Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Neurobiology

Models and processes of multisensory cue combination

Robert L Seilheimer[1], Ari Rosenberg[1] and Dora E Angelaki

Fundamental to our perception of a unified and stable environment is the capacity to combine information across the senses. Although this process appears seamless as an adult, the brain's ability to successfully perform multisensory cue combination takes years to develop and relies on a number of complex processes including cue integration, cue calibration, causal inference, and reference frame transformations. Further complexities exist because multisensory cue combination is implemented across time by populations of noisy neurons. In this review, we discuss recent behavioral studies exploring how the brain combines information from different sensory systems, neurophysiological studies relating behavior to neuronal activity, and a theory of neural sensory encoding that can account for many of these experimental findings.

Addresses
Baylor College of Medicine, Houston, TX, United States

Corresponding author: Angelaki, Dora E (angelaki@bcm.edu)
[1] Equal author contribution.

Current Opinion in Neurobiology 2014, **25**:38–46

This review comes from a themed issue on **Theoretical and computational neuroscience**

Edited by **Adrienne Fairhall** and **Haim Sompolinsky**

0959-4388/$ – see front matter,

http://dx.doi.org/10.1016/j.conb.2013.11.008

Introduction

To make sense of a world that is noisy and ambiguous, neural systems combine information across sensory modalities to create unified and stable percepts. Numerous examples highlight the vital role of this process. When driving, we decide whether it is safe to change lanes based on a combination of sights and sounds, our perceived speed, and the force applied to the gas pedal. To better comprehend what someone is saying, we often look at their lips while listening to them speak. If you tilt your head to the side, the scene does not appear rotated because information from the inner ear is used to stabilize your visual perception of the world.

Because the brain often integrates the senses seamlessly, it is easy to overlook the complexities of multisensory cue combination. When presented with two sensory signals (say, light and sound), the brain must determine if they have a common source, reconcile differences in the reference frames in which they are encoded, and integrate information across time to form a coherent percept (Figure 1a). In this review, we discuss how information is combined across senses and examine how theoretical and computational neuroscience has informed our understanding of the neural underpinnings of multisensory cue combination.

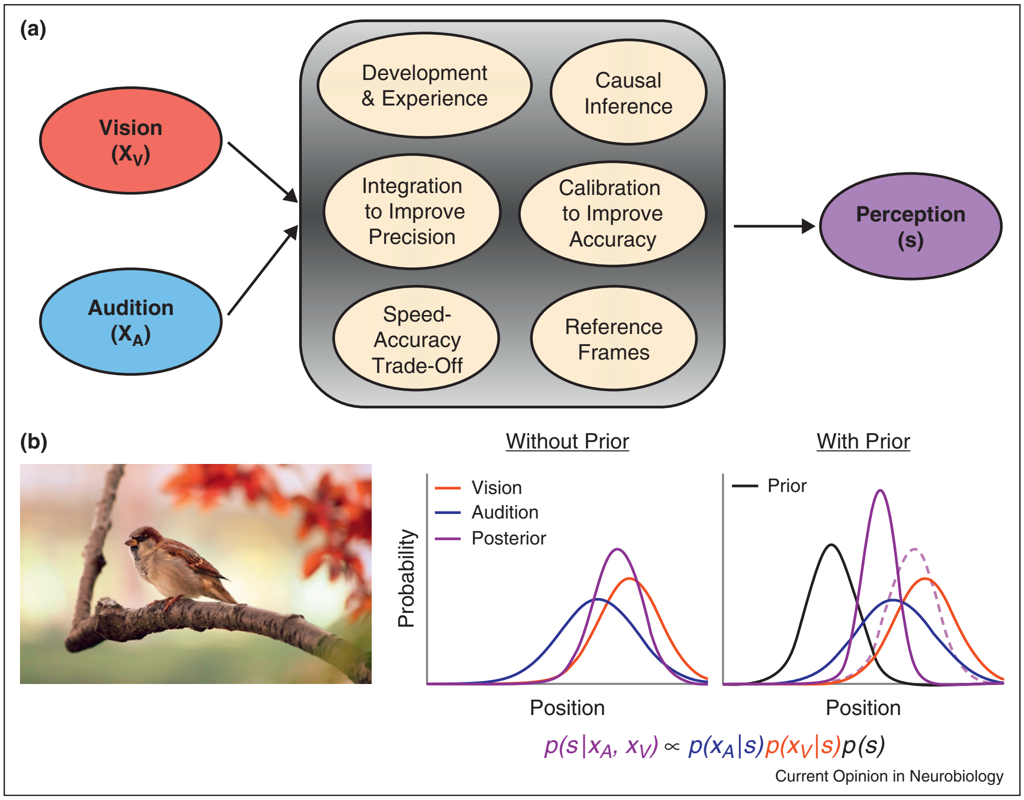
Bayesian cue integration

Because sensory information is noisy and subject to ambiguity, we must infer the state of the world [1]. To improve this inference, information from different senses is combined through multisensory integration. Behavioral studies suggest that sensory signals are often combined in a Bayes-optimal (or nearly optimal) fashion [2,3,4••,5•,6•] to create a probability distribution over the range of possible stimuli that could have given rise to the signals. This process is probabilistic in the sense that the reliability of each sensory cue is taken into account, and Bayesian because prior information can be combined with available sensory information [7,8••,9••] (Figure 1b). Choosing the stimulus with the highest probability results in optimal inference in that it maximizes the observer's precision [10•].

In recent studies, monkeys judging their direction of self-motion were shown to be near-optimal in integrating visual and vestibular information, and to reweight each cue according to its reliability on a trial-by-trial basis [4••,11]. To examine the neural underpinnings of this behavior, the activity of single neurons in the dorsal medial superior temporal area (MSTd) was recorded while the task was performed. These neurons respond to both visual and vestibular signals and were found to modulate their weighting of each cue dynamically with changes in reliability, demonstrating a neural correlate of reliability-based cue combination [4••].

Humans may also be near-optimal in deciding whether or not information *should* be integrated. This process, called causal inference, judges whether different sensory signals (e.g., visual and auditory) originated from either the same or separate sources. Ideally, different sensory signals should be integrated only if they originated from the same source, but otherwise kept separate. To investigate how this inference is performed, one study examined data from human subjects who were presented with synchronized visual flashes and auditory clicks that originated from either the same or different locations, and asked to indicate both the locations of the stimuli and whether they had one or two causes [12]. Behavior in this task could be largely accounted for by a model of Bayesian causal inference in which the probability that two sensory cues have the same underlying cause is computed first

Figure 1



Multisensory cue combination. **(a)** Multisensory combination (e.g., of visual and auditory information) entails a number of processes which the brain learns to implement during development. These processes include, but are not limited to, causal inference to determine if the sensory cues have a common source, integration to improve precision, calibration to improve accuracy, reconciliation of the reference frames in which each sense is encoded, a speed-accuracy trade-off (including accumulation of evidence from each cue across time), and the incorporation of prior information. Together, these processes result in a coherent percept of the sensory stimulus. **(b)** The most well-studied aspect of multisensory cue combination (a general term broadly encompassing situations in which information from different sensory systems is combined) is cue integration, which improves precision. Consider the task of localizing a bird in a tree (s) using auditory (x_A) and visual (x_V) cues. Behavioral experiments suggest that the brain represents each sensory cue probabilistically with a likelihood function – $p(x_A|s)$ and $p(x_V|s)$ – and combines them with prior information $p(s)$ to produce a posterior $p(s|x_A,x_V)$ describing how likely the bird is to be perceived at a particular location. Bayes' rule states that when the noise in each sense is independent, the posterior (purple) is proportional to the product of the likelihood of each sensory cue (blue and red) and the prior (black). The graph on the left shows the likelihood functions for each sensory cue (blue and red) and the resulting posterior (purple). Without prior information, this is equivalent to maximum-likelihood estimation. The graph on the right shows the same sensory likelihood functions, but also includes a prior reflecting past experience (here, a tendency for the bird to be further to the left than the current sensory information suggests). The inclusion of this prior information produces a Bayesian estimate, shifting the posterior to the left (for comparison, the dashed curve re-plots the posterior without prior information).

and then Bayesian cue integration is performed taking into account the observer's belief about the number of causes [13]. However, it should be noted that the reliability of the sensory cues was not varied, leaving it uncertain whether causal inference is indeed implemented optimally. In the next section, we discuss a theoretical framework that describes how neural systems can implement Bayesian inference and multisensory integration.

A theory of how neurons implement multisensory integration

The behavioral observation that cue integration is probabilistic suggests that the brain may directly encode the reliability of sensory information. This led to the investigation of how the brain can simultaneously represent multiple pieces of sensory information along with their reliabilities, and combine them optimally to implement Bayesian cue integration [14].

An intriguing possibility is that this is achieved by populations of neurons whose combined activity describes the likelihood of a sensory input. Given that the inherent variability of neural responses can be described as $p(r|s)$ (i.e., the likelihood that a stimulus s will elicit a population activity r), a neural population can encode a posterior probability distribution over possible stimuli, $p(s|r)$, through Bayes' rule [15]. Specifically, the posterior can be encoded simply through multiplication: $p(s|r) \propto p(r|s)\cdot p(s)$, where $p(s)$ is a prior probability distribution describing how likely particular stimuli are to be encountered. This idea is formalized mathematically by a framework called the Poisson-like probabilistic population code (PPC), in which variability in neural populations follows distributions of the form

$$p(\boldsymbol{r}|s,g) = \phi(\boldsymbol{r},g)\cdot\exp(\boldsymbol{h}(s)\cdot\boldsymbol{r}) \quad (1)$$

where $\boldsymbol{h}(s)$ is a neuronal weighting function, g is the gain of the population (proportional to the reliability of s), and $\phi(\boldsymbol{r},g)$ is a function of the population activity and gain (Box 1). Such distributions have the property that all information about the stimulus s is contained in a weighted linear sum of the population activity ($\boldsymbol{h}(s)\cdot\boldsymbol{r}$), and this information can be decoded by taking the logarithm of $p(\boldsymbol{r}|s,g)$. The weighting function $\boldsymbol{h}(s)$ depends on the neurons' tuning curves and the correlations in the population, but is independent of stimulus reliability (e.g., image contrast). This generalizes a widely used model which assumes that neurons' firing rates are independent and governed by Poisson statistics to allow for, among other things, correlated neural variability [16,17] and different mean–variance relationships in firing rates [18]. Consistent with the defining properties of the Poisson-like PPC described here, recent studies have shown that primary visual cortex can represent stimuli with a linear, contrast-invariant code [19] that takes into account neural correlations [20].

Box 1 The Poisson-like PPC formalizes the idea that variability in neuronal populations reflects the encoding of probability distributions over a set of stimuli. This provides a framework in which many processes of multisensory combination can be performed through biologically plausible computations. To provide a concrete example, we walk through the case of a population of neurons with independent Poisson variability for a fixed gain [14,25••].

For a stimulus *s*, the probability that response *r* (a count of action potentials fired) is elicited from the *i*th neuron in the population is given by a Poisson distribution:

$$p(r_i|s) = \frac{e^{-f_i(s)}\,f_i(s)^{r_i}}{r_i!}$$

where $f_i(s)$ is that neuron's tuning curve over the possible stimuli.

Because probabilities multiply and neural variability is assumed to be independent, the probability of observing a particular population response $\boldsymbol{r}$ is given by the product of the individual $p(r_i|s)$:

$$p(\boldsymbol{r}|s) = \prod_{i=1}^{n}\frac{e^{-f_i(s)}\,f_i(s)^{r_i}}{r_i!}$$

where *n* is the number of neurons in the population.

With some algebra, this can be rewritten in the more general form of the Poisson-like PPC presented in the text (Eq. (1)):

$$p(\boldsymbol{r}|s) = \prod_{i=1}^{n}\frac{e^{-f_i(s)}}{r_i!}\cdot e^{\sum r_i \log(f_i(s))}.$$

In this case,

$$h_i(s) = \log(f_i(s))$$

and assuming that the sum of tuning curves is constant ($\sum_i^n f_i(s) = c$), then

$$\phi(\boldsymbol{r}) = \frac{e^{-c}}{\prod_{i=1}^{n} r_i!}$$

Here ϕ is a function of $\boldsymbol{r}$ (not $\boldsymbol{r}$ and g) because the gain was assumed to be fixed.

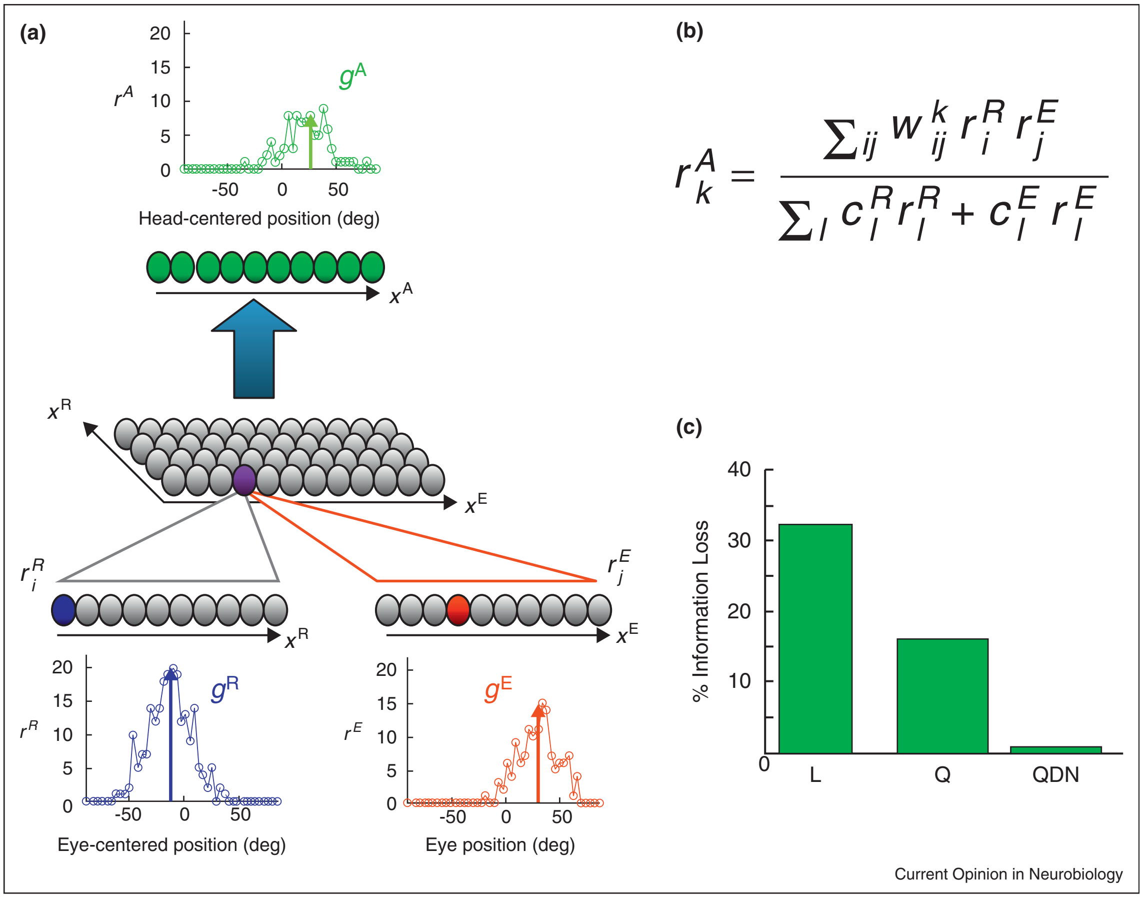
Lastly, following Bayes' rule, the probability of stimulus *s* given a population response $\boldsymbol{r}$ is $p(s|\boldsymbol{r}) \propto p(\boldsymbol{r}|s)\cdot p(s)$, where $p(s)$ is the probability of encountering stimulus *s*.

The Poisson-like PPC provides a surprisingly straightforward neural solution to optimal cue integration. If two neural populations each represent a stimulus through a Poisson-like PPC, then Bayesian cue integration can be achieved by a third population which simply sums the activities of the other two populations [14]. The integration of visual and vestibular signals by MSTd neurons is generally consistent with this framework, but the weights placed on each cue appear to be dependent on reliability [4••]. This finding can, however, be accounted for by incorporating divisive normalization at the level of multisensory integration [21•]. Divisive normalization is a network-level computation found throughout the nervous system that scales the responses of individual neurons by the population activity [22••,23•]. When the multisensory responses of MSTd neurons are modeled as a linear combination of unisensory visual and vestibular responses [4••], divisive normalization can explain the dependency of sensory weights on cue reliability as follows. An increase in the reliability of one sensory cue (say, an increase in visual motion coherence) facilitates the associated unisensory (i.e., visual) response by increasing the response gain [4••,14,24], but has no effect on the other unisensory (i.e., vestibular) response. At the same time, the increase in visual cue reliability also increases the magnitude of the normalization term acting on the population of multisensory neurons. This suppresses both unisensory responses equally at the level of cue integration (i.e., when they are summed together by a multisensory neuron). Combined with the increased gain of the visual response, this directly translates into an increased visual weight and a decreased vestibular weight. Divisive normalization can additionally account for a number of properties of multisensory neurons found in the superior colliculus (SC) [21•], and is important for tasks involving marginalization such as visual search and reference frame transformations [23•,25••,26] (Figure 2).

Reference frame transformations

In primates, the posterior parietal cortex is an important locus of multisensory cue combination. Individual parietal neurons often encode information from multiple senses; for example, neurons in the ventral intraparietal area (VIP) can respond to visual, vestibular, tactile, and auditory stimuli [27–30]. Considering that different sensory systems encode information relative to different egocentric reference frames (e.g., the eyes, head, or body), an important question to ask is: how can information represented in different reference frames be combined?

Figure 2



Poisson-like PPCs and reference frame transformations. **(a)** Neural network performing a reference frame transformation in which the activity of a population of units representing eye position is used to transform an eye-centered representation of object position into a head-centered representation. The network has two input layers (bottom) using Poisson-like PPCs to represent the object's eye-centered position (r^R; blue) and the position of the eyes in the head (r^E; red). The response curves show the activity of each unit for a single object position and eye position. The height of the activity represents the population gain (g^R and g^E), which is proportional to stimulus reliability. The activity of the two input layers is combined by an intermediate layer (middle) that serves as a set of basis functions for computing the object's head-centered position in the output layer at the top (r^A; green). The output layer also encodes object position using a Poisson-like PPC, and its gain (g^A) is less than that of either input layer due to divisive normalization. **(b)** This equation shows that the activity of the output layer units (r^A) can be expressed as a sum of weighted products (a quadratic nonlinearity) of the activity of input layer units (r^E and r^R) divided by the weighted sum of activity in each layer (i.e., divisive normalization). Here, the *w* and *c* terms are weight parameters. **(c)** The percentage of information loss in a simulated neural network, calculated as the difference between the true posterior and that estimated by the output layer, depends on the computations performed by the network. With a quadratic nonlinearity and divisive normalization (QDN) as depicted in **a**, the network loses less than 1% of information. With only the quadratic nonlinearity (Q), there is a 16% loss of information. With neither divisive normalization nor a quadratic nonlinearity (L), there is a 32% loss of information. This demonstrates the importance of both the quadratic nonlinearity and divisive normalization in maintaining information when performing reference frame transformations.
Figure adapted with permission from Beck *et al*. [25••].

While it was previously thought that the brain must remap sensory signals into a common reference frame in order for multisensory cue combination to occur (see [31] for an example) this does not seem to be the general case. For example, single VIP neurons represent tactile signals in a head-centered reference frame, visual signals in a range of intermediate reference frames distributed between eye-centered and head-centered, and vestibular signals in a body-centered reference frame [27,32,33]. Visual and vestibular signals in MSTd are also encoded in different reference frames, with visual signals in an eye-centered frame and vestibular signals in a range of intermediate reference frames distributed between eye-centered and head-centered [32,33,34].

These findings indicate that neural signals need not be in a common reference frame to be combined [25••,35,36]. Computational studies have shown how the activities of two unisensory populations encoding information in different reference frames with Poisson-like PPCs can be combined to form a population of multisensory units that perform optimal statistical inference [25••,27,35]

(Figure 2a). Many of these multisensory units represent sensory information in intermediate reference frames, as observed in MSTd, VIP, and other parietal areas [27,32–34]. A fourth population also implementing a Poisson-like PPC can then combine the activity of the multisensory units to re-express the sensory information in a different reference frame than either of the unisensory populations [25••,35]. Compared to cue integration, performing reference frame transformations with a Poisson-like PPC requires more complex (but widely observed) neural computations including a quadratic nonlinearity (multiplying the activity of neurons) and divisive normalization [25••] (Figure 2b,c). Once these biologically plausible nonlinearities are incorporated into the computations, the Poisson-like PPC framework can account for multiple contemporary observations regarding both optimality and the combination of sensory signals represented in different reference frames.

Whereas sensory information is first encoded relative to egocentric reference frames, the perceptual stability of the environment is suggestive of an allocentric (world-centered) representation in the brain. An object's spatial orientation, for example, is perceived to remain constant relative to the gravitational vector even when your head is tilted to the side. This reflects that the brain uses gravitational (vestibular/proprioceptive) signals to transform the visual representation of the scene from an eye-centered into a world-centered reference frame. Recently, the visual responses of surface orientation selective neurons in the macaque caudal intraparietal area (CIP) [37] were found to be modified by gravitational signals such that surface orientation was encoded in a range of reference frames distributed between head-centered, eye-centered, and world-centered (Rosenberg & Angelaki, abstract in Computational and Systems Neuroscience 2013, Salt Lake City, UT, February 2013). A neural network like the one in Figure 2a reproduced this finding in the intermediate layer and created a purely world-centered representation of surface orientation in the output layer.

Decision making and speed-accuracy trade-off

In many studies, the dynamics of the decision process are hidden because subjects only report a final percept. A common approach to studying how a decision is *formed* is to use a reaction-time paradigm, in which the subjects control when the decision is reported. Previous work using this paradigm showed that observers make trade-offs between speed and accuracy [38] and that more reliable evidence leads to faster decisions [39], suggesting that perceptual evidence is accumulated over time until a decision boundary is reached. The activity of neurons in the macaque lateral intraparietal area (LIP) correlates with this decision process, temporally integrating sensory information until a decision is made [39,40,41••]. A Poisson-like PPC can reproduce this property of LIP neurons [42], and may be superior to other models in describing the decision process in that it allows for moment-to-moment fluctuations in the reliability of sensory evidence and can account for observer uncertainty [43•].

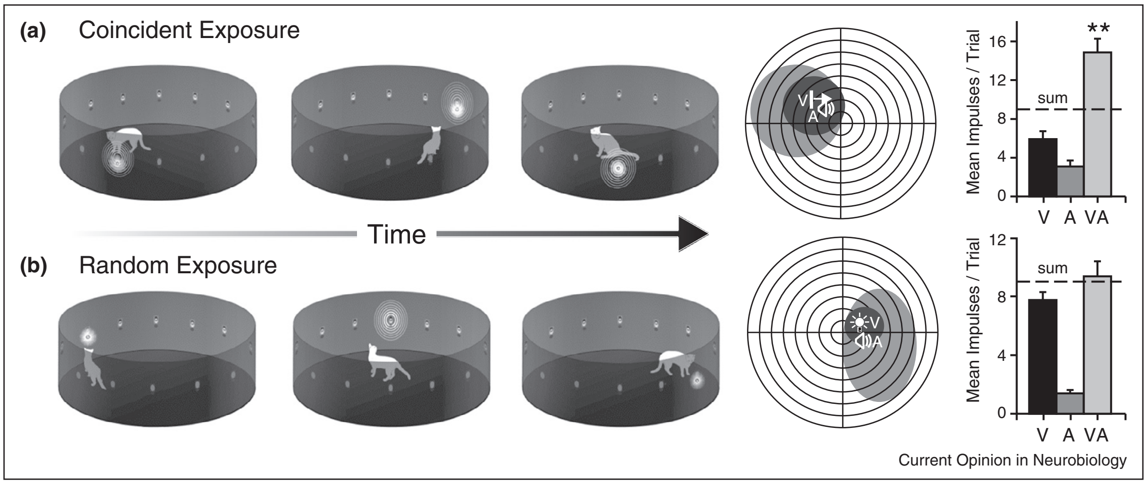
While evidence accumulation is well studied in unisensory perceptual tasks, it is unclear how evidence from multiple senses is accumulated and combined across time. A recent human psychophysical study using a reaction-time version of a heading discrimination task reported that visual–vestibular discrimination thresholds during cue combination were worse than those predicted by conventional optimal cue combination theory, and sometimes worse than that of the more reliable unisensory cue alone (Drugowitsch *et al.*, abstract in Computational and Systems Neuroscience 2011, Salt Lake City, UT, February 2011). This unanticipated result could be explained by a model in which evidence from each sense was weighted and accumulated according to its reliability at each point in time in order to maximize the correct decision rate, implying a more general notion of optimal cue combination which takes into account the time required to integrate information and the time-dependent reliability of the senses. As of now, the neural correlates of these properties remain unexplored.

Development and calibration of multisensory integration

Although Bayesian multisensory integration appears normative in adults, children are far from optimal. Instead, one sense dominates childrens' judgments, suggesting that the brain may forgo multisensory integration while it is learning to calibrate sensory systems relative to each other [44,45•,46]. Consider, for example, the use of vision and touch to perceive an object. Recent studies have shown that children with congenital visual deficits have an impaired ability to determine the object's orientation by touch [47], and children with movement disorders have an impaired ability to visually discriminate the object's size [48]. These studies provide evidence that impairments in one sense may hinder the calibration of another. There is also evidence that cross-sensory calibration is a normative process in adults [49]. For example, during a heading discrimination task, the presentation of discrepant visual and vestibular information leads to a re-calibration of the perceived heading elicited by either sensory signal on its own, with each estimate shifting towards the other [50]. Interestingly, cross-sensory calibration can also influence the interaction of mothers with their newborns: the odor of a newborn mouse pup can induce changes in the auditory cortex of its mother, allowing the mother to better detect the pup's vocalizations [51].

Neurophysiological experiments conducted in the cat SC, a non-cortical locus of multisensory integration for

Figure 3



Development of multisensory integration. **(a)** (Left) Cats in the coincident exposure group were reared in an environment in which auditory and visual stimulation always occurred at coincident locations and times. (Middle) Visual (V) and auditory (A) receptive fields of a superior colliculus neuron from an animal reared in the coincident exposure group. (Right) Unisensory (visual V; auditory A) and multisensory (VA) responses of the same neuron, demonstrating a superadditive response to multisensory stimulation. **(b)** (Left) Cats in the random exposure group were reared in an environment in which auditory and visual stimuli were presented separately at random locations and times. (Middle) Visual and auditory receptive fields of a superior colliculus neuron from an animal reared in the random exposure group. (Right) Unisensory and multisensory responses of the same neuron, demonstrating the lack of a superadditive multisensory response.
Figure adapted with permission from Xu *et al*. [52••].

sensory detection and orienting responses, have illuminated some aspects of the development of multisensory integration. When visual and auditory stimuli are simultaneously presented, neurons in the SC normally display multisensory responses that are superadditive (greater than the sum of the unisensory responses). However, this only develops if the animal is reared in an environment with spatiotemporally coherent multisensory stimulation [52••] (Figure 3). Likewise, when reared in an environment in which multisensory stimuli are only presented with a fixed spatial disparity, SC neurons only develop a superadditive response at that disparity [53]. These results indicate that multisensory integration is learned, but how does it develop?

Ideas originating from machine learning theory and statistics may help us understand how multisensory cue combination develops in the brain. For example, artificial neural networks can be trained to perform Bayesian cue integration and causal inference using reinforcement learning [54]. Specifically, the network learns to optimally combine sensory information by predicting the reward that an action will produce for a given set of sensory information. Another study showed that a class of neural networks called restricted Boltzmann machines (RBMs) can learn optimal cue integration, causal inference, reference frame transformations, and the encoding of priors via density estimation [55•]. This is appealing since density estimation is a statistical technique for learning probability distributions of hidden variables, thus allowing the network to encode posterior distributions. In multisensory integration, an RBM learns to estimate a posterior distribution using a feedback loop in which the multisensory units learn a set of weights capturing all of the relevant information contained in the unisensory units. Such studies thus describe computational mechanisms the brain may use to develop the ability to perform multisensory cue combination in a probabilistic, Bayesian fashion.

Conclusions

In this review we discussed several key components of multisensory cue combination, explored our understanding of each at the behavioral and neural levels, and examined a theoretical framework describing how single neurons might combine sensory information. However, we are far from fully understanding the complexities of how information from different senses is combined. For example, while several studies have considered the influence of naturally occurring priors on perception [7,56–58], little has been done to directly manipulate priors in the study of multisensory integration. Without this manipulation, Bayesian inference is indistinguishable from maximum-likelihood estimation. Thus, full validation of the Bayesian model for multisensory integration (Figure 1b) still requires this manipulation.

Additionally, our focus on the Poisson-like PPC as a theory for how the brain combines sensory information in part reflects that there is currently no clear alternative theory. While the Poisson-like PPC framework does account for several important aspects of multisensory cue combination, some of its underlying assumptions may not always be valid. For example, neural weights appear to depend on stimulus reliability in MSTd [4••,24], and although the theory assumes unisensory representations are independent, sensory interactions may begin before multisensory integration occurs [59•,60,61]. Furthermore, computations like causal inference cannot be performed explicitly using Poisson-like PPCs [62]. Thus, significant challenges in understanding multisensory cue combination remain, but the continuing endeavor to combine experimental neuroscience with computation and theory promises to elucidate this complex process. In the near future, such work is likely to reveal how stable allocentric representations of the environment are created, how evidence is temporally accumulated across multiple senses, and how the brain develops the ability to effectively integrate information from different sensory systems.

Acknowledgements

We thank Greg DeAngelis, Eliana Klier, Wei Ji Ma, and Adhira Sunkara for their comments on the manuscript. This work was supported by NIH grants T32EY007001 (R.L.S) as well as EY019087 and EY022538 (D.E.A.).

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Knill DC, Richards W: *Perception as Bayesian inference*. Cambridge: Cambridge University Press; 1996, .

2. Ernst MO, Banks MS: **Humans integrate visual and haptic information in a statistically optimal fashion**. *Nature* 2002, **415**:429-433.

3. Alais D, Burr D: **The ventriloquist effect results from near-optimal bimodal integration**. *Curr Biol* 2004, **14**:257-262.

4. •• Fetsch CR, Pouget A, DeAngelis GC, Angelaki DE: **Neural correlates of reliability-based cue weighting during multisensory integration**. *Nat Neurosci* 2011, **15**:146-154.
This study provides neural evidence for reliability-based cue reweighting. Using a visual–vestibular cue conflict experiment, the authors demonstrate that monkeys integrate the two sensory signals near optimally. Simultaneously measured neural activity suggests that multisensory neurons reweight the individual sensory signals on a trial-by-trial basis depending on cue reliability.

5. • Raposo D, Sheppard JP, Schrater PR, Churchland AK: **Multisensory decision-making in rats and humans**. *J Neurosci* 2012, **32**:3726-3735.
This paper presents a novel auditory-visual rate discrimination task in which information is integrated over time to judge the rate of events as either ''high'' or ''low.'' The use of rodents to study multisensory decision making offers a number of advantages for future research, such as optogenetic manipulations.

6. • Sheppard JP, Raposo D, Churchland AK: **Dynamic weighting of multisensory stimuli shapes decision-making in rats and humans**. *J Vis* 2013:13.
Using an auditory-visual rate discrimination task, this paper demonstrated dynamic reweighting of multisensory stimuli in rodents. Tests in both humans and rats using a time-varying stimulus further revealed that both species integrate information over the entire stimulus duration.

7. Laurens J, Angelaki DE: **The functional significance of velocity storage and its dependence on gravity**. *Exp Brain Res* 2011, **210**:407-422.

8. •• Laurens J, Meng H, Angelaki DE: **Computation of linear acceleration through an internal model in the macaque cerebellum**. *Nat Neurosci* 2013, **16**:1701-1708.
This study provides evidence that the cerebellum is involved in constructing internal models of physical principles to interpret sensory signals. Using unnatural motion stimuli inducing incorrect self-motion perceptions, Purkinje cells in the caudal cerebellar vermis and neurons in the cerebellar nuclei of the macaque monkey were found to encode erroneous linear acceleration, consistent with the internal model hypothesis.

9. •• Laurens J, Meng H, Angelaki DE: **Neural representation of gravity in the macaque vestibulocerebellum**. *Neuron* 2013 http://dx.doi.org/10.1016/j.neuron.2013.09.029. (in press).
Because gravitational and inertial accelerations are physically indistinguishable (i.e., Einstein's equivalence principle), sensing gravity is a complicated process which is thought to be solved by tracking head orientation through multisensory integration. In this study, the responses of a group of Purkinje cells in the caudal cerebellar vermis of the macaque monkey were shown to reflect an estimate of the orientation of the self relative to gravity.

10. • Ma WJ: **Organizing probabilistic models of perception**. *Trends Cogn Sci* 2012, **16**:511-518.
This review provides a valuable guide explaining probabilistic, Bayesian, and optimal models of perception. The author defines each of these models, describes how they differ, and discusses supporting experimental results.

11. Gu Y, Angelaki DE, DeAngelis GC: **Neural correlates of multisensory cue integration in macaque MSTd**. *Nat Neurosci* 2008, **11**:1201-1210.

12. Kording KP, Beierholm U, Ma WJ, Quartz S, Tenenbaum JB, Shams L: **Causal inference in multisensory perception**. *PLoS ONE* 2007, **2**:e943.

13. Shams L, Beierholm UR: **Causal inference in perception**. *Trends Cogn Sci* 2010, **14**:425-432.

14. Ma WJ, Beck JM, Latham PE, Pouget A: **Bayesian inference with probabilistic population codes**. *Nat Neurosci* 2006, **9**:1432-1438.

15. Zemel RS, Dayan P, Pouget A: **Probabilistic interpretation of population codes**. *Neural Comput* 1998, **10**:403-430.

16. Zohary E, Shadlen MN, Newsome WT: **Correlated neuronal discharge rate and its implications for psychophysical performance**. *Nature* 1994, **370**:140-143.

17. Gu Y, Liu S, Fetsch CR, Yang Y, Fok S, Sunkara A, DeAngelis GC, Angelaki DE: **Perceptual learning reduces interneuronal correlations in macaque visual cortex**. *Neuron* 2011, **71**:750-761.

18. Gur M, Snodderly DM: **High response reliability of neurons in primary visual cortex (V1) of alert, trained monkeys**. *Cereb Cortex* 2006, **16**:888-895.

19. Berens P, Ecker AS, Cotton RJ, Ma WJ, Bethge M, Tolias AS: **A fast and simple population code for orientation in primate V1**. *J Neurosci* 2012, **32**:10618-10626.

20. Graf AB, Kohn A, Jazayeri M, Movshon JA: **Decoding the activity of neuronal populations in macaque primary visual cortex**. *Nat Neurosci* 2011, **14**:239-245.

21. • Ohshiro T, Angelaki DE, DeAngelis GC: **A normalization model of multisensory integration**. *Nat Neurosci* 2011, **14**:775-782.
This computational paper shows that divisive normalization can account for several properties of multisensory neurons in the superior colliculus and area MSTd, including reliability-based cue reweighting (see Ref. [4••]).

22. •• Carandini M, Heeger DJ: **Normalization as a canonical neural computation**. *Nat Rev Neurosci* 2011, **13**:51-62.
This review provides a thorough treatment of divisive normalization in sensory representations, covering physiological evidence that it occurs at multiple levels of the nervous system and across multiple species, potential cellular circuits that could implement the computation, and behavioral evidence that divisive normalization contributes to perception.

23. • Louie K, Khaw MW, Glimcher PW: **Normalization is a general neural mechanism for context-dependent decision making**. *Proc Natl Acad Sci U S A* 2013, **110**:6139-6144.
Motivated by electrophysiological results on divisive normalization, this paper presents a divisive normalization-based model for decision making. The authors show that this model can account for non-rational human behavior seen in value-based decisions.

24. Morgan ML, DeAngelis GC, Angelaki DE: **Multisensory integration in macaque visual cortex depends on cue reliability**. *Neuron* 2008, **59**:662-673.

25. Beck JM, Latham PE, Pouget A: **Marginalization in neural**
•• **circuits with divisive normalization**. *J Neurosci* 2011, **31**:15310-15319.
This theoretical study shows how neurons implementing a Poisson-like PPC can perform marginalization using two canonical neural phenomena: divisive normalization and quadratic nonlinearities.

26. Mazyar H, van den Berg R, Seilheimer RL, Ma WJ: **Independence is elusive: set size effects on encoding precision in visual search**. *J Vis* 2013:13.

27. Avillac M, Deneve S, Olivier E, Pouget A, Duhamel JR: **Reference frames for representing visual and tactile locations in parietal cortex**. *Nat Neurosci* 2005, **8**:941-949.

28. Schlack A, Sterbing-D'Angelo SJ, Hartung K, Hoffmann KP, Bremmer F: **Multisensory space representations in the macaque ventral intraparietal area**. *J Neurosci* 2005, **25**:4616-4625.

29. Chen A, DeAngelis GC, Angelaki DE: **Representation of vestibular and visual cues to self-motion in ventral intraparietal cortex**. *J Neurosci* 2011, **31**:12036-12052.

30. Chen A, DeAngelis GC, Angelaki DE: **Functional specializations of the ventral intraparietal area for multisensory heading discrimination**. *J Neurosci* 2013, **33**:3567-3581.

31. Cohen YE, Andersen RA: **A common reference frame for movement plans in the posterior parietal cortex**. *Nat Rev Neurosci* 2002, **3**:553-562.

32. Chen X, DeAngelis GC, Angelaki DE: **Diverse spatial reference frames of vestibular signals in parietal cortex**. *Neuron* 2013, **80**:1310-1321 (in press).

33. Chen X, DeAngelis GC, Angelaki DE: **Eye-centered representation of optic flow tuning in the ventral intraparietal area**. *J Neurosci* 2013, **33**:18574-18582.

34. Fetsch CR, Wang S, Gu Y, DeAngelis GC, Angelaki DE: **Spatial reference frames of visual, vestibular, and multimodal heading signals in the dorsal subdivision of the medial superior temporal area**. *J Neurosci* 2007, **27**:700-712.

35. Deneve S, Latham PE, Pouget A: **Efficient computation and cue integration with noisy population codes**. *Nat Neurosci* 2001, **4**:826-831.

36. De Meyer K, Spratling MW: **A model of partial reference frame transforms through pooling of gain-modulated responses**. *Cereb Cortex* 2013, **23**:1230-1239.

37. Rosenberg A, Cowan NJ, Angelaki DE: **The visual representation of 3D object orientation in parietal cortex**. *J Neurosci* 2013, **33**:19352-19361.

38. Palmer J, Huk AC, Shadlen MN: **The effect of stimulus strength on the speed and accuracy of a perceptual decision**. *J Vis* 2005, **5**:376-404.

39. Roitman JD, Shadlen MN: **Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task**. *J Neurosci* 2002, **22**:9475-9489.

40. Kiani R, Shadlen MN: **Representation of confidence associated with a decision by neurons in the parietal cortex**. *Science* 2009, **324**:759-764.

41. Churchland AK, Kiani R, Chaudhuri R, Wang XJ, Pouget A,
•• Shadlen MN: **Variance as a signature of neural computations during decision making**. *Neuron* 2011, **69**:818-831.
In this paper, the authors introduce a measure called the variance of conditional expectation (VarCE) to analyze variability in neuronal responses. The VarCE measure removes variability attributable to the stochastic nature of spiking to isolate the variability component attributable to underlying changes in the quantity a neuron's firing rate represents. Using this measure, they uncover changes in neural state throughout the course of a decision making task.

42. Beck JM, Ma WJ, Kiani R, Hanks T, Churchland AK, Roitman J, Shadlen MN, Latham PE, Pouget A: **Probabilistic population codes for Bayesian decision making**. *Neuron* 2008, **60**:1142-1152.

43. Drugowitsch J, Pouget A: **Probabilistic vs. non-probabilistic**
• **approaches to the neurobiology of perceptual decision-making**. *Curr Opin Neurobiol* 2012, **22**:963-969.
This review covers experimental work and computational modeling of evidence accumulation for making perceptual decisions. The authors conclude that the Poisson-like PPC is superior to drift-diffusion models and neural networks with attractor dynamics in modeling perceptual decision-making.

44. Gori M, Del Viva M, Sandini G, Burr DC: **Young children do not integrate visual and haptic form information**. *Curr Biol* 2008, **18**:694-698.

45. Gori M, Giuliana L, Sandini G, Burr D: **Visual size perception and**
• **haptic calibration during development**. *Dev Sci* 2012, **15**:854-862.
This comprehensive study examined visual–haptic integration throughout development. The authors tested size discrimination in children and adults and found that visual size discrimination is better when the object is within arm's reach, providing evidence for haptic calibration of vision. For objects outside of arms' reach, the authors found age-related changes in visual biases from underestimation to overestimation of object size.

46. Nardini M, Begus K, Mareschal D: **Multisensory uncertainty reduction for hand localization in children and adults**. *J Exp Psychol Hum Percept Perform* 2013, **39**:773-787.

47. Gori M, Sandini G, Martinoli C, Burr D: **Poor haptic orientation discrimination in nonsighted children may reflect disruption of cross-sensory calibration**. *Curr Biol* 2010, **20**:223-225.

48. Gori M, Tinelli F, Sandini G, Cioni G, Burr D: **Impaired visual size-discrimination in children with movement disorders**. *Neuropsychologia* 2012, **50**:1838-1843.

49. Zaidel A, Ma WJ, Angelaki DE: **Supervised calibration relies on the multisensory percept**. *Neuron* 2013 http://dx.doi.org/10.1016/j.neuron.2013.09.026.

50. Zaidel A, Turner AH, Angelaki DE: **Multisensory calibration is independent of cue reliability**. *J Neurosci* 2011, **31**:13949-13962.

51. Cohen L, Rothschild G, Mizrahi A: **Multisensory integration of natural odors and sounds in the auditory cortex**. *Neuron* 2011, **72**:357-369.

52. Xu J, Yu L, Rowland BA, Stanford TR, Stein BE: **Incorporating**
•• **cross-modal statistics in the development and maintenance of multisensory integration**. *J Neurosci* 2012, **32**:2287-2298.
This paper demonstrated that spatiotemporally coherent multisensory stimulation is required for the development of superadditive multisensory responses in the superior colliculus. Cats were reared in environments in which auditory and visual stimuli were either spatiotemporally coincident or not. Only cats that were exposed to coincident auditory and visual stimulation developed neurons with superadditive multisensory responses, reflecting that multisensory integration is a developmental phenomenon.

53. Wallace MT, Stein BE: **Early experience determines how the senses will interact**. *J Neurophysiol* 2007, **97**:921-926.

54. Weisswange TH, Rothkopf CA, Rodemann T, Triesch J: **Bayesian cue integration as a developmental outcome of reward mediated learning**. *PLoS ONE* 2011, **6**:e21575.

55. Makin JG, Fellows MR, Sabes PN: **Learning multisensory**
• **integration and coordinate transformation via density estimation**. *PLoS Comput Biol* 2013, **9**:e1003035.
In this work, the authors use a restricted Boltzmann machine implementing density estimation (a technique for estimating parameters of a probability distribution) to model how several key features of multisensory combination including cue integration, reference frame transformations, and the encoding of prior information can arise through a single developmental process.

56. Stocker AA, Simoncelli EP: **Noise characteristics and prior expectations in human visual speed perception**. *Nat Neurosci* 2006, **9**:578-585.

57. MacNeilage PR, Banks MS, Berger DR, Bulthoff HH: **A Bayesian model of the disambiguation of gravitoinertial force by visual cues**. *Exp Brain Res* 2007, **179**:263-290.

58. De Vrijer M, Medendorp WP, Van Gisbergen JA: **Shared computational mechanism for tilt compensation accounts for biased verticality percepts in motion and pattern vision**. *J Neurophysiol* 2008, **99**:915-930.

59. Kim R, Peters MA, Shams L: **0 + 1 > 1: How adding**
• **noninformative sound improves performance on a visual task**. *Psychol Sci* 2012, **23**:6-12.
This paper used behavioral evidence to argue for the existence of interactions between the senses at the level of sensory representations, before they are combined. The authors had subjects perform a two-alternative forced choice task for detecting visual motion in the presence of various types of auditory stimuli (congruent/incongruent with the visual stimulus and informative/non-informative about the visual stimulus). They found that a congruent, non-informative auditory stimulus improved performance on the visual motion task.

60. Hedger SC, Nusbaum HC, Lescop O, Wallisch P, Hoeckner B: **Music can elicit a visual motion aftereffect**. *Atten Percept Psychophys* 2013, **75**:1039-1047.

61. van der Groen O, van der Burg E, Lunghi C, Alais D: **Touch influences visual perception with a tight orientation-tuning**. *PLOS One* 2013, **8**:e79558.

62. Ma WJ, Rahmati M: **Towards a neural implementation of causal inference in cue combination**. *Multisensory Res* 2013, **26**:159-176.