06)54005-1>*. *Progress in Brain Research*, *154*, 93–120.

Angelucci, A., Levitt, J. B., & Lund, J. S. (2002). Anatomical origins of the classical receptive field and modulatory surround field of single neurons in macaque visual cortical area V1. *Progress in Brain Research*, *136*, 373–388.

Averbeck, B. B., Latham, P. E., & Pouget, A. (2006). Neural correlations, population coding and computation *<http:// dx.doi.org/10.1038/nrn1888>*. *Nature Reviews Neuroscience*, *7*(5), 358–366.

Ba, J. L., Kiros, J. R., & Hinton, G. E. (2016). Layer normalization. *arXiv preprint arXiv:1607.06450*.

Baccus, S. A., & Meister, M. (2002). Fast and slow contrast adaptation in retinal circuitry. *Neuron*, *36*(5), 909–919.

Badel, L., Ohta, K., Tsuchimoto, Y., & Kazama, H. (2016). Decoding of context-dependent olfactory behavior in drosophila *<http://dx.doi.org/10.1016/j.neuron.2016.05.022>*. *Neuron*, *91*(1), 155–167.

Bargmann, C. I. (2006). Comparative chemosensation from receptors to ecology *<http://dx.doi.org/10.1038/ nature05402>*. *Nature*, *444*(7117), 295–301.

Barlow, H. B. (1961). Possible principles underlying the transformation of sensory messages. In W. A. Rosenblith (Ed.), *Sensory communication*. Cambridge, MA: MIT Press.

Basso, M. A., & Wurtz, R. H. (1997). Modulation of neuronal activity by target uncertainty *<http://dx.doi.org/ 10.1038/37975>*. *Nature*, *389*(6646), 66–69.

Basso, M. A., & Wurtz, R. H. (1998). Modulation of neuronal activity in superior colliculus by changes in target probability. *Journal of Neuroscience*, *18*(18), 7519–7534.

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., & Friston, K. J. (2012). Canonical microcircuits for predictive coding *<http://dx.doi.org/10.1016/j.neuron.2012.10.038>*. *Neuron*, *76*(4), 695–711.

Bateson, M., Healy, S. D., & Hurly, T. A. (2003). Context-dependent foraging decisions in rufous hummingbirds. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *270*(1521), 1271–1276.

Beck, J. M., Latham, P. E., & Pouget, A. (2011). Marginalization in neural circuits with divisive normalization *<http:// dx.doi.org/10.1523/JNEUROSCI.1706%9611.2011>*. *Journal of Neuroscience*, *31*(43), 15310–15319.

Beck, J. M., Ma, W. J., Kiani, R., Hanks, T., Churchland, A. K., Roitman, J., . . . Pouget, A. (2008). Probabilistic population codes for Bayesian decision making _<http://dx.doi.org/10.1016/j.neuron.2008.09.021>_. *Neuron*, *60*(6), 1142–1152.

Bethge, M. (2006). Factorial coding of natural images: how effective are linear models in removing higher-order dependencies? *Journal of the Optical Society of America A*, *23*(6), 1253–1268.

Blaisdell, A. P., Sawa, K., Leising, K. J., & Waldmann, M. R. (2006). Causal reasoning in rats _<http://dx.doi.org/10.1126/science.1121872>_. *Science*, *311*(5763), 1020–1022.

Blakemore, C., & Tobin, E. A. (1972). Lateral inhibition between orientation detectors in the cat's visual cortex. *Experimental Brain Research*, *15*(4), 439–440.

Bogacz, R., Usher, M., Zhang, J., & McClelland, J. L. (2007). Extending a biologically inspired model of choice: Multi-alternatives, nonlinearity and value-based multidimensional choice _<http://dx.doi.org/10.1098/rstb.2007.2059>_. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1485), 1655–1670.

Bonds, A. B. (1989). Role of inhibition in the specification of orientation selectivity of cells in the cat striate cortex. *Visual Neuroscience*, *2*(1), 41–55.

Bonin, V., Mante, V., & Carandini, M. (2005). The suppressive field of neurons in lateral geniculate nucleus _<http://dx.doi.org/10.1523/JNEUROSCI.3562%9605.2005>_. *Journal of Neuroscience*, *25*(47), 10844–10856.

Bonin, V., Mante, V., & Carandini, M. (2006). The statistical computation underlying contrast gain control _<http://dx.doi.org/10.1523/JNEUROSCI.0284%9606.2006>_. *Journal of Neuroscience*, *26*(23), 6346–6353.

Boynton, G. M. (2009). A framework for describing the effects of attention on visual responses _<http://dx.doi.org/10.1016/j.visres.2008.11.001>_. *Vision Research*, *49*(10), 1129–1143.

Boynton, R. M., & Whitten, D. N. (1970). Visual adaptation in monkey cones: Recordings of late receptor potentials. *Science*, *170*(3965), 1423–1426.

Brecht, M., Roth, A., & Sakmann, B. (2003). Dynamic receptive fields of reconstructed pyramidal cells in layers 3 and 2 of rat somatosensory barrel cortex _<http://dx.doi.org/10.1113/jphysiol.2003.044222>_. *Journal of Physiology*, *553*(Pt 1), 243–265.

Britten, K. H., & Heuer, H. W. (1999). Spatial summation in the receptive fields of MT neurons. *Journal of Neuroscience*, *19*(12), 5074–5084.

Brouwer, G. J., Arnedo, V., Offen, S., Heeger, D. J., & Grant, A. C. (2015). Normalization in human somatosensory cortex _<http://dx.doi.org/10.1152/jn.00939.2014>_. *Journal of Neurophysiology*, *114*(5), 2588–2599.

Brouwer, G. J., & Heeger, D. J. (2009). Decoding and reconstructing color from responses in human visual cortex _<http://dx.doi.org/10.1523/JNEUROSCI.3577%9609.2009>_. *Journal of Neuroscience*, *29*(44), 13992–14003.

Brouwer, G. J., & Heeger, D. J. (2011). Cross-orientation suppression in human visual cortex _<http://dx.doi.org/10.1152/jn.00540.2011>_. *Journal of Neurophysiology*, *106*(5), 2108–2119.

Bubl, E., Kern, E., Ebert, D., Bach, M., & Tebartz van Elst, L. (2010). Seeing gray when feeling blue? Depression can be measured in the eye of the diseased _<http://dx.doi.org/10.1016/j.biopsych.2010.02.009>_. _Biological Psychiatry_, _68_(2), 205–208.

Busse, L., Wade, A. R., & Carandini, M. (2009). Representation of concurrent stimuli by population activity in visual cortex _<http://dx.doi.org/10.1016/j.neuron.2009.11.004>_. _Neuron_, _64_(6), 931–942.

Butler, P. D., Silverstein, S. M., & Dakin, S. C. (2008). Visual perception and its impairment in schizophrenia _<http://dx.doi.org/10.1016/j.biopsych.2008.03.023>_. _Biological Psychiatry_, _64_(1), 40–47.

Butler, P. D., Zemon, V., Schechter, I., Saperstein, A. M., Hoptman, M. J., Lim, K. O., . . . Javitt, D. C. (2005). Early-stage visual processing and cortical amplification deficits in schizophrenia _<http://dx.doi.org/10.1001/archpsyc.62.5.495>_. _Archives of General Psychiatry_, _62_(5), 495–504.

Campbell, F. W., Cooper, G. F., & Enroth-Cugell, C. (1969). The spatial selectivity of the visual cells of the cat. _Journal of Physiology_, _203_(1), 223–235.

Carandini, M. (2004). Amplification of trial-to-trial response variability by neurons in visual cortex _<http://dx.doi.org/10.1371/journal.pbio.0020264>_. _PLOS Biology_, _2_(9), E264.

Carandini, M. (2012). From circuits to behavior: A bridge too far? _<http://dx.doi.org/10.1038/nn.3043>_. _Nature Neuroscience_, _15_(4), 507–509.

Carandini, M., & Heeger, D. J. (1994). Summation and division by neurons in primate visual cortex. _Science_, _264_(5163), 1333–1336.

Carandini, M., & Heeger, D. J. (2012). Normalization as a canonical Neural Computation _<http://dx.doi.org/10.1038/nrn3136>_. _Nature Reviews Neuroscience_, _13_(1), 51–62.

Carandini, M., Heeger, D. J., & Movshon, J. A. (1997). Linearity and normalization in simple cells of the macaque primary visual cortex. _Journal of Neuroscience_, _17_(21), 8621–8644.

Carandini, M., Heeger, D. J., & Senn, W. (2002). A synaptic explanation of suppression in visual cortex. _Journal of Neuroscience_, _22_(22), 10053–10065.

Cavanaugh, J. R., Bair, W., & Movshon, J. A. (2002a). Nature and interaction of signals from the receptive field center and surround in macaque V1 neurons _<http://dx.doi.org/10.1152/jn.00692.2001>_. _Journal of Neurophysiology_, _88_(5), 2530–2546.

Cavanaugh, J. R., Bair, W., & Movshon, J. A. (2002b). Selectivity and spatial distribution of signals from the receptive field surround in macaque V1 neurons _<http://dx.doi.org/10.1152/jn.00693.2001>_. _Journal of Neurophysiology_, _88_(5), 2547–2556.

deCharms, R. C., Blake, D. T., & Merzenich, M. M. (1998). Optimizing sound features for cortical neurons. _Science_, _280_(5368), 1439–1443.

Chichilnisky, E. J. (2001). A simple white noise analysis of neuronal light responses _<http://dx.doi.org/10.1088/0954-898x/12/2/306>_. _Network: Computation in Neural Systems_, _12_(2), 199–213.

Chubb, C., Sperling, G., & Solomon, J. A. (1989). Texture interactions determine perceived contrast. *Proceedings of the National Academy of Sciences of the United States of America*, *86*(23), 9631–9635.

Churchland, A. K., Kiani, R., & Shadlen, M. N. (2008). Decision-making with multiple alternatives *<http://dx.doi.org/10.1038/nn.2123>*. *Nature Neuroscience*, *11*(6), 693–702.

Cisek, P. (2006). Integrated neural processes for defining potential actions and deciding between them: a computational model *<http://dx.doi.org/10.1523/JNEUROSCI.5605%9605.2006>*. *Journal of Neuroscience*, *26*(38), 9761–9770.

Cleland, T. A., Johnson, B. A., Leon, M., & Linster, C. (2007). Relational representation in the olfactory system *<http://dx.doi.org/10.1073/pnas.0608564104>*. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(6), 1953–1958.

Coen-Cagli, R., Dayan, P., & Schwartz, O. (2012). Cortical surround interactions and perceptual salience via natural scene statistics *<http://dx.doi.org/10.1371/journal.pcbi.1002405>*. *PLOS Computational Biology*, *8*(3), e1002405.

Coen-Cagli, R., Kohn, A., & Schwartz, O. (2015). Flexible gating of contextual influences in natural vision *<http://dx.doi.org/10.1038/nn.4128>*. *Nature Neuroscience*, *18*(11), 1648–1655.

Cox, K. M., & Kable, J. W. (2014). BOLD subjective value signals exhibit robust range adaptation *<http://dx.doi.org/10.1523/JNEUROSCI.3927%9614.2014>*. *Journal of Neuroscience*, *34*(49), 16533–16543.

Crespi, L. P. (1942). Quantitative variation of incentive and performance in the white rat. *American Journal of Psychology*, *55*, 467–517.

Dakin, S., & Frith, U. (2005). Vagaries of visual perception in autism *<http://dx.doi.org/10.1016/j.neuron.2005.10.018>*. *Neuron*, *48*(3), 497–507.

DeAngelis, G. C., Freeman, R. D., & Ohzawa, I. (1994). Length and width tuning of neurons in the cat's primary visual cortex. *Journal of Neurophysiology*, *71*(1), 347–374.

DeAngelis, G. C., Ohzawa, I., & Freeman, R. D. (1995). Receptive-field dynamics in the central visual pathways. *Trends in Neuroscience*, *18*(10), 451–458.

DeAngelis, G. C., Robson, J. G., Ohzawa, I., & Freeman, R. D. (1992). Organization of suppression in receptive fields of neurons in cat visual cortex *<http://dx.doi.org/10.1152/jn.1992.68.1.144>*. *Journal of Neurophysiology*, *68*(1), 144–163.

Deneve, S., Latham, P. E., & Pouget, A. (2001). Efficient computation and cue integration with noisy population codes *<http://dx.doi.org/10.1038/90541>*. *Nature Neuroscience*, *4*(8), 826–831.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention *<http://dx.doi.org/10.1146/annurev.ne.18.030195.001205>*. *Annual Review of Neuroscience*, *18*, 193–222.

DiCarlo, J. J., & Johnson, K. O. (2002). Receptive field structure in cortical area 3b of the alert monkey. *Behavioural Brain Research*, *135*(1–2), 167–178.

DiCarlo, J. J., Zoccolan, D., & Rust, N. C. (2012). How does the brain solve visual object recognition? *<http://dx.doi.org/10.1016/j.neuron.2012.01.010>*. *Neuron*, *73*(3), 415–434.

Dorris, M. C., & Glimcher, P. W. (2004). Activity in posterior parietal cortex is correlated with the relative subjective desirability of action <http://dx.doi.org/10.1016/j.neuron.2004.09.009>. *Neuron*, *44*(2), 365–378.

Douglas, R. J., & Martin, K. A. (1991). A functional microcircuit for cat visual cortex. *Journal of Physiology*, *440*, 735–769.

Douglas, R. J., & Martin, K. A. (2004). Neuronal circuits of the neocortex <http://dx.doi.org/10.1146/annurev.neuro.27.070203.144152>. *Annual Review of Neuroscience*, *27*, 419–451.

Flaherty, C. F. (1982). Incentive contrast: A review of behavioral changes following shifts in reward. *Animal Learning & Behavior*, *10*(4), 409–440.

Flevaris, A. V., & Murray, S. O. (2014). Orientation-specific surround suppression in the primary visual cortex varies as a function of autistic tendency <http://dx.doi.org/10.3389/fnhum.2014.01017>. *Frontiers in Human Neuroscience*, *8*, 1017.

Foss-Feig, J. H., Tadin, D., Schauder, K. B., & Cascio, C. J. (2013). A substantial and unexpected enhancement of motion perception in autism <http://dx.doi.org/10.1523/JNEUROSCI.1608%9612.2013>. *Journal of Neuroscience*, *33*(19), 8243–8249.

Freeman, T. C., Durand, S., Kiper, D. C., & Carandini, M. (2002). Suppression without inhibition in visual cortex. *Neuron*, *35*(4), 759–771.

Furl, N. (2016). Facial-attractiveness choices are predicted by divisive normalization <http://dx.doi.org/10.1177/0956797616661523>. *Psychological Science*, *27*(10), 1379–1387.

Gao, D., & Vasconcelos, N. (2009). Decision-theoretic saliency: computational principles, biological plausibility, and implications for neurophysiology and psychophysics <http://dx.doi.org/10.1162/neco.2009.11-06-391>. *Neural Computation*, *21*(1), 239–271.

Geisler, W. S. (2008). Visual perception and the statistical properties of natural scenes <http://dx.doi.org/10.1146/annurev.psych.58.110405.085632>. *Annual Review of Psychology*, *59*, 167–192.

Ghose, G. M., & Maunsell, J. H. (2008). Spatial summation can explain the attentional modulation of neuronal responses to multiple stimuli in area V4 <http://dx.doi.org/10.1523/JNEUROSCI.0138%9608.2008>. *Journal of Neuroscience*, *28*(19), 5115–5126.

Giraldo, L. G. S., & Schwartz, O. (2018). Integrating flexible normalization into mid-level representations of deep convolutional neural networks. *arXiv preprint arXiv:1806.01823*.

Gnadt, J. W., & Andersen, R. A. (1988). Memory related motor planning activity in posterior parietal cortex of macaque. *Experimental Brain Research*, *70*(1), 216–220.

Golomb, J. D., McDavitt, J. R., Ruf, B. M., Chen, J. I., Saricicek, A., Maloney, K. H., . . . Bhagwagar, Z. (2009). Enhanced visual motion perception in major depressive disorder <http://dx.doi.org/10.1523/JNEUROSCI.1003%9609.2009>. *Journal of Neuroscience*, *29*(28), 9072–9077.

Grabska-Barwinska, A., Barthelme, S., Beck, J., Mainen, Z. F., Pouget, A., & Latham, P. E. (2017). A probabilistic approach to demixing odors <http://dx.doi.org/10.1038/nn.4444>. *Nature Neuroscience*, *20*(1), 98–106.

Griffiths, T. L., & Tenenbaum, J. B. (2009). Theory-based causal induction *<http://dx.doi.org/10.1037/a0017201>*. *Psychological Review*, *116*(4), 661–716.

Grossberg, S. (1988). Nonlinear neural networks: Principles, mechanisms, and architectures. *Neural Networks*, *1*(1), 17–61.

Hanes, D. P., & Schall, J. D. (1996). Neural control of voluntary movement initiation. *Science*, *274*(5286), 427–430.

Hartline, H. K. (1940). The receptive fields of optic nerve fibers. *American Journal of Physiology*, *130*(4), 0690–0699.

Hassabis, D., Kumaran, D., Summerfield, C., & Botvinick, M. (2017). Neuroscience-inspired artificial intelligence *<http://dx.doi.org/10.1016/j.neuron.2017.06.011>*. *Neuron*, *95*(2), 245–258.

Heeger, D. J. (1992). Normalization of cell responses in cat striate cortex *<http://dx.doi.org/10.1017/S0952523800009640>*. *Visual Neuroscience*, *9*(2), 181–197.

Heeger, D. J. (1993). Modeling simple-cell direction selectivity with normalized, half-squared, linear operators *<http://dx.doi.org/10.1152/jn.1993.70.5.1885>*. *Journal of Neurophysiology*, *70*(5), 1885–1898.

Heeger, D. J., Behrmann, M., & Dinstein, I. (2017). Vision as a beachhead *<http://dx.doi.org/10.1016/j.biopsych.2016.09.019>*. *Biological Psychiatry*, *81*(10), 832–837.

Heeger, D. J., Simoncelli, E. P., & Movshon, J. A. (1996). Computational models of cortical visual processing. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(2), 623–627.

Hochberg, L. R., Serruya, M. D., Friehs, G. M., Mukand, J. A., Saleh, M., Caplan, A. H., . . . Donoghue, J. P. (2006). Neuronal ensemble control of prosthetic devices by a human with tetraplegia *<http://dx.doi.org/10.1038/nature04970>*. *Nature*, *442*(7099), 164–171.

Holper, L., Van Brussel, L. D., Schmidt, L., Schulthess, S., Burke, C. J., Louie, K., . . . Tobler, P. N. (2017). Adaptive value normalization in the prefrontal cortex is reduced by memory load *<http://dx.doi.org/10.1523/ENEURO.0365%9617.2017>*. *eNeuro*, *4*(2).

Hopfield, J. J. (1995). Pattern recognition computation using action potential timing for stimulus representation *<http://dx.doi.org/10.1038/376033a0>*. *Nature*, *376*(6535), 33–36.

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, *160*, 106–154.

Huber, J., Payne, J. W., & Puto, C. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of Consumer Research*, *9*(1), 90–98.

Hunt, L. T., Dolan, R. J., & Behrens, T. E. (2014). Hierarchical competitions subserving multi-attribute choice *<http://dx.doi.org/10.1038/nn.3836>*. *Nature Neuroscience*, *17*(11), 1613–1622.

Huys, Q. J., Maia, T. V., & Frank, M. J. (2016). Computational psychiatry as a bridge from neuroscience to clinical applications *<http://dx.doi.org/10.1038/nn.4238>*. *Nature Neuroscience*, *19*(3), 404–413.

Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*.

Itthipuripat, S., Cha, K., Rangsipat, N., & Serences, J. T. (2015). Value-based attentional capture influences context-dependent decision-making *<http://dx.doi.org/10.1152/jn.00343.2015>*. *Journal of Neurophysiology*, *114*(1), 560–569.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*(10–12), 1489–1506.

Jarrett, K., Kavukcuoglu, K., & LeCun, Y. (2009). *What is the best multi-stage architecture for object recognition?* Paper presented at the Computer Vision, 2009 IEEE 12th International Conference on.

Kahneman, D., & Tversky, A. (1979). Prospect theory—analysis of decision under risk. *Econometrica*, *47*(2), 263–291.

Kaliukhovich, D. A., & Vogels, R. (2016). Divisive normalization predicts adaptation-induced response changes in macaque inferior temporal cortex *<http://dx.doi.org/10.1523/JNEUROSCI.2011%9615.2016>*. *Journal of Neuroscience*, *36*(22), 6116–6128.

Katzner, S., Busse, L., & Carandini, M. (2011). GABAA inhibition controls response gain in visual cortex *<http://dx.doi.org/10.1523/JNEUROSCI.5753-10.2011>*. *Journal of Neuroscience*, *31*(16), 5931–5941.

Kay, K. N., Naselaris, T., Prenger, R. J., & Gallant, J. L. (2008). Identifying natural images from human brain activity *<http://dx.doi.org/10.1038/nature06713>*. *Nature*, *452*(7185), 352–355.

Khaw, M. W., Glimcher, P. W., & Louie, K. (2017). Normalized value coding explains dynamic adaptation in the human valuation process *<http://dx.doi.org/10.1073/pnas.1715293114>*. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(48), 12696–12701.

Klein, J. T., Deaner, R. O., & Platt, M. L. (2008). Neural correlates of social target value in macaque parietal cortex *<http://dx.doi.org/10.1016/j.cub.2008.02.047>*. *Current Biology*, *18*(6), 419–424.

Kobayashi, S., Pinto de Carvalho, O., & Schultz, W. (2010). Adaptation of reward sensitivity in orbitofrontal neurons *<http://dx.doi.org/10.1523/JNEUROSCI.4009-09.2010>*. *Journal of Neuroscience*, *30*(2), 534–544.

Kohn, A., Coen-Cagli, R., Kanitscheider, I., & Pouget, A. (2016). Correlations and neuronal population information *<http://dx.doi.org/10.1146/annurev-neuro-070815-013851>*. *Annual Review of Neuroscience*, *39*, 237–256.

Koszegi, B., & Rabin, M. (2006). A model of reference-dependent preferences. *Quarterly Journal of Economics*, *121*(4), 1133–1165.

Koszegi, B., & Rabin, M. (2007). Reference-dependent risk attitudes. *American Economic Review*, *97*(4), 1047–1073.

Kouh, M., & Poggio, T. (2008). A canonical neural circuit for cortical nonlinear operations *<http://dx.doi.org/10.1162/neco.2008.02-07-466>*. *Neural Computation*, *20*(6), 1427–1451.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). *Imagenet classification with deep convolutional neural networks*. Paper presented at the advances in neural information processing systems.

Kuffler, S. W. (1953). Discharge patterns and functional organization of mammalian retina _<http://dx.doi.org/10.1152/jn.1953.16.1.37>_. *Journal of Neurophysiology*, *16*(1), 37–68.

Laparra, V., Berardino, A., Balle, J., & Simoncelli, E. P. (2017). Perceptually optimized image rendering _<http://dx.doi.org/10.1364/JOSAA.34.001511>_. *Journal of the Optical Society of America A Optics Image Science Vision*, *34*(9), 1511–1525.

Laparra, V., Munoz-Mari, J., & Malo, J. (2010). Divisive normalization image quality metric revisited _<http://dx.doi.org/10.1364/JOSAA.27.000852>_. *Journal of the Optical Society of America A Optics Image Science Vision*, *27*(4), 852–864.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning _<http://dx.doi.org/10.1038/nature14539>_. *Nature*, *521*(7553), 436–444.

Lee, J., & Maunsell, J. H. (2009). A normalization model of attentional modulation of single unit responses _<http://dx.doi.org/10.1371/journal.pone.0004651>_. *PLOS One*, *4*(2), e4651.

Lee, T. S., & Mumford, D. (2003). Hierarchical Bayesian inference in the visual cortex. *Journal of the Optical Society of America A Optics Image Science Vision*, *20*(7), 1434–1448.

Li, H. H., Carrasco, M., & Heeger, D. J. (2015). Deconstructing interocular suppression: Attention and divisive normalization _<http://dx.doi.org/10.1371/journal.pcbi.1004510>_. *PLOS Computational Biology*, *11*(10), e1004510.

Li, Z. (2002). A saliency map in primary visual cortex. *Trends in Cognitive Sciences*, *6*(1), 9–16.

Ling, S., & Blake, R. (2012). Normalization regulates competition for visual awareness _<http://dx.doi.org/10.1016/j.neuron.2012.05.032>_. *Neuron*, *75*(3), 531–540.

Livingstone, M. S., Pettine, W. W., Srihasam, K., Moore, B., Morocz, I. A., & Lee, D. (2014). Symbol addition by monkeys provides evidence for normalized quantity coding _<http://dx.doi.org/10.1073/pnas.1404208111>_. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(18), 6822–6827.

LoFaro, T., Louie, K., Webb, R., & Glimcher, P. (2014). The temporal dynamics of cortical normalization models of decision-making. *Letters in Biomathematics*, *1*(2), 209–220.

Louie, K., & Glimcher, P. W. (2010). Separating value from choice: Delay discounting activity in the lateral intraparietal area _<http://dx.doi.org/10.1523/JNEUROSCI.5742-09.2010>_. *Journal of Neuroscience*, *30*(16), 5498–5507.

Louie, K., & Glimcher, P. W. (2012). Efficient coding and the neural representation of value _<http://dx.doi.org/10.1111/j.1749-6632.2012.06496.x>_. *Annals of the New York Academy of Sciences*, *1251*, 13–32.

Louie, K., Glimcher, P. W., & Webb, R. (2015). Adaptive neural coding: From biological to behavioral decision-making _<http://dx.doi.org/10.1016/j.cobeha.2015.08.008>_. *Current Opinion in Behavioral Sciences*, *5*, 91–99.

Louie, K., Grattan, L. E., & Glimcher, P. W. (2011). Reward value-based gain control: Divisive normalization in parietal cortex _<http://dx.doi.org/10.1523/JNEUROSCI.1237-11.2011>_. *Journal of Neuroscience*, *31*(29), 10627–10639.

Louie, K., Khaw, M. W., & Glimcher, P. W. (2013). Normalization is a general neural mechanism for context-dependent decision making _<http://dx.doi.org/10.1073/pnas.1217854110>_. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(15), 6139–6144.

Louie, K., LoFaro, T., Webb, R., & Glimcher, P. W. (2014). Dynamic divisive normalization predicts time-varying value coding in decision-related circuits *<http://dx.doi.org/10.1523/JNEUROSCI.2851-14.2014>*. *Journal of Neuroscience*, *34*(48), 16046–16057.

Luce, R. D. (1959). *Individual choice behavior: A theoretical analysis*. New York, NY: Wiley.

Luo, S. X., Axel, R., & Abbott, L. F. (2010). Generating sparse and selective third-order responses in the olfactory system of the fly *<http://dx.doi.org/10.1073/pnas.1005635107>*. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(23), 10713–10718.

Lyu, S. (2011). Dependency reduction with divisive normalization: Justification and effectiveness *<http://dx.doi.org/10.1162/NECO_a_00197>*. *Neural Computation*, *23*(11), 2942–2973.

Lyu, S., & Simoncelli, E. P. (2008, 23–28 June). *Nonlinear image representation using divisive normalization*. Paper presented at the 2008 IEEE Conference on Computer Vision and Pattern Recognition.

Ma, W. J., Beck, J. M., Latham, P. E., & Pouget, A. (2006). Bayesian inference with probabilistic population codes *<http://dx.doi.org/10.1038/nn1790>*. *Nature Neuroscience*, *9*(11), 1432–1438.

Malo, J., Epifanio, I., Navarro, R., & Simoncelli, E. P. (2006). Nonlinear image representation for efficient perceptual coding. *IEEE Transactions on Image Processing*, *15*(1), 68–80.

Marr, D. (1982). *Vision*. San Francisco, CA: W.H. Freeman.

Martinez-Trujillo, J., & Treue, S. (2002). Attentional modulation strength in cortical area MT depends on stimulus contrast. *Neuron*, *35*(2), 365–370.

Martinez-Trujillo, J. C., & Treue, S. (2004). Feature-based attention increases the selectivity of population responses in primate visual corte *<http://dx.doi.org/10.1016/j.cub.2004.04.028>*. *Current Biology*, *14*(9), 744–751.

McAdams, C. J., & Maunsell, J. H. (1999). Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *Journal of Neuroscience*, *19*(1), 431–441.

Michaelis, L., Menten, M. L., Johnson, K. A., & Goody, R. S. (2011). The original Michaelis constant: translation of the 1913 Michaelis-Menten paper *<http://dx.doi.org/10.1021/bi201284u>*. *Biochemistry*, *50*(39), 8264–8269.

Mikaelian, S., & Simoncelli, E. P. (2001). Modeling temporal response characteristics of V1 neurons with a dynamic normalization model. *Neurocomputing*, *38–40*, 1461–1467.

Montague, P. R., Dolan, R. J., Friston, K. J., & Dayan, P. (2012). Computational psychiatry *<http://dx.doi.org/10.1016/j.tics.2011.11.018>*. *Trends in Cognitive Sciences*, *16*(1), 72–80.

Moradi, F., & Heeger, D. J. (2009). Inter-ocular contrast normalization in human visual cortex *<http://dx.doi.org/10.1167/9.3.13>*. *Journal of Vision*, *9*(3), 13.

Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science*, *229*(4715), 782–784.

Morrone, M. C., Burr, D. C., & Maffei, L. (1982). Functional implications of cross-orientation inhibition of cortical visual cells. I. Neurophysiological evidence. *Proceedings of the Royal Scoiety of London. Series B: Biological Sciences*, *216*(1204), 335–354.

Movshon, J. A., Adelson, E. H., Gizzi, M. S., & Newsome, W. T. (1985). The analysis of moving visual patterns. In C. Chagas, R. Gattass, & C. G. Gross (Eds.), *Pattern recognition mechanisms* (pp. 117–151). New York, NY: Springer.

Movshon, J. A., Thompson, I. D., & Tolhurst, D. J. (1978). Spatial summation in the receptive fields of simple cells in the cat's striate cortex. *Journal of Physiology*, *283*, 53–77.

Mullins, C., Fishell, G., & Tsien, R. W. (2016). Unifying views of autism spectrum disorders: A consideration of autoregulatory feedback loops *<http://dx.doi.org/10.1016/j.neuron.2016.02.017>*. *Neuron*, *89*(6), 1131–1156.

Naka, K. I., & Rushton, W. A. (1966). S-potentials from colour units in the retina of fish (*Cyprinidae*). *Journal of Physiology*, *185*(3), 536–555.

Nassi, J. J., Avery, M. C., Cetin, A. H., Roe, A. W., & Reynolds, J. H. (2015). Optogenetic activation of normalization in alert macaque visual cortex *<http://dx.doi.org/10.1016/j.neuron.2015.05.040>*. *Neuron*, *86*(6), 1504–1517.

Nassi, J. J., Gomez-Laberge, C., Kreiman, G., & Born, R. T. (2014). Corticocortical feedback increases the spatial extent of normalization *<http://dx.doi.org/10.3389/fnsys.2014.00105>*. *Frontiers in System Neuroscience*, *8*, 105.

Nassi, J. J., Lomber, S. G., & Born, R. T. (2013). Corticocortical feedback contributes to surround suppression in V1 of the alert primate *<http://dx.doi.org/10.1523/JNEUROSCI.5124-12.2013>*. *Journal of Neuroscience*, *33*(19), 8504–8517.

Ni, A. M., & Maunsell, J. H. R. (2017). Spatially tuned normalization explains attention modulation variance within neurons *<http://dx.doi.org/10.1152/jn.00218.2017>*. *Journal of Neurophysiology*, *118*(3), 1903–1913.

Ni, A. M., Ray, S., & Maunsell, J. H. (2012). Tuned normalization explains the size of attention modulations *<http://dx.doi.org/10.1016/j.neuron.2012.01.006>*. *Neuron*, *73*(4), 803–813.

Nieder, A., Freedman, D. J., & Miller, E. K. (2002). Representation of the quantity of visual items in the primate prefrontal cortex *<http://dx.doi.org/10.1126/science.1072493>*. *Science*, *297*(5587), 1708–1711.

Nieder, A., & Miller, E. K. (2003). Coding of cognitive magnitude: compressed scaling of numerical information in the primate prefrontal cortex. *Neuron*, *37*(1), 149–157.

Nieder, A., & Miller, E. K. (2004). A parieto-frontal network for visual numerical information in the monkey *<http://dx.doi.org/10.1073/pnas.0402239101>*. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(19), 7457–7462.

Normann, R. A., & Perlman, I. (1979). The effects of background illumination on the photoresponses of red and green cones. *Journal of Physiology*, *286*, 491–507.

Norton, D. J., McBain, R. K., Pizzagalli, D. A., Cronin-Golomb, A., & Chen, Y. (2016). Dysregulation of visual motion inhibition in major depression *<http://dx.doi.org/10.1016/j.psychres.2016.04.028>*. *Psychiatry Research*, *240*, 214–221.

Ohshiro, T., Angelaki, D. E., & DeAngelis, G. C. (2011). A normalization model of multisensory integration *<http://dx.doi.org/10.1038/nn.2815>*. *Nature Neuroscience*, *14*(6), 775–782.

Ohshiro, T., Angelaki, D. E., & DeAngelis, G. C. (2017). A neural signature of divisive normalization at the level of multisensory integration in primate cortex *<http://dx.doi.org/10.1016/j.neuron.2017.06.043>*. *Neuron*, *95*(2), 399–411, e398.

Olsen, S. R., Bhandawat, V., & Wilson, R. I. (2010). Divisive normalization in olfactory population codes *<http://dx.doi.org/10.1016/j.neuron.2010.04.009>*. *Neuron*, *66*(2), 287–299.

Olsen, S. R., & Wilson, R. I. (2008). Lateral presynaptic inhibition mediates gain control in an olfactory circuit *<http://dx.doi.org/10.1038/nature06864>*. *Nature*, *452*(7190), 956–960.

Ozeki, H., Finn, I. M., Schaffer, E. S., Miller, K. D., & Ferster, D. (2009). Inhibitory stabilization of the cortical network underlies visual surround suppression *<http://dx.doi.org/10.1016/j.neuron.2009.03.028>*. *Neuron*, *62*(4), 578–592.

Ozeki, H., Sadakane, O., Akasaki, T., Naito, T., Shimegi, S., & Sato, H. (2004). Relationship between excitation and inhibition underlying size tuning and contextual response modulation in the cat primary visual cortex *<http://dx.doi.org/10.1523/JNEUROSCI.3852-03.2004%2024/6/1428>*. *Journal of Neuroscience*, *24*(6), 1428–1438.

Padoa-Schioppa, C. (2009). Range-adapting representation of economic value in the orbitofrontal cortex *<http://dx.doi.org/10.1523/JNEUROSCI.3751-09.2009>*. *Journal of Neuroscience*, *29*(44), 14004–14014.

Paninski, L., Pillow, J., & Lewi, J. (2007). Statistical models for neural encoding, decoding, and optimal stimulus design *<http://dx.doi.org/10.1016/S0079-6123(06)65031-0>*. *Progress in Brain Research*, *165*, 493–507.

Pastor-Bernier, A., & Cisek, P. (2011). Neural correlates of biased competition in premotor cortex *<http://dx.doi.org/10.1523/JNEUROSCI.5681-10.2011>*. *Journal of Neuroscience*, *31*(19), 7083–7088.

Petrov, Y., Carandini, M., & McKee, S. (2005). Two distinct mechanisms of suppression in human vision *<http://dx.doi.org/10.1523/JNEUROSCI.2871-05.2005>*. *Journal of Neuroscience*, *25*(38), 8704–8707.

Platt, M. L., & Glimcher, P. W. (1999). Neural correlates of decision variables in parietal cortex *<http://dx.doi.org/10.1038/22268>*. *Nature*, *400*(6741), 233–238.

Porciatti, V., Bonanni, P., Fiorentini, A., & Guerrini, R. (2000). Lack of cortical contrast gain control in human photosensitive epilepsy *<http://dx.doi.org/10.1038/72972>*. *Nature Neuroscience*, *3*(3), 259–263.

Pouget, A., Beck, J. M., Ma, W. J., & Latham, P. E. (2013). Probabilistic brains: knowns and unknowns *<http://dx.doi.org/10.1038/nn.3495>*. *Nature Neuroscience*, *16*(9), 1170–1178.

Pouget, A., & Sejnowski, T. J. (1997). Spatial transformations in the parietal cortex using basis functions *<http://dx.doi.org/10.1162/jocn.1997.9.2.222>*. *Journal of Cognitive Neuroscience*, *9*(2), 222–237.

Pouget, A., & Snyder, L. H. (2000). Computational approaches to sensorimotor transformations *<http://dx.doi.org/10.1038/81469>*. *Nature Neuroscience*, *3*(11s), 1192–1198.

Rabinowitz, N. C., Willmore, B. D., Schnupp, J. W., & King, A. J. (2011). Contrast gain control in auditory cortex *<http://dx.doi.org/10.1016/j.neuron.2011.04.030>*. *Neuron*, *70*(6), 1178–1191.

Rangel, A., & Clithero, J. A. (2012). Value normalization in decision making: Theory and evidence *<http://dx.doi.org/10.1016/j.conb.2012.07.011>*. *Current Opinion in Neurobiology*, *22*(6), 970–981.

Rao, R. P. N., Olshausen, B. A., & Lewicki, M. S. (2002). *Probabilistic models of the brain: Perception and neural function/ edited by* Rajesh P.N. Rao, Bruno A. Olshausen, Michael S. Lewicki. Cambridge, MA: MIT Press.

Ray, S., Ni, A. M., & Maunsell, J. H. (2013). Strength of gamma rhythm depends on normalization *<http://dx.doi.org/10.1371/journal.pbio.1001477>*. *PLOS Biology*, *11*(2), e1001477.

Recanzone, G. H., Wurtz, R. H., & Schwarz, U. (1997). Responses of MT and MST neurons to one and two moving objects in the receptive field *<http://dx.doi.org/10.1152/jn.1997.78.6.2904>*. *Journal of Neurophysiology*, *78*(6), 2904–2915.

Ren, M., Liao, R., Urtasun, R., Sinz, F. H., & Zemel, R. S. (2016). Normalizing the normalizers: Comparing and extending network normalization schemes. *arXiv preprint arXiv:1611.04520*.

Reynaud, A., Masson, G. S., & Chavane, F. (2012). Dynamics of local input normalization result from balanced short- and long-range intracortical interactions in area V1 *<http://dx.doi.org/10.1523/JNEUROSCI.1618-12.2012>*. *Journal of Neuroscience*, *32*(36), 12558–12569.

Reynolds, J. H., Chelazzi, L., & Desimone, R. (1999). Competitive mechanisms subserve attention in macaque areas V2 and V4. *Journal of Neuroscience*, *19*(5), 1736–1753.

Reynolds, J. H., & Desimone, R. (2003). Interacting roles of attention and visual salience in V4. *Neuron*, *37*(5), 853–863.

Reynolds, J. H., & Heeger, D. J. (2009). The normalization model of attention *<http://dx.doi.org/10.1016/j.neuron.2009.01.002>*. *Neuron*, *61*(2), 168–185.

Reynolds, J. H., Pasternak, T., & Desimone, R. (2000). Attention increases sensitivity of V4 neurons. *Neuron*, *26*(3), 703–714.

Rieke, F., & Rudd, M. E. (2009). The challenges natural images pose for visual adaptation *<http://dx.doi.org/10.1016/j.neuron.2009.11.028>*. *Neuron*, *64*(5), 605–616.

Rieke, F., Warland, D., de Ruyter van Steveninck, R., & Bialek, W. (1997). *Spikes: Exploring the neural code*. Cambridge, MA: MIT Press.

Rigoli, F., Friston, K. J., Martinelli, C., Selakovic, M., Shergill, S. S., & Dolan, R. J. (2016). A Bayesian model of context-sensitive value attribution *<http://dx.doi.org/10.7554/eLife.16127>*. *Elife*, *5*.

Rigoli, F., Mathys, C., Friston, K. J., & Dolan, R. J. (2017). A unifying Bayesian account of contextual effects in value-based choice *<http://dx.doi.org/10.1371/journal.pcbi.1005769>*. *PLOS Computational Biology*, *13*(10), e1005769.

Rigoli, F., Rutledge, R. B., Dayan, P., & Dolan, R. J. (2016). The influence of contextual reward statistics on risk preference *<http://dx.doi.org/10.1016/j.neuroimage.2015.12.016>*. *Neuroimage*, *128*, 74–84.

Ringach, D. L., & Malone, B. J. (2007). The operating point of the cortex: Neurons as large deviation detectors *<http://dx.doi.org/10.1523/JNEUROSCI.1048-07.2007>*. *Journal of Neuroscience*, *27*(29), 7673–7683.

Robertson, C. E., Kravitz, D. J., Freyberg, J., Baron-Cohen, S., & Baker, C. I. (2013). Tunnel vision: Sharper gradient of spatial attention in autism *<http://dx.doi.org/10.1523/JNEUROSCI.51200-12.2013>*. *Journal of Neuroscience*, *33*(16), 6776–6781.

Rodman, H. R., & Albright, T. D. (1989). Single-unit analysis of pattern-motion selective properties in the middle temporal visual area (MT). *Experimental Brain Research*, *75*(1), 53–64.

Roe, R. M., Busemeyer, J. R., & Townsend, J. T. (2001). Multialternative decision field theory: A dynamic connectionist model of decision making. *Psychological Review*, *108*(2), 370–392.

Roitman, J. D., Brannon, E. M., & Platt, M. L. (2007). Monotonic coding of numerosity in macaque lateral intraparietal area <http://dx.doi.org/10.1371/journal.pbio.0050208>. *PLOS Biology*, *5*(8), e208.

Roitman, J. D., & Shadlen, M. N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *Journal of Neuroscience*, *22*(21), 9475–9489.

Rorie, A. E., Gao, J., McClelland, J. L., & Newsome, W. T. (2010). Integration of sensory and reward information during perceptual decision-making in lateral intraparietal cortex (LIP) of the macaque monkey <http://dx.doi.org/10.1371/journal.pone.0009308>. *PLOS One*, *5*(2), e9308.

Rosenberg, A., Patterson, J. S., & Angelaki, D. E. (2015). A computational perspective on autism <http://dx.doi.org/10.1073/pnas.1510583112>. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(30), 9158–9165.

Rubenstein, J. L., & Merzenich, M. M. (2003). Model of autism: Increased ratio of excitation/inhibition in key neural systems. *Genes, Brain and Behavior*, *2*(5), 255–267.

Rubin, D. B., Van Hooser, S. D., & Miller, K. D. (2015). The stabilized supralinear network: A unifying circuit motif underlying multi-input integration in sensory cortex <http://dx.doi.org/10.1016/j.neuron.2014.12.026>. *Neuron*, *85*(2), 402–417.

Ruff, D. A., Alberts, J. J., & Cohen, M. R. (2016). Relating normalization to neuronal populations across cortical areas <http://dx.doi.org/10.1152/jn.00017.2016>. *Journal of Neurophysiology*, *116*(3), 1375–1386.

Ruff, D. A., & Cohen, M. R. (2017). A normalization model suggests that attention changes the weighting of inputs between visual areas <http://dx.doi.org/10.1073/pnas.1619857114>. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(20), E4085–E4094.

Rust, N. C., Mante, V., Simoncelli, E. P., & Movshon, J. A. (2006). How MT cells analyze the motion of visual patterns <http://dx.doi.org/10.1038/nn1786>. *Nature Neuroscience*, *9*(11), 1421–1431.

Sato, T. K., Haider, B., Hausser, M., & Carandini, M. (2016). An excitatory basis for divisive normalization in visual cortex <http://dx.doi.org/10.1038/nn.4249>. *Nature Neuroscience*, *19*(4), 568–570.

Sato, T. K., Hausser, M., & Carandini, M. (2014). Distal connectivity causes summation and division across mouse visual cortex <http://dx.doi.org/10.1038/nn.3585>. *Nature Neuroscience*, *17*(1), 30–32.

Sawada, T., & Petrov, A. A. (2017). The divisive normalization model of V1 neurons: A comprehensive comparison of physiological data and model predictions <http://dx.doi.org/10.1152/jn.00821.2016>. *Journal of Neurophysiology*, *118*(6), 3051–3091.

Schneeweis, D. M., & Schnapf, J. L. (1999). The photovoltage of macaque cone photoreceptors: Adaptation, noise, and kinetics. *Journal of Neuroscience*, *19*(4), 1203–1216.

Schwartz, A. B., Cui, X. T., Weber, D. J., & Moran, D. W. (2006). Brain-controlled interfaces: movement restoration with neural prosthetics *<http://dx.doi.org/10.1016/j.neuron.2006.09.019>*. *Neuron*, *52*(1), 205–220.

Schwartz, O., & Simoncelli, E. P. (2001). Natural signal statistics and sensory gain control *<http://dx.doi.org/10.1038/9052690526>*. *Nature Neuroscience*, *4*(8), 819–825.

Sclar, G., Maunsell, J. H., & Lennie, P. (1990). Coding of image contrast in central visual pathways of the macaque monkey. *Vision Research*, *30*(1), 1–10.

Sejnowski, T. J., Koch, C., & Churchland, P. S. (1988). Computational neuroscience. *Science*, *241*(4871), 1299–1306.

Shadlen, M. N., & Newsome, W. T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *Journal of Neurophysiology*, *86*(4), 1916–1936.

Shafir, S., Waite, T. A., & Smith, B. H. (2002). Context-dependent violations of rational choice in honeybees (*Apis mellifera*) and gray jays (*Perisoreus canadensis*). *Behavioral Ecology and Sociobiology*, *51*(2), 180–187.

Shapley, R. M., & Enroth-Cugell, C. (1984). Visual adaptation and retinal gain control. *Progress in Retinal and Eye Research*, *3*, 263–346.

Shapley, R. M., & Victor, J. D. (1978). The effect of contrast on the transfer properties of cat retinal ganglion cells. *Journal of Physiology*, *285*, 275–298.

Shapley, R. M., & Victor, J. D. (1981). How the contrast gain control modifies the frequency responses of cat retinal ganglion cells. *Journal of Physiology*, *318*, 161–179.

Shushruth, S., Mangapathy, P., Ichida, J. M., Bressloff, P. C., Schwabe, L., & Angelucci, A. (2012). Strong recurrent networks compute the orientation tuning of surround modulation in the primate primary visual cortex *<http://dx.doi.org/10.1523/JNEUROSCI.3789-11.2012>*. *Journal of Neuroscience*, *32*(1), 308–321.

Silver, R. A. (2010). Neuronal arithmetic *<http://dx.doi.org/10.1038/nrn2864>*. *Nature Reviews Neuroscience*, *11*(7), 474–489.

Simoncelli, E. P., & Heeger, D. J. (1998). A model of neuronal responses in visual area MT. *Vision Research*, *38*(5), 743–761.

Simoncelli, E. P., & Olshausen, B. A. (2001). Natural image statistics and neural representation *<http://dx.doi.org/10.1146/annurev.neuro.24.1.1193>*. *Annual Review of Neuroscience*, *24*, 1193–1216.

Simoncelli, E. P., Paninski, L., Pillow, J., & Schwartz, O. (2004). Characterization of neural responses with stochastic stimuli. *The Cognitive Neurosciences*, *3*(327–338), 1.

Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. *Journal of Consumer Research*, *16*(2), 158–174.

Singer, W., & Gray, C. M. (1995). Visual feature integration and the temporal correlation hypothesis *<http://dx.doi.org/10.1146/annurev.ne.18.030195.003011>*. *Annual Review of Neuroscience*, *18*, 555–586.

Sinz, F., & Bethge, M. (2013). Temporal adaptation enhances efficient contrast gain control on natural images *<http://dx.doi.org/10.1371/journal.pcbi.1002889>*. *PLOS Computational Biology*, *9*(1), e1002889.

Sit, Y. F., Chen, Y., Geisler, W. S., Miikkulainen, R., & Seidemann, E. (2009). Complex dynamics of V1 population responses explained by a simple gain-control model *<http://dx.doi.org/10.1016/j.neuron.2009.08.041>*. *Neuron*, *64*(6), 943–956.

Smith, P. L., Sewell, D. K., & Lilburn, S. D. (2015). From shunting inhibition to dynamic normalization: Attentional selection and decision-making in brief visual displays *<http://dx.doi.org/10.1016/j.visres.2014.11.001>*. *Vision Research*, *116*(Pt B), 219–240.

Snow, M., Coen-Cagli, R., & Schwartz, O. (2016). Specificity and timescales of cortical adaptation as inferences about natural movie statistics *<http://dx.doi.org/10.1167/16.13.1>*. *Journal of Vision*, *16*(13).

Snyder, L. H., Batista, A. P., & Andersen, R. A. (1997). Coding of intention in the posterior parietal cortex *<http://dx.doi.org/10.1038/386167a0>*. *Nature*, *386*(6621), 167–170.

Solomon, J. A., Sperling, G., & Chubb, C. (1993). The lateral inhibition of perceived contrast is indifferent to on-center/off-center segregation, but specific to orientation. *Vision Research*, *33*(18), 2671–2683.

Soltani, A., De Martino, B., & Camerer, C. (2012). A range-normalization model of context-dependent choice: A new model and evidence *<http://dx.doi.org/10.1371/journal.pcbi.1002607>*. *PLOS Computational Biology*, *8*(7), e1002607.

Stein, B. E., Stanford, T. R., & Rowland, B. A. (2014). Development of multisensory integration from the perspective of the individual neuron. *Nature Reviews Neuroscience*, *15*(8), 520–535.

Stephens, D. W., & Krebs, J. R. (1986). *Foraging theory*. Princeton, NJ: Princeton University Press.

Sugrue, L. P., Corrado, G. S., & Newsome, W. T. (2004). Matching behavior and the representation of value in the parietal cortex *<http://dx.doi.org/10.1126/science.1094765>*. *Science*, *304*(5678), 1782–1787.

Tadin, D., Kim, J., Doop, M. L., Gibson, C., Lappin, J. S., Blake, R., . . . Park, S. (2006). Weakened center-surround interactions in visual motion processing in schizophrenia *<http://dx.doi.org/10.1523/JNEUROSCI.2592-06.2006>*. *Journal of Neuroscience*, *26*(44), 11403–11412.

Teo, P. C., & Heeger, D. J. (1994). Perceptual image distortion. *Icip-94—Proceedings*, *2*, 982–986.

Theunissen, F., & Miller, J. P. (1995). Temporal encoding in nervous systems: A rigorous definition. *Journal of Computational Neuroscience*, *2*(2), 149–162.

Tremblay, L., & Schultz, W. (1999). Relative reward preference in primate orbitofrontal cortex *<http://dx.doi.org/10.1038/19525>*. *Nature*, *398*(6729), 704–708.

Treue, S., & Martinez Trujillo, J. C. (1999). Feature-based attention influences motion processing gain in macaque visual cortex *<http://dx.doi.org/10.1038/21176>*. *Nature*, *399*(6736), 575–579.

Trueblood, J. S., Brown, S. D., & Heathcote, A. (2014). The multiattribute linear ballistic accumulator model of context effects in multialternative choice *<http://dx.doi.org/10.1037/a0036137>*. *Psychological Review*, *121*(2), 179–205.

Tsai, J. J., Norcia, A. M., Ales, J. M., & Wade, A. R. (2011). Contrast gain control abnormalities in idiopathic generalized epilepsy _<http://dx.doi.org/10.1002/ana.22462>_. _Annals of Neurology_, _70_(4), 574–582.

Tversky, A. (1972). Elimination by aspects—a theory of choice. _Psychological Review_, _79_(4), 281.

Tversky, A., & Simonson, I. (1993). Context-dependent preferences. _Management Science_, _39_(10), 1179–1189.

Tymula, A., & Plassmann, H. (2016). Context-dependency in valuation _<http://dx.doi.org/10.1016/j.conb.2016.06.015>_. _Current Opinion in Neurobiology_, _40_, 59–65.

Van Essen, D. C., Anderson, C. H., & Felleman, D. J. (1992). Information processing in the primate visual system: An integrated systems perspective. _Science_, _255_(5043), 419–423.

Verhoef, B. E., & Maunsell, J. H. R. (2017). Attention-related changes in correlated neuronal activity arise from normalization mechanisms _<http://dx.doi.org/10.1038/nn.4572>_. _Nature Neuroscience_, _20_(7), 969–977.

Vinje, W. E., & Gallant, J. L. (2002). Natural stimulation of the nonclassical receptive field increases information transmission efficiency in V1 _<http://dx.doi.org/20026216>_. _Journal of Neuroscience_, _22_(7), 2904–2915.

Von Neumann, J., & Morgenstern, O. (1944). _Theory of games and economic behavior_. Princeton, NJ: Princeton University Press.

Wainwright, M. J., Schwartz, O., & Simoncelli, E. P. (2002). Natural image statistics and divisive normalization: Modeling nonlinearity and adaptation in cortical neurons. In R. P. N. Rao, B. A. Olshausen, & M. S. Lewicki (Eds.), _Probabilistic models of the brain: Perception and neural function_, (pp. 203–222). Cambridge, MA: MIT Press.

Wang, X. J. (2008). Decision making in recurrent neuronal circuits _<http://dx.doi.org/10.1016/j.neuron.2008.09.034>_. _Neuron_, _60_(2), 215–234.

Wang, X. J., & Krystal, J. H. (2014). Computational psychiatry _<http://dx.doi.org/10.1016/j.neuron.2014.10.018>_. _Neuron_, _84_(3), 638–654.

Westrick, Z. M., Heeger, D. J., & Landy, M. S. (2016). Pattern adaptation and normalization reweighting _<http://dx.doi.org/10.1523/JNEUROSCI.1067>_. _Journal of Neuroscience_, _36_(38), 9805–9816.

Williford, T., & Maunsell, J. H. (2006). Effects of spatial attention on contrast response functions in macaque area V4 _<http://dx.doi.org/10.1152/jn.01207.2005>_. _Journal of Neurophysiology_, _96_(1), 40–54.

Wilson, H. R., & Humanski, R. (1993). Spatial frequency adaptation and contrast gain control _<http://dx.doi.org/10.1016/0042-6989(93)90248-U>_. _Vision Research_, _33_(8), 1133–1149.

Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An internal model for sensorimotor integration. _Science_, _269_(5232), 1880–1882.

Xiao, J., Niu, Y. Q., Wiesner, S., & Huang, X. (2014). Normalization of neuronal responses in cortical area MT across signal strengths and motion directions _<http://dx.doi.org/10.1152/jn.00700.2013>_. _Journal of Neurophysiology_, _112_(6), 1291–1306.

Xing, J., & Heeger, D. J. (2001). Measurement and modeling of center-surround suppression and enhancement. *Vision Research*, *41*(5), 571–583.

Yamada, H., Louie, K., Tymula, A., & Glimcher, P. W. (2018). Free choice shapes normalized value signals in medial orbitofrontal cortex *<http://dx.doi.org/10.1038/s41467-017-02614-w>*. *Nature Communications*, *9*(1), 162.

Yamins, D. L., & DiCarlo, J. J. (2016). Using goal-driven deep learning models to understand sensory cortex *<http://dx.doi.org/10.1038/nn.4244>*. *Nature Neuroscience*, *19*(3), 356–365.

Yamins, D. L., Hong, H., Cadieu, C. F., Solomon, E. A., Seibert, D., & DiCarlo, J. J. (2014). Performance-optimized hierarchical models predict neural responses in higher visual cortex *<http://dx.doi.org/10.1073/pnas.1403112111>*. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(23), 8619–8624.

Yizhar, O., Fenno, L. E., Prigge, M., Schneider, F., Davidson, T. J., O'Shea, D. J., . . . Deisseroth, K. (2011). Neocortical excitation/inhibition balance in information processing and social dysfunction *<http://dx.doi.org/10.1038/nature10360>*. *Nature*, *477*(7363), 171–178.

Zeaman, D. (1949). Response latency as a function of the amount of reinforcement. *Journal of Experimental Psychology*, *39*(4), 466–483.

Zhu, P., Frank, T., & Friedrich, R. W. (2013). Equalization of odor representations by a network of electrically coupled inhibitory interneurons *<http://dx.doi.org/10.1038/nn.3528>*. *Nature Neuroscience*, *16*(11), 1678–1686.

Zoccolan, D., Cox, D. D., & DiCarlo, J. J. (2005). Multiple object response normalization in monkey inferotemporal cortex. *Journal of Neuroscience, 25*(36), 8150–8164.

Related Articles

General Principles for Sensory Coding

Deep Neural Networks in Computational Neuroscience

What Is a Neuronal Ensemble?