UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

The circuit architecture of cortical multisensory processing: Distinct functions jointly operating within a common anatomical network

Meijer, G.T.; Mertens, P.E.C.; Pennartz, C.M.A.; Olcese, U.; Lansink, C.S.

DOI
10.1016/j.pneurobio.2019.01.004

Publication date
2019

Document Version
Final published version

Published in
Progress in Neurobiology

License
CC BY-NC-ND

Link to publication

Citation for published version (APA):
Meijer, G. T., Mertens, P. E. C., Pennartz, C. M. A., Olcese, U., & Lansink, C. S. (2019). The circuit architecture of cortical multisensory processing: Distinct functions jointly operating within a common anatomical network. *Progress in Neurobiology*, *174*, 1-15. https://doi.org/10.1016/j.pneurobio.2019.01.004

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:08 Nov 2026

Progress in Neurobiology 174 (2019) 1–15

Contents lists available at ScienceDirect

Progress in Neurobiology

journal homepage: www.elsevier.com/locate/pneurobio

Review article

The circuit architecture of cortical multisensory processing: Distinct functions jointly operating within a common anatomical network

Guido T. Meijer[a], Paul E.C. Mertens[a], Cyriel M.A. Pennartz[a,b], Umberto Olcese[a,b,1,*], Carien S. Lansink[a,b,1,*]

[a] Swammerdam Institute for Life Sciences, University of Amsterdam, Science Park 904, 1098XH Amsterdam, the Netherlands
[b] Research Priority Program Brain and Cognition, University of Amsterdam, Science Park 904, 1098XH Amsterdam, the Netherlands

ARTICLE INFO

Keywords:
Multisensory processing
Multisensory integration
Cue integration
Cross-modal
Sensory systems

ABSTRACT

Our perceptual systems continuously process sensory inputs from different modalities and organize these streams of information such that our subjective representation of the outside world is a unified experience. By doing so, they also enable further cognitive processing and behavioral action. While cortical multisensory processing has been extensively investigated in terms of psychophysics and mesoscale neural correlates, an in depth understanding of the underlying circuit-level mechanisms is lacking. Previous studies on circuit-level mechanisms of multisensory processing have predominantly focused on *cue integration,* i.e. the mechanism by which sensory features from different modalities are combined to yield more reliable stimulus estimates than those obtained by using single sensory modalities. In this review, we expand the framework on the circuit-level mechanisms of *cortical* multisensory processing by highlighting that multisensory processing is a family of functions – rather than a single operation – which involves not only the integration but also the segregation of modalities. In addition, multisensory processing not only depends on stimulus features, but also on cognitive resources, such as attention and memory, as well as behavioral context, to determine the behavioral outcome. We focus on rodent models as a powerful instrument to study the circuit-level bases of multisensory processes, because they enable combining cell-type-specific recording and interventional techniques with complex behavioral paradigms. We conclude that distinct multisensory processes share overlapping anatomical substrates, are implemented by diverse neuronal micro-circuitries that operate in parallel, and are flexibly recruited based on factors such as stimulus features and behavioral constraints.

1. Introduction

Throughout our lives, we are continuously exposed to a wealth of sensory information of various modalities, such as vision, audition and touch. Information coming from these different sensory modalities is processed by our brains to generate a subjective representation of the outside world that is a unified percept, and the development of this representation also involves cognitive processes such as learning, decision making and selective attention to operate across the multiple sensory modalities. Multisensory processing (MP) has been extensively investigated for over a century, primarily via studies on human subjects (Todd, 1912; Raab, 1962; Diederich and Colonius, 2004; Spence, 2011). Hence, the neuronal basis of MP has been mostly mapped at the mesoscopic level with techniques such as EEG and fMRI (Calvert and Thesen, 2004; Murray et al., 2016a). Therefore, our understanding of the circuit-level mechanisms underlying MP is still limited. The development of methods to record and manipulate the activity of neuronal circuits with cell-type specificity *in vivo* makes it now possible to overcome this. Rodents are particularly suitable to investigate MP at the micro-scale, because an array of behavioral tasks has recently been developed that may be combined with neurophysiological recordings at cellular resolution (Carandini and Churchland, 2013; Raposo et al., 2014; Lee et al., 2016; Song et al., 2017; Meijer et al., 2018; Nikbakht et al., 2018). Our goal is to review the current advancements on the

Abbreviations: MLE, Maximum-Likelihood Estimate; MP, multisensory processing; MSTd, medial superior temporal lobe; PPC, posterior parietal cortex; PV, parvalbumine positive; SC, superior colliculus

* Corresponding authors at: Swammerdam Institute for Life Sciences, University of Amsterdam, Science Park 904, 1098XH Amsterdam, the Netherlands.
E-mail addresses: g.t.meijer@uva.nl (G.T. Meijer), p.e.c.mertens@uva.nl (P.E.C. Mertens), c.m.a.pennartz@uva.nl (C.M.A. Pennartz), u.olcese@uva.nl (U. Olcese), c.s.lansink@uva.nl (C.S. Lansink).

[1] Equal author contribution.

https://doi.org/10.1016/j.pneurobio.2019.01.004
Received 21 December 2017; Received in revised form 21 December 2018; Accepted 21 January 2019
Available online 21 January 2019
0301-0082/ © 2019 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/BY-NC-ND/4.0/).

cortical, neuron-level mechanisms of MP and propose the systems neuroscience approach, with rodents as the prime model, to investigate how the neocortex combines sensory stimuli of different modalities.

At the neuronal level, MP has primarily been investigated in terms of *cue integration*. This is a mechanism by which the integration of external cross-modal cues can provide a more reliable estimate of an object or event as compared to modality-specific (stand-alone) cues, leading to behavioral benefits such as responding faster and more accurately in a given situation (Gielen et al., 1983; Gingras et al., 2009). For instance, when we simultaneously see a car moving and hear the looming sound of an engine, we can rapidly integrate these streams of information and attribute them to the same event, i.e. an approaching car. It has to be kept in mind, however, that cue integration can be expressed in different forms, such as the detection versus discrimination of stimuli from different modalities. In addition, cue integration is only one of the many types of MP. For instance, MP includes also sensory selection, a form of attention in which processing of one sensory modality is prioritized at the expense of others (Lakatos et al., 2008; Schroeder and Lakatos, 2009; Wimmer et al., 2015; Ahrens et al., 2015), but also processes such as object recognition impinging on multiple sensory modalities. Moreover, MP includes not only integration of cross-modal stimuli, but also segregation, for instance to account for the qualitatively varied nature of conscious experience (2009; Pennartz, 2015) - Box 3. Thus, not all multisensory processes can be modeled in terms of simple integration of external sensory features.

The neural mechanisms of multisensory cue integration have historically been investigated in the model system of the cat superior colliculus (SC) (Wallace et al., 1993;Stein and Stanford, 2008; Stein et al., 2009a), a phylogenetically ancient component of the vertebrate midbrain that plays a key role in directing eye movements and other orientation responses towards a target stimulus (Bell et al., 2005). The multisensory coding principles derived from research on cue integration in the cat SC inspired the search for neuronal principles of MP in cortical areas. Single unit responses specific to cross-modal stimulus combinations were first observed in cortical association areas, where projections from sensory cortices converge across several mammalian species including rodents (Wallace et al., 2004), carnivores (Jiang et al., 2001, 2002; Stein and Stanford, 2008) and non-human primates (Bruce et al., 1981; Beauchamp, 2005; Ghazanfar and Schroeder, 2006; Cappe et al., 2009). These findings were extended by other studies indicating that multisensory responses occur even in primary sensory cortices (Bizley et al., 2007; Kayser and Logothetis, 2007; Lakatos et al., 2008; Kayser et al., 2008; Ghazanfar and Lemus, 2010; Lemus et al., 2010; Iurilli et al., 2012; Ibrahim et al., 2016). Altogether, this prompted some authors to ask whether the whole neocortex should be considered multisensory (Ghazanfar and Schroeder, 2006).

Provocative as this question is, because it challenges the traditional view of cortical anatomical parcellation and segregation of functions, in essence it invites to embrace a view on the functional implications of multisensory processing and its neuronal substrates that is broader than just cue integration (see Fig. 1 and Box 1). This view is implicit to the definition of multisensory or cross-modal integration as any process in which information across sensory modalities is combined to make a perceptual inference (Bizley et al., 2016). This definition encompasses functions which both precede and incorporate cue integration and other forms of MP, such as sensory selection. Altogether, multisensory processes are crucial components of higher order cognitive processes including memory, decision making and consciousness (Raposo et al., 2012; Pennartz, 2015; Jacklin et al., 2016; Licata et al., 2017). At the same time, these cognitive processes feed back into MP (Fig. 1). Here we will refer to the larger family of multisensory processes as MP.

The detection and discrimination of stimulus features may be considered to be largely stimulus driven, bottom-up processes which are mediated by feedforward projections from sensory to association areas, and by lateral connectivity between sensory areas. In contrast, sensory selection was shown to depend on top-down modulation originating in

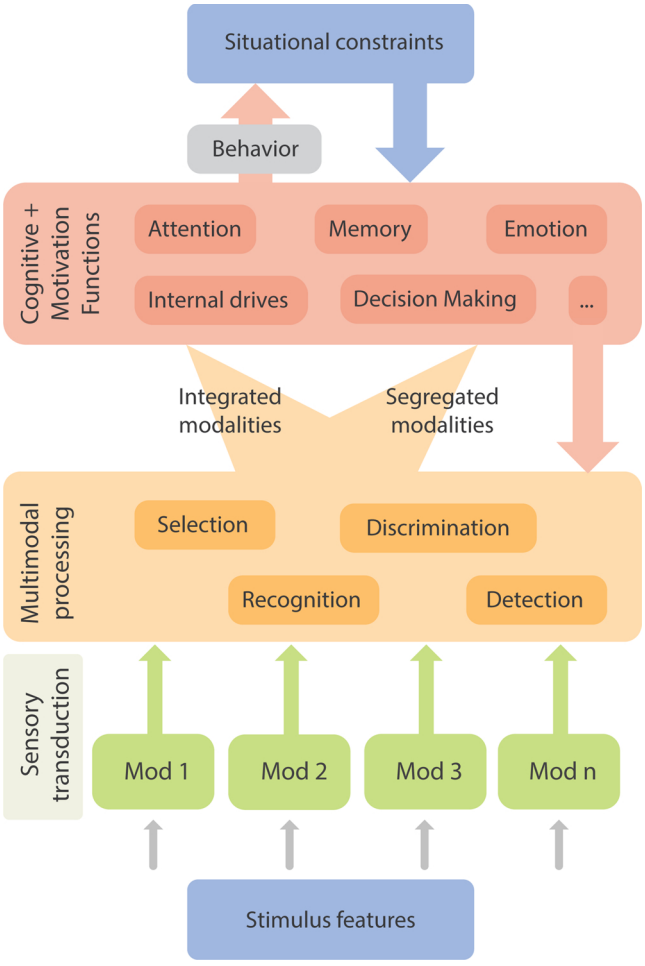

Fig. 1. Multisensory processing is determined by the features of sensory stimuli and relevant cognitive resources. Multisensory processing results in either the integration of stimulus features, attributing them to the same object or event, or the segregation of these features leading to distinct representations associated with multiple objects (yellow module). The outcome of multisensory processes such as detection, discrimination, recognition and selection is first determined by the neural representations of the available stimulus features transmitted in a bottom-up fashion by the sensory systems (green modules; Mod = sensory modality). For example, a salient object that is detected in the surroundings of an individual will attract attention via bottom-up mechanisms. The second factor that determines the outcome of multisensory processing comprise the (higher-order) cognitive and motivational (value-related) functions (red module), which can recruit resources in a top-down fashion in order to deal with specific situational constraints that arise as the animal interacts with its environment. For example, representations of an object consisting of relationships between features of different sensory modalities acquired through experience may aid recognition of an object or location. Perception of an object is considered to be the result of the whole processing cascade illustrated in this scheme.

the medial prefrontal cortex and impinging on primary sensory cortices via the reticular thalamic formation (Birrell and Brown, 2000; Wimmer et al., 2015; Ahrens et al., 2015). In this review we will expand the framework on cortical MP at the neuronal level beyond cue integration, by addressing the question of how multisensory contributions to cognitive functions expressed in the subject's behavior are supported by basic neuronal, local-circuit and population mechanisms. Compared to the SC, the neuronal architecture and dynamics at the local level as well as the inter-area connectivity are much more elaborate in the cortex (Cappe et al., 2009). This requires additional arrangements of multisensory processing mechanisms within and between brain areas. Importantly, rodents are an exceptionally suited experimental model to

Box 1

To integrate or to segregate? A hierarchical Bayesian inference approach.

When crossing an intersection, you are tasked with judging the location of approaching cars in order to determine when it is safe to cross the street. Your perceptual systems receive a myriad of sensory inputs from several senses which should be correctly assigned to distinct external stimuli. For example, you might receive the visual input of an approaching car and also hear the sound of a horn. When these two sensory inputs originate from the same physical entity in the outside world, the brain must integrate these two inputs into a combined percept. However, in case these two sensory stimuli originate from different entities, the brain should not combine the information from these sensory domains. In this example, the decision whether to integrate can be of vital importance because if the sound of the horn actually originates from another car which is coming from the other side of the road this could result in a dangerous accident.

In the situation in which the two sensory inputs originate from the same external entity (i.e. the car that you see also honks; *S* in Fig. Box 1), the way by which these inputs can be integrated into a combined estimate can be described by Bayesian optimal cue integration (Ernst and Bülthoff, 2004). The collection of estimates of the sensory systems -about the location of the car are typically described as a normal distribution in location space (Fig. Box 1). The visual estimate is depicted here as a blue Gaussian/bell-shaped curve and the auditory estimate in red, and the reliability of the estimates is reflected by the inverse of the standard deviation of the distributions. The theory of Bayesian optimal cue integration states that the statistically optimal combination of estimates from individual sensory systems, i.e. the Maximum-Likelihood Estimate (MLE), is equal to the sum of these estimates, each weighted according to the reliability of the corresponding sensory system. The resulting optimal combined estimate -about the location of the stimulus (purple line in Fig. Box 1) is more reliable compared to both unisensory estimates, as is reflected in a lower standard deviation of the normal distribution. According to the MLE rule, the reliability of the combined multisensory estimate is the sum of each unisensory reliability and is therefore by definition more reliable. Theoretical and experimental work has shown that this framework is applicable both at the psychophysical (Ernst and Banks, 2002) and neuronal level (Ma et al., 2006; Fetsch et al., 2012).

Now consider the situation in which the two sensory inputs originate from two different external entities (S_V and S_A in Fig. Box 1); i.e., the honking originates from a different car which was approaching from the other side of the road, outside of your field of view. Both this and the previous scenario present the nervous system with a two-stage problem. First, it must determine whether the sensory signals originated from the same object and subsequently it must perform either integration or segregation of the visual and auditory signal. This problem can be formalized as a hierarchical Bayesian inference problem (Shams and Beierholm, 2010) in which there are two possible schemes: the signals originate from a common source ($C = 1$) or from different sources ($C = 2$). In the case of assumed unity, the combined estimate $\hat{S}$ is determined as described above whereas in the case of multiple objects the two unisensory estimates are not combined and give rise to two independent estimates of the location of two objects ($\hat{S}_V$ & $\hat{S}_A$).

Whether to integrate or segregate can be determined by computing the probability of either scenario (single or dual sources). The probability that two sensory signals (X_V & X_A) originate from the same source (S) depends on the similarity between the two signals, for example, when the visual and auditory system both give a similar estimate of the location of the stimulus it is likely that these signals originated from the same object. In case of ambiguity, resampling of the sensory inputs – aided by eye, ear or head movements, is desirable. Moreover, the timing of the two sensory signals can determine the likelihood that they should be fused: for example, stimuli that are separated in time by more than 80 ms are likely to be segregated (Meredith et al., 1987). Furthermore, the probability of a single or of multiple causes is dependent on the prior expectation that the observed signals originated from the same object (e.g. the sound of a train has a lower a priori likelihood to originate from a car). All these factors contribute to the evidence favoring either a single source (evidence $C = 1 > C = 2$) or dual sources (evidence $C = 1 < C = 2$). A possible neural substrate which serves this function can be found in the posterior parietal cortex, where the mismatch between information originating from the visual and auditory domain is computed through feed-forward inhibition originating from the respective primary sensory cortices (Song et al., 2017).

investigate the nature of these processes. We pursue our main aim by considering a scheme in which the outcome of multisensory processing is dependent on both physical stimulus properties (low level factors), cognitive resources and situational demands (high level factors) - Fig. 1 (Talsma et al., 2010; Pennartz, 2015; De Meo et al., 2015; Murray et al., 2016b; ten Oever et al., 2016). While similar schemes were previously discussed mostly in terms of cognitive processes or mesoscopic brain dynamics – see e.g. (Talsma et al., 2010; ten Oever et al., 2016) – we will primarily focus on the mechanistic implementation of MP at the level of neuronal circuits. Therefore, we review studies using neurophysiological techniques with single cell resolution predominantly using rodent subjects. We will refer to studies with cats and non-human primates if no rodent data are available in the literature, or if results have been seminal to the field. However, we do provide links to MP research with human subjects to discuss how animal research can inform studies on corresponding processes in the human brain. Most of the research discussed here pertains to visual, auditory and somatosensory processing, but also other sensory modalities, such as odor and taste, have been shown to interact with each other in the cortex (Maier et al., 2015; Vincis and Fontanini, 2016).

In the next section we discuss how integration versus segregation of multisensory information is expressed in the behavior of animals. Then, we investigate the neuronal responses and mechanisms at the single neuron, local circuit and population levels associated with various low- and high-level factors that are fundamental to shaping the various behaviors that impinge on MP.

2. The behavioral relevance of multisensory processing

The availability of validated behavioral tasks probing multisensory processes is essential for gaining understanding of the neuronal mechanisms associated with various forms of MP. In the scheme shown in Fig. 1, the yellow module represents the outcome of MP: information conveyed by different sensory organs can be integrated and treated as a unified representation or separated into two or more distinct representations. Multisensory processes such as stimulus detection, discrimination and selection, which lead to inferences based on sensory information, operate in this module; i.e., the resulting representations are based on the outcome of the integration/segregation process. Ultimately, these representations are then flexibly used within the context of higher order functions such as decision making and memory formation. In this section we discuss the recent advancements in the development of behavioral tasks for animals which directly test factors that determine the outcome of multisensory operations.

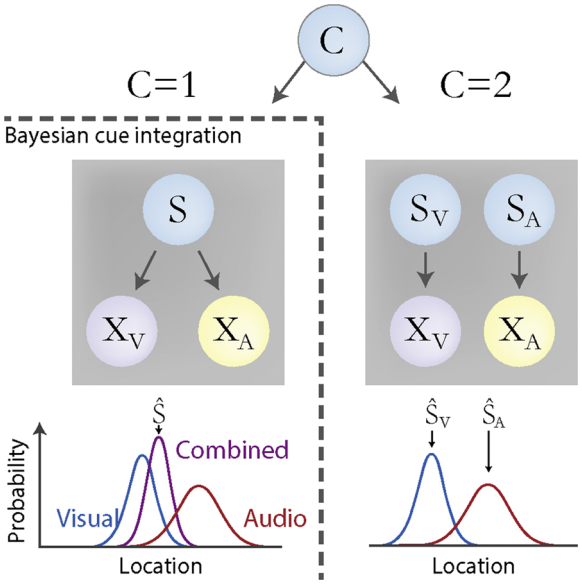

Fig. Box 1. A hierarchical Bayesian inference framework to determine whether to integrate or to segregate. Whether a visual sensory signal (X_V; purple) and an auditory sensory signal (X_A; *yellow*) originate from the same source (S) or from two independent sources (S_V & S_A; *blue*) can be determined by Bayesian inference. The evidence favoring unity of the two signals ($C = 1$) is compared to the evidence in favor of two separate sources ($C = 2$) by a hidden variable C. This evidence is dependent on temporal, spatial and contextual congruency between the two signals. In the former case, localizing a single object in space based on information from multiple modalities can be described by the Bayesian optimal cue integration framework (Ernst and Banks, 2002). A visual (blue curve) and an auditory source (red) both provide information -about the location of a single object (left panel). According to the maximum likelihood estimate rule, the optimal combination ($\hat{S}$) of these two sources of information is a sum of both estimates which is weighted by their respective reliabilities. In the latter scenario (right panel), the visual and auditory signal both originate from a different source (S_V & S_A respectively) in which case they should be segregated instead of integrated resulting in separate visual ($\hat{S}_V$) and auditory ($\hat{S}_A$) estimates.

The first class of behavioral tasks investigates MP that is primarily stimulus driven, i.e. directed by external stimuli and therefore involving mostly bottom-up processing (Fig. 1, *bottom half*). These tasks are mainly accounted for by different forms of cue integration. In *stimulus detection*, the availability of multisensory processing may reduce the uncertainty about the presence or identity of an external object. Indeed, stimulus detection has been shown to improve when rodents use stimuli from more than one sensory modality, resulting in lower detection thresholds (Fig. 2a) and shorter reaction times (Hirokawa et al., 2008; Cappe et al., 2010; Gleiss and Kayser, 2012; Hollensteiner et al., 2015; Siemann et al., 2015; Meijer et al., 2018), especially when stimulus amplitudes in both modalities are around the threshold for detection (Meijer et al., 2018). This mechanism is ecologically relevant, for example both predators and animals of prey, such as rodents, employ stimulus detection to spot their opponent (Stein and Stanford, 2008). The mechanisms of stimulus detection have also been extensively characterized in humans – see for instance (Raab, 1962; Todd, 1912; Diederich and Colonius, 2004).

It should be noted, however, that most cue integration paradigms contrast the behavior in relation to composite sensory stimuli to behavioral responses to the modality-specific stimulus constituents. This procedure entails a higher combined stimulus saliency (here intended as a compound term to indicate stimulus intensity: volume, luminosity, contrast, etc.) in the cross-modal compared to the unimodal condition, which may explain the changes in detection threshold and reaction time. A study by Gingras et al. (2009), however, showed that the increase in accuracy with which cats responded to composite stimuli compared to modality-specific stimuli was much larger for audio-visual stimuli than for within modality (visual-visual or auditory-auditory) combined stimuli. These results indicate that multisensory cue integration increases the behavioral performance to an extent that cannot be fully explained by stimulus saliency or redundancy. A possible explanation is that sensory noise is largely non-overlapping between different sensory organs whereas it is mostly shared within a specific sensory modality. Therefore, combining sensory information over multiple sensory domains results in a higher behavioral improvement compared to sensory integration of stimuli within a single sensory domain.

Furthermore, the chance of detecting a composite stimulus compared to its modality-specific constituents increases automatically without the need for integration (Miller, 1982, 1986; Colonius and Diederich, 2004; Diederich and Colonius, 2008), even without considering increases in overall saliency. This may occur simply as a consequence of the fact that stimuli from different modalities can be detected independently – yet simultaneously – by the corresponding sensory systems. We recently investigated whether the multisensory gain we observed in an audio-visual stimulus detection task with mice was indeed dependent on cue integration (Meijer et al., 2018) by assessing how well models of cue-combination based on signal detection theory fitted to our behavioral data (Jones, 2016) – Fig. 2a. We showed that models -involving cue integration according to – (non-) linear combinations of visual and auditory inputs fitted significantly better to the average behavioral performance of all mice as well as to the performance of the majority of individual mice than the model assuming independent processing of stimuli. For the other mice, all models fitted the behavioral data equally well. These results indicate that the multisensory gain was more likely to be achieved by the integration of stimuli than by stimulus redundancy or competition.

Another bottom-up, cue integration-based process is *stimulus discrimination* (Fig. 2b). Discriminating between two or more stimuli (or the specific properties of a stimulus, such as its direction of motion) has been shown to be more accurate when the configurations impinge on more than one sensory system (Ernst and Bülthoff, 2004; Lippert et al., 2007; Nikbakht et al., 2018). In a series of studies, the laboratory of Angelaki and DeAngelis showed that monkeys judge their heading direction more accurately based upon combined visual and vestibular cues as compared to only one of the two constituents (Fetsch et al., 2009, 2012; Dokka et al., 2015). Similarly, rats were shown to judge more reliably whether a train of short stimuli had a high or low rate when the stimulus train consisted of audio-visual stimuli rather than only visual or auditory stimuli (Raposo et al., 2012; Sheppard et al., 2013; Brunton et al., 2013) – Fig. 2b. In both cases, the increase in discriminability between multi- and unisensory stimuli closely approximated the prediction made by Bayesian optimal cue integration (Pouget et al., 2002; Knill and Pouget, 2004; Ma et al., 2006; Meyniel et al., 2015). The Bayesian framework provides a prediction as to how to optimally combine two sources of sensory information, which may differ in terms of reliability, in order to predict whether a stimulus feature is present (detection) or how it should be categorized (discrimination). Humans (Ernst and Banks, 2002), non-human primates (Fetsch et al., 2012), ferrets (Hollensteiner et al., 2015) and rats (Raposo et al., 2012; Sheppard et al., 2013) were all shown to integrate cues in a way that results in a near-optimal combination of cue reliability. Moreover, the Bayesian framework has also been proposed as a computational mechanism via which the brain can decide whether two stimuli should be segregated or integrated (see Box 1 for further details). Thus, the Bayesian framework has the potential to provide a mechanistic explanation of cortical MP, and not just multisensory cue integration. However, recent studies showed that rodents are also able to integrate sensory stimuli in a supra-linear, better than Bayes-optimal fashion (Nikbakht et al., 2018). This suggests that, besides mechanisms in line with Bayesian integration, the sampling of cross-modal stimuli is

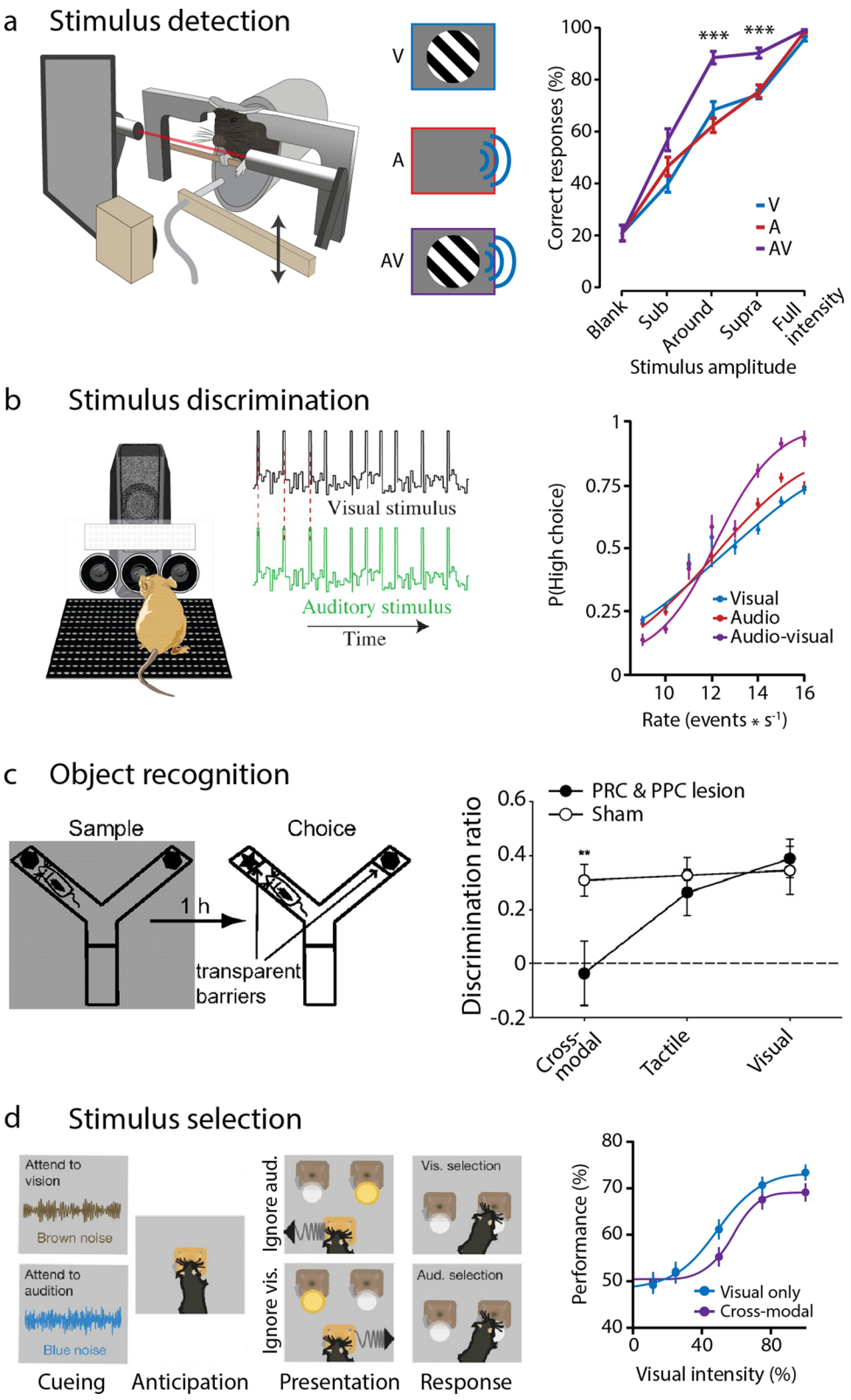

Fig. 2. Behavioral paradigms probing multisensory processes. a) *Left* Stimulus detection paradigm in which head-fixed mice were required to make a licking response in relation to visual moving gratings (V), auditory stimuli (A) of different contrasts and amplitudes, respectively, presented alone or in concert (AV). *Right* Mice more accurately responded to the audiovisual stimuli compared to visual or audio stimuli presented in isolation when the amplitudes of the stimuli were around or above the detection threshold, but not for subthreshold and full intensity stimuli. Blank: trials in which no stimulus was presented. Correct responses: the number of correct responses as percentage of the total number of trials performed. Adapted from (Meijer et al., 2018). *Left* In this stimulus discrimination paradigm, rats were trained to report whether a stimulus train, consisting of visual, audio or audiovisual components, was presented at a high or low rate by poking their nose in the associated port. *Right* The probability of 'high rate' choices increases more sharply for audio-visual stimulus trains (purple curve) than for unisensory trains (red and blue curves) with increasing stimulus rate. From (Raposo et al., 2012). c) *Left* In this cross-modal tactile-to-visual object recognition task, rats sampled objects in the arms of a completely dark Y-shaped track using tactile stimuli only. An hour later rats identified objects using visual stimuli only (transparent barriers are inserted to prevent access to the objects). *Right* Sham injected rats were able to recognize visual features of the object on the basis of a multisensory representation with an accuracy comparable to tactile-tactile object recognition. Excitotoxic disconnection lesions of the perirhinal cortex and posterior parietal cortex (PRC & PPC lesion) disrupted cross-modal -, but not tactile-tactile and visual-visual object recognition as indicated by a significant decrease in the ratio of time the rat spent exploring the novel object in these conditions (discrimination ratio; ** $p <$ 0.01). From (Winters and Reid, 2010). d) *Left* Stimulus selection paradigm in which mice were cued (Cueing) with an auditory stimulus to either ignore the visual or the auditory stimulus (Presentation) while making a nose poke response at the associated location (Response). *Right* The number of correct responses in this task increased with intensity of the visual stimulus and was better across stimulus intensities when visual stimuli were presented in isolation (Visual only; blue curve) than when visual and auditory stimuli were presented concurrently (Cross-modal; purple curve). From (Wimmer et al., 2015).

Box 2

Potential pitfalls while quantifying the principle of inverse effectiveness.

In their seminal paper, Meredith & Stein (1983) observed that the superior colliculus contains neurons which enhance their response when a visual stimulus was paired with an auditory cue. Importantly, they made the following additional observation: "When individual stimuli were just above threshold for eliciting discharges, the enhancement was greater than when the individual stimuli were highly effective. Thus, as the effectiveness (number of discharges elicited) of the individual stimuli increased, the percentage of enhancement produced by combining them decreased." ((Meredith and Stein, 1983), p. 391). This inverse relationship between stimulus drive and the magnitude of multisensory integration has been adopted as a - principle of multisensory integration and is commonly referred to as the 'principle of inverse effectiveness' (see also section 3).

The principle of inverse effectiveness is commonly investigated by either *(i)* presenting stimuli with different levels of intensity and quantifying the magnitude of multisensory integration for each stimulus intensity, or *(ii)* by sorting neurons according to their responsiveness to a stimulus and determining the relationship between neuronal responsiveness and multisensory enhancement. While the former approach is quite robust (but see below), the latter is subject to several statistical considerations which, if not accounted for with the proper control analyses, may result in the erroneous conclusion that neurons show inverse effectiveness while the observed correlation is actually spurious (Holmes, 2009, 2007), cf. (Meijer et al., 2017). Analyses using the latter approach often show a significant negative correlation between unisensory responsiveness and multisensory modulation, such that weakly responsive neurons enhance their firing whereas more strongly responding neurons do so to a lesser degree or suppress their firing (Perrault et al., 2003; Alvarado et al., 2007; Kayser et al., 2008). This manner of quantifying inverse effectiveness is liable to several considerations which will be discussed below.

First, computing the correlation between two variables which are not independent from one another might suffer from a 'regression to the mean' effect. In the example illustrated above, neurons are sorted according to their responsiveness in condition A (unisensory) and their percentage change in response from condition A to condition B (multisensory) is determined. However, it is reasonable to assume that neuronal populations show a normal distribution of their population response to a collection of stimuli (Ma et al., 2006). Therefore, even without any quantitative difference between condition A and B, neurons that have been selected to be on the low end of the normal distribution in condition A will 'regress to the mean' in condition B resulting in a seemingly higher response and vice versa.

Second, neurons have a dynamic range in which they can optimally change their firing rate bi-directionally. This is also commonly expressed in a sigmoid transfer function, plotting a neuron's firing-rate output as a function of its total synaptic input. For example, when the activity of a neuron is close to the upper limit of its dynamic range the probability of its activity to increase even further is small because of biophysical constraints of the generation of action potentials. Therefore, neurons which show either very low or very high firing rates should be treated with caution because they can predominantly modulate their firing rate only in a unidirectional fashion which will artificially inflate the correlation between unisensory responsiveness and multisensory response modulation. Basically, the addition of any input, whether modality-specific or cross-modal, is likely to show a nonlinear response enhancement if the initial response to a single input is low. Note that both approaches *(i)* and *(ii)* outlined above potentially suffer from this problem.

Finally, the metric that is used to quantify multisensory response modulation can have a profound effect on the observed results (Holmes, 2009). A common practice is to express the magnitude of multisensory modulation as a percentage increase of the neural response elicited by the multisensory stimulus as compared to the strongest elicited unisensory response (Stevenson et al., 2014). When quantifying multisensory modulation in this manner, one should take note that percentual differences are especially large when the quantities that are being compared are small. For example, a single extra spike will result in a large percentual increase when the original number of spikes was small compared to when it was large. However, downstream neurons often require an absolute minimum number of presynaptic excitatory inputs in order to reach spiking threshold. Therefore, a neuron that increases its spiking from one to two spikes might show a 100% increase in activity but this does not necessarily translate to a proportionate increase in driving downstream neurons to fire.

These considerations imply that, when investigating concepts such as inverse effectiveness in MP, interpretation of the obtained results should always be corroborated by appropriate control analyses, which can take several forms. First, to control for the possibility of regression to the mean, the observed correlation between unisensory responsiveness and multisensory response modulation should be tested against a null-distribution which is obtained by a shuffling procedure. Only when the observed correlation is significantly stronger compared to the correlational values in the null-distribution it can be claimed that inverse effectiveness is present in the data. Second, to control for the potential problem of neurons operating at the limits of their dynamic range, it should be assured that inverse effectiveness persists after removal of neurons with the lowest and highest firing rates. Nonlinear enhancement should be stronger than that obtained by unisensory manipulation. To address the potential influence of units of quantification, several ways of quantifying -multisensory modulation should be used to assure that the observed effect is robust against these different methods of expressing multisensory modulation. The considerations discussed above do not solely apply to the quantification of inverse effectiveness, but to any case in which the two variables that are compared are not independent from one another (e.g. baseline brain-area volume vs. change in volume after an intervention).

an active process in which information from one modality can influence the sampling of the other in a synergistic, non-linear manner. For example, rodents could more optimally scan an object with their whiskers based upon the visual input pertaining to this object. Finally, in the literature on human MP (Regenbogen et al., 2016) and on animal unisensory processing (Hanks et al., 2015) other frameworks, such as the family of drift diffusion models, are also used to quantify sensory behaviors. These models, however, have so far not been adopted in the context of MP in animal studies.

Whereas the format of cross-modal information commonly used in experimental paradigms often leads to a MP outcome compatible with cue integration, different stimulus configurations may lead to the opposite outcome: *modality segregation*. Iurilli et al. (Iurilli et al., 2012) showed that the conditioned fear-response of mice to a visual stimulus was significantly weakened when a high-volume sound (effectively a distractor) was concurrently presented. This indicates that the loud auditory noise overshadowed the visual stimulus, resulting in a reduced behavioral output in response to the visual cue. Thus, when one modality provides disproportionally more salient or reliable stimulus information than a simultaneously present different one, it may capture

available resources and exclude the second modality from prioritized processing, thereby segregating sensory modalities in a bottom-up manner. Such dominance of one sensory modality may also occur when sensory stimuli predicting conflicting outcomes are concurrently presented. In mice experiencing audio-visual conflicts, innate circuit mechanisms at the level of sensory and association cortices enabled an auditory cue to determine the behavioral response at the expense of the visual cue. This occurred even if this meant that no reward would be obtained (Song et al., 2017). Thus, based on bottom-up, feature-dependent mechanisms, cross-modal information may be integrated or segregated in detection and discrimination paradigms.

The second class of behavioral tasks includes situational demand-driven, top-down MP (Fig. 1) – see for example (Talsma et al., 2010; De Meo et al., 2015). The various forms of top-down MP require that cognitive resources, such as attention or memory, interact with the properties of sensory stimuli to determine the behavioral outcome. Rats, monkeys and humans were all shown to be able to identify the same object using cues from distinct sensory modalities separately, an operation called *cross-modal object recognition* – Fig. 2c (Molholm et al., 1991; Lehmann and Murray, 2005; Seitz et al., 2006; Shams and Seitz, 2008; Winters and Reid, 2010; Vasconcelos et al., 2011; Shams et al., 2011; Moran et al., 2013; Rossi-Pool et al., 2016; Matusz et al., 2017). In a basic cross-modal object recognition task (Winters and Reid, 2010), rats first learned object features using one sensory modality only (e.g. tactile features only). Then, they were asked to identify the familiar object using cues from a different modality (e.g. visual cues) only (Fig. 2c). Rats were shown to be able to transfer object information across sensory modalities when the testing phase followed the acquisition by up to one hour. The retention interval over which the rats were able to transfer tactile information to the visual domain was significantly prolonged, however, when rats were allowed to see and touch the objects before task execution (Jacklin et al., 2016). This prolongation of cross-modal object recognition is hypothesized to depend on the formation of a multisensory object representation harboring direct associations between tactile and visual information during pre-exposure. Thus, subjects are able to transfer object information between senses (either via abstract, amodal object representations or via intrinsically multisensory ones), thereby increasing their behavioral flexibility. In addition, they incorporate previous experiences enhancing the multisensory behavioral gain.

Top-down MP also includes *sensory selection,* in which cross-modal information is segregated using a form of attention that prioritizes processing of one sensory modality at the expense of others (Lakatos et al., 2008; Schroeder and Lakatos, 2009; Wimmer et al., 2015; Ahrens et al., 2015). In an elegant behavioral paradigm by Wimmer et al. (2015) (Fig. 2d) mice were trained on a two-alternative forced choice operant conditioning task in which a visual and/or auditory stimulus signaled reward to be available in one of two reward ports associated with each modality. Before stimulus onset, the mouse was cued on whether the visual stimulus had to be attended and the auditory tone had to be ignored (as the former was indicative of reward position while the latter was non-informative) or vice versa (Fig. 2d). When the visual and auditory stimuli were presented concurrently, and thus provided conflicting information on reward availability, the mouse was able to selectively attend to the cued modality and responded accordingly, by licking in the reward port where the cued modality was presented.

On the basis of the above overview of MP-related studies, it is not exaggerated to state that in recent years major advancements were made in developing cross-modal behavioral tasks for rodents, capable of testing a broad range of MP (cf. Figs. 1 and 2). Some of these tasks are highly similar to human behavioral tasks and therefore allow for translational research (Raposo et al., 2012). The availability of these validated behavioral tasks is an important stepping stone for gaining understanding of the neuronal mechanisms associated with each type of MP, such as bottom-up versus top-down MP. The cortical circuitry underlying the whole range of MP is currently only partially understood at best, although recent proposals have begun to address this issue (van Atteveldt et al., 2014). Going beyond cue integration offers the opportunity to better understand the neuron-level mechanisms underlying cortical MP, and uncover if and how the computational principles for MP in the cortex differ from those of subcortical structures such as the SC (Stein and Stanford, 2008).

3. Neuronal mechanisms of bottom-up multisensory cue integration

In the last decade, many behavioral tasks have been developed, especially for rodents, to systematically test how the outcome of a broad range of MP is determined by the interaction between bottom-up and top-down factors (Fig. 1). This makes it now possible to characterize the neuronal architecture associated with the various multisensory behaviors. In this section we will address neuronal mechanisms recorded in animals which were presented with multimodal stimuli but were not required to make an active behavioral response to obtain reward. We refer to this experimental set up as “passively observing”. These neuronal mechanisms, which mediate cue integration, are largely stimulus-driven and, at the level of the neocortex, are mediated by inter-areal connections. As such, they provide insight in the influence of stimulus properties (i.e. bottom-up influences) on MP outcome (Fig. 1).

The SC of the cat has been the traditional model system for studying cue integration at the level of single neurons. Unisensory neurons in this structure, which are sensitive to visual, auditory or somatosensory stimulation, converge on multisensory neurons that respond with a significantly different firing rate change to cross-modal stimuli than to any one of the modality specific stimulus components. This has often been used as an operational definition of *multisensory integration*, which can be either positive (enhancement) or negative (depression), based on whether multisensory responses are higher or lower than the highest or lowest unisensory response, respectively (Stein et al., 2009b). In case of *multisensory enhancement*, this can result in firing responses that are either larger (supra-additive) or smaller (sub-additive) than the sum of responses to modality-specific stimulus components (Meredith and Stein, 1983; Stein et al., 2009b). The magnitude of multisensory enhancement was shown to be inversely proportional to the ability of the modality-specific stimulus component to elicit a neuronal response, which is known as the principle of *inverse effectiveness* (Meredith and Stein, 1986) – but see Box 2 for a discussion on the care that needs to be taken to properly assess this principle. Multisensory enhancement was furthermore shown to be maximal when the unisensory stimulus constituents (visual, auditory or somatosensory) were presented from (roughly) the same location in space. This is known as the *spatial principle* of multisensory integration (Meredith and Stein, 1986; Kadunce et al., 2001). Because the receptive fields for the different sensory modalities are aligned in bimodal neurons, they are optimally activated under this stimulus configuration. Lastly, for multisensory enhancement to occur, stimulus constituents must be presented closely related in time – the *temporal principle* of multisensory integration (Meredith et al., 1987). Only near-simultaneous activations of the bimodal neurons by the stimulus constituents will produce neuronal responses that can interact. For elaborate reviews on the principles of multisensory integration we refer to (Stein and Stanford, 2008; Stein et al., 2009b). These three neuron-level principles along which multisensory cue integration operates were also shown to hold at the behavioral level, when cats were detecting cross-modal stimuli (Gingras et al., 2009) – but see (Meijer et al., 2018). Notably, multisensory integration is different from *uni*sensory integration, which is defined as the integration of multiple stimuli from a single modality – e.g. two co-occurring visual stimuli (Alvarado et al., 2007; Stein and Stanford, 2008). This occurs as a result of several differences in the way stimuli from the same or different modalities interact with each other – for example shared vs. uncorrelated noise sources, or specific patterns of

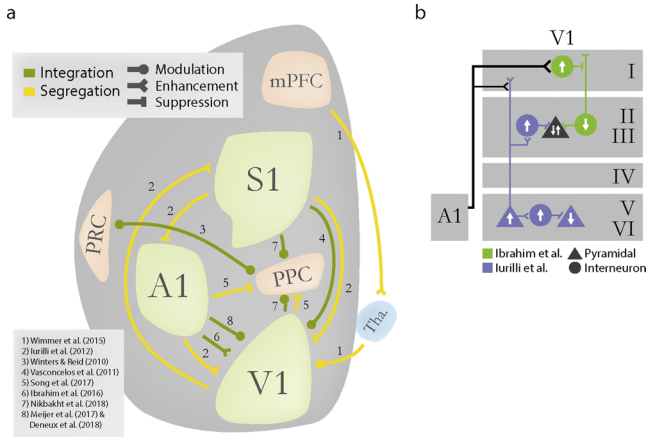

Fig. 3. Current understanding of multisensory processing in single neurons, microcircuits and neural systems. a) Multisensory processing occurs across many nodes along the corticothalamic hierarchy, including primary sensory areas (green), association cortices (pink) and thalamic nuclei (blue). The lines specify the known *physiological* interactions between these areas. Information contributed from different senses may be integrated (green lines) or segregated (yellow lines). At the single neuron level, cortico-cortical multisensory influences are manifested as enhancement or suppression of neuronal responses to multisensory compared to unisensory stimuli. If both response enhancements and suppressions have been observed between a pair of brain areas this is indicated as modulation. See section 3 and 4 for explanation of the specific connections between brain areas. V1: primary visual cortex; A1: primary auditory cortex; S1: primary somatosensory cortex; mPFC: medial prefrontal cortex; PRC: perirhinal cortex, PPC: posterior parietal cortex; Tha.: thalamic nuclei. b) V1 contains (at least) two micro-circuits via which sounds influence visual processing. Sound-induced hyperpolarization of layer II/III pyramidal neurons (triangles) was shown to be mediated by cortico-cortico projections impinging on V1 layer V pyramidal neurons which then recruit inhibitory neurons (purple) (Iurilli et al., 2012). Sound induced response enhancement in layer II/III pyramidal neurons was shown to be mediated by a suppression of interneurons (circles) in the primary visual cortex mediated by direct cortico-cortical input from A1 to layer 1 of V1 (green) (Ibrahim et al., 2016). Upwards arrows indicate increase in activity, downwards arrows decrease in activity. Both pathways were verified in (Deneux et al., 2018) and may be used under different contextual conditions (e.g. light and dark environment).

responsiveness to stimuli (e.g. center-surround effects). As a consequence of these factors, and in stark contrast with multisensory integration, unisensory integration often results in a weaker neuronal response compared with that elicited by single stimuli (Alvarado et al., 2007; Foxworthy et al., 2013a). This indicates that different circuits evolved for multi- versus unisensory processing, and that multisensory integration cannot just be seen as a simple consequence of additional inputs converging onto multisensory neurons.

3.1. Association cortices

The principles of multisensory cue integration established in the SC inspired the search for cortical neural correlates of multisensory processing. Historically speaking, much attention has been granted to the association cortices, which in the context of this review refers to cortical areas that are neither primary sensory nor motor, and receive converging inputs from sensory areas. Many single unit studies have highlighted a striking similarity between the principles underlying single-neuron-level multisensory integration in the SC and in association cortices of monkeys and cats, such as in the superior temporal sulcus (Bruce et al., 1981; Beauchamp, 2005; Ghazanfar and Schroeder, 2006; Cappe et al., 2009), the intraparietal sulcus (areas LIP and VIP) (Andersen et al., 1997; Bremmer et al., 2002; Gifford and Cohen, 2004; Avillac et al., 2005; Ghazanfar and Schroeder, 2006), the medial superior temporal lobe (i.e. area MSTd) (Gu et al., 2008, 2012) and frontal areas such as the cat antero-ectosylvian sulcus (Jiang et al., 2001, 2002; Ghazanfar and Schroeder, 2006; Stein and Stanford, 2008). In addition, studies investigating multisensory convergence and cue integration in rodents and ferrets (Wallace et al., 2004; Hirokawa et al., 2008; Foxworthy et al., 2013a; 2013b; Lippert et al., 2013; Olcese et al., 2013) have indicated that association cortices, mostly parietal but also temporal regions, contain a combination of unisensory and multisensory neurons and show multisensory enhancement predominantly in superficial layers (Foxworthy et al., 2013b; Olcese et al., 2013) – Fig. 3. In the cortex, as in the SC, multisensory enhancement and suppression responses can be supra- or sub-additive in relation to the sum of the two unisensory responses, depending on the specific stimulus settings. This has been shown to be compatible with the principle of inverse effectiveness (Olcese et al., 2013; Fetsch et al., 2013). Neuronal responses to bimodal visual-vestibular cues indicating heading direction were shown to be mostly sub-additive and could often be explained by a linear weighted summation of the unisensory responses (Morgan et al., 2008). The weights assigned to each sensory modality were not fixed across stimulus conditions, but varied with the relative reliability of visual and vestibular cues, which is in agreement with Bayesian cue integration. Multisensory processing in rat parietal cortex also complies with the SC temporal principle: current source density responses in layer 4 were enhanced when somatosensory stimuli preceded visual stimuli closely in time (0–100 ms), but suppressed when the stimuli were presented in opposite order (Lippert et al., 2013). Recently, the rat parietal cortex was also proposed to implement a form of “supramodal” integration (Nikbakht et al., 2018), in which single neurons encode the category of an object (e.g. the orientation of a grid) irrespective of the sensory modality which is used to display the object (visual, tactical or visuotactile). This suggests that cortical association areas are able to perform more advanced forms of multisensory processing than cue integration, namely object categorization. Strikingly, studies performed in non-human primates reported a similar supramodal form of integration in premotor and frontal areas (Rossi-Pool et al., 2016; Vergara et al., 2016). An open question is thus at which stage in cortical processing information from different sensory modalities is integrated to construct a supramodal representation.

The finding that cue weights dynamically adapt to the current situation (Morgan et al., 2008) formed the basis of a population-level interpretation of multisensory cue integration in association cortices. While previous studies mostly did not go beyond the single-neuron level (Stein and Stanford, 2008), a convergence of theoretical and experimental work led some groups to conceive a population-level, Bayesian framework to interpret how different sensory cues are effectively integrated in association cortices (Pouget et al., 2002; Fetsch et al., 2013) – Box 1. Such population-level computations are fully compatible with the single-neuron-level principles of cue integration, and have been hypothesized to be derivable from a circuit model implementing divisive normalization: an inhibition-mediated neuronal operation for combining different signal sources at the microcircuit level (Ohshiro et al., 2011; Carandini and Heeger, 2012). Divisive normalization operates by having a pool of surround neurons performing a normalization of the integrated outputs of a given neuron (Carandini et al., 1997; Carandini and Heeger, 2012; Fetsch et al., 2013; Ohshiro et al., 2017). An important feature of the model is that the normalization step is implemented as a division, allowing output signals to be scaled. By contrast, a subtraction procedure creates output signals that are not necessarily related to the existing signal in a linear way. Nevertheless, it remains to be seen how general and valid this normalization operation for explaining multisensory integration is. For example, it is not known whether it applies to cortical areas beyond association cortices (or to the SC (Ursino et al., 2014)), or whether it can be generalized across all types of stimuli (e.g. highly salient stimuli (Iurilli et al., 2012)) and all forms of MP.

Finally, there is evidence that neurons in the association cortices, in addition to stimulus properties such as amplitude, spatial and temporal

coherence – which lie at the basis of the -principles of multisensory cue integration in the SC – also encode congruency and complexity of stimulus constituents. Ghazanfar et al. (2005) showed that local field potential (LFP) amplitude responses in the auditory belt showed enhancement primarily when pictures of vocalizing monkeys were paired with the correct vocalizations. It is thus likely that cortical areas can perform more complex – and acquired (non-innate) – forms of cue integration, and that the applicability of SC-derived principles of cue integration will have to be reconsidered on the basis of higher-order sensory features such as category (e.g. face vs. object) or motivational value. While multisensory cue integration in the SC has been linked to stimulus detection and localization (Rowland et al., 2007; Stein and Stanford, 2008; Gingras et al., 2009), the exact functional role of cortical multisensory cue integration has not been clearly identified, as most association cortices integrating multiple modalities have only been shown to be causally affecting behavior for one sensory modality alone (Raposo et al., 2014; Song et al., 2017; Licata et al., 2017).

3.2. Primary sensory cortices

Primary sensory cortices were recently shown to contain neurons differentially responsive to modality-specific and cross-modal cues (Kadunce et al., 2001; Bizley et al., 2007; Kayser et al., 2008; Ghazanfar and Lemus, 2010; Kayser and Logothetis, 2007; Lemus et al., 2010; Iurilli et al., 2012; Ibrahim et al., 2016;) – Fig. 3. This is in line with connectivity studies in primates (Falchier et al., 2002; Cappe et al., 2009), which indicated direct projections from A1 to V1, specifically towards the portion of V1 corresponding to the peripheral visual field. Hence, it was hypothesized that such connections might aid the detection of peripheral stimuli, in line with the proposed role for multisensory integration in the SC. Moreover, functional studies also indicated primary areas as potential loci of multisensory integration (Calvert et al., 1997; Murray et al., 2016a). However, the changes in response patterns of single neurons to stimuli of a non-primary modality were shown to be highly heterogeneous, and ultimately not in line with the above mentioned hypothesis on the function of MP in primary cortices pertaining to the detection of peripheral stimuli.

In fact, neuronal correlates of MP in primary cortices indicate different functional applications. As a first case, primary cortices have been shown to integrate sensory inputs from other modalities in a way that increases their discriminatory power, i.e. stimuli from the non-primary modality enable neurons in one sensory area to better discriminate similar stimuli from its primary sensory modality (Lippert et al., 2007; Bizley and King, 2008; Kayser et al., 2010; Feng et al., 2014; Ibrahim et al., 2016; Meijer et al., 2017). This is firstly reflected by neurons showing either enhanced or suppressed firing rates in relation to bimodal compared to modality-specific stimuli, resulting in sub-additive, additive or supra-additive responses (Wallace et al., 2004; Bizley et al., 2007; Kayser et al., 2008; Kayser et al., 2010; Vasconcelos et al., 2011; Ibrahim et al., 2016; Meijer et al., 2017). In a recent study using two-photon calcium imaging in passively observing awake mice presented with cross-modal stimulus sets, we further characterized this by showing that auditory modulation of visual processing was strongly dependent on the contrast of the visual stimulus and the temporal congruency of cross-modal stimuli (Meijer et al., 2017). Presentation of a visual moving grating in concert with a tone that was modulated at the same rate gave rise to a response enhancement in a sub-population of neurons, and an equally large subpopulation showing response suppression compared with unisensory visual stimuli. Stimulus pairs which were temporally incongruent or did not share temporal patterning (i.e. white noise bursts) at full contrast resulted in a predominant response suppression across the neuronal population. Visual contrast did not influence multisensory processing when the audio-visual stimulus pairs were temporally congruent. However, the response suppression that was observed when pairing a full contrast visual stimulus with a white noise burst was absent when the contrast of the visual stimulus was low. A second indication of multisensory integration, leading to increased discriminative power, was apparent from more reliable fine-grained firing patterns and reduced trial-to-trial variability to bimodal compared to unimodal stimuli. This was specifically found for neurons that suppressed their firing rates in the bimodal condition (Bizley et al., 2007; Kayser et al., 2010) compared to the unimodal conditions. Lastly, a small portion of the neurons in sensory cortices was found to be responsive to stimuli of the non-primary modality, even without a stimulus in the primary modality (Bizley et al., 2007; Meijer et al., 2017).

As a second case of MP in primary sensory areas, studies have produced indications for competition between different primary cortices (Kayser et al., 2008; Iurilli et al., 2012) – Fig. 3. Specifically, high-volume auditory bursts have been shown to hyperpolarize the membranes of neurons in superficial layers of the primary visual cortex of mice, via a direct inhibitory circuit from primary auditory to visual cortex, thereby disrupting visual processing (Iurilli et al., 2012). Similar inhibitory influences, specifically activated by highly salient sensory stimuli, were found to exist between most – but not all – primary cortices. For example, while at least some lateral connections from both primary auditory and somatosensory cortex impinging on the primary visual cortex are inhibitory, projections from the primary visual cortex onto the primary somatosensory cortex were shown to be excitatory (Iurilli et al., 2012) – Fig. 2. This suggests that the outcome of the competition between sensory modalities is highly dependent on the characteristics of the cortical areas under scrutiny and on the connectivity patterns between cortical regions (Fig. 3).

Furthermore, even for primary sensory cortices which were shown to compete with each other – e.g. sound-induced hyperpolarizations mediated by auditory neurons in V1 (Iurilli et al., 2012) – different forms of multisensory interaction may occur. Specifically, (Ibrahim et al. (2016)) recently showed that sounds were able to enhance neuronal responses in V1 to stimuli of the preferred orientation, in apparent contrast with the results reported by Iurilli et al. (Iurilli et al., 2012). How is it possible that sound-induced activity A1 has been shown to both inhibit and enhance V1 activity? Several key differences between the two studies may explain the opposing results, such as type of auditory stimuli (high-volume continuous white noise in (Iurilli et al., 2012) vs. a rhythmic pattern in (Ibrahim et al., 2016), brain state (shallow (Iurilli et al., 2012) vs. deep (Ibrahim et al. (2016)) anesthesia), absence (Iurilli et al., 2012) vs. presence (Ibrahim et al. (2016)) of simultaneous visual stimuli. Importantly, broadband white noise was used as an auditory stimulus in both studies; this rules out the possibility that a different affective tone of the auditory stimuli used in the two studies (Knutson et al., 2002; Portfors, 2007) might play a role in determining their different results. However, as pointed out above, stimulus configuration and temporal congruency of the constituents were shown to determine the response patterns to simultaneous audio-visual stimulation in awake mice (Meijer et al., 2017). Another recent study showed that A1 neurons projecting to V1 in awake mice preferentially respond to abrupt sounds, yet have a differential effect on V1 activity based on visual context: inhibitory in darkness and excitatory during illumination (Deneux et al., 2018). Thus, the interactions between A1 and V1 are highly dynamic, which may also explain the apparent contrast between the studies by Iurilli et al. and Ibrahim et al.. Moreover, distinct circuits may mediate sound-induced excitatory and inhibitory influences on V1 (Fig. 3b). Excitatory influences were mediated by projections from layer 5 in A1 to layer 1 interneurons in V1; these interneurons would then further inhibit layer 2–3 interneurons impinging on pyramidal cells, which would thus be disinhibited, resulting in enhanced visual responses (Ibrahim et al. (2016)) – Fig. 3b. Concurrent visual stimuli may facilitate this depolarizing effect and turn it into an hyperpolarizing mechanism during darkness (Deneux et al., 2018). The hyperpolarizing effect described in (Iurilli et al., 2012) was instead elicited by A1 layer 5 projections onto layer 5 pyramidal neurons in V1; these were shown to enhance the activity of

V1 interneurons, which would then mediate the reported sound-induced hyperpolarization (Fig. 3b). It is therefore plausible that different anatomical circuits between A1 and V1 coexist, and that the net effect of auditory influences on V1 is due to the combination of a plethora of factors, all of which determine how these distinct circuits interact. Thus, the same anatomical circuit in the primary sensory cortex (e.g. direct connections between V1 and A1) can harbor distinct connection patterns to exert different effects based for instance on the type (volume, frequency, temporal pattern, cross-modal synchrony) of sensory stimuli, and thus implement distinct forms of context-dependent cross-modal behaviors.

In addition to single neuron mechanisms, multisensory interactions were shown to be implemented at the local circuit level in the primary sensory cortices by oscillatory phase resetting (Lakatos et al., 2008; Schroeder and Lakatos, 2009). In passively observing monkeys, a salient (punctate) somatosensory stimulus was shown to reset ongoing rhythmic activity in several frequency bands (especially in delta, theta and gamma bands) in the primary auditory cortex. In this way, the stimulus instantiated a rhythmic fluctuation of the excitability of the auditory cortical circuit (mediated via neurons in supragranular layers), by which responses to auditory stimuli occurring at the peaks of the oscillation were enhanced and responses occurring at the troughs were depressed. This phase-resetting mechanism has been hypothesized to enhance signal transmission across cortical areas (Schroeder and Lakatos, 2009; van Atteveldt et al., 2014; Fries, 2015), and could thus lie at the basis of the principles of multisensory cue integration.

In conclusion, what becomes clear from this overview of bottom-up multisensory processing in primary and association cortices is that the cortex not only has mechanisms reminiscent of the original SC principles, but also several neuronal strategies to implement bottom-up MP, both at the single neuron and population level. Each of these strategies is highly dependent on specific stimulus features, as well as on properties of individual neurons and cortical regions involved. Whether specific mechanisms pertain to association cortex, sensory cortex, or represent more general coding schemes awaits further investigation.

4. Cognitive and environmental influences on the neuronal mechanisms of multisensory processing

In the previous section we outlined how processing of signals originating from different sensory systems is spread beyond multisensory-specific areas and how individual brain regions have many neuronal mechanisms at their disposal to implement bottom-up MP. Besides low-level (bottom-up) factors, the outcome of multisensory processing is, however, often also dependent on high-level (top-down) factors (De Meo et al., 2015). Here, we examine recent developments in the research of neuronal mechanisms of MP in actively behaving animals, which are determined by the interaction between stimulus features and high-level factors such as behavioral demands and cognitive resources (Fig. 1).

Rodent research on task-driven MP focuses mainly on the (posterior) parietal cortex (PPC; Fig. 3), which consists of a collection of distinct but related areas (Wang and Burkhalter, 2013), each receiving visual, auditory and/or somatosensory inputs forwarded from primary sensory cortices and the thalamus (Olcese et al., 2013; Wang and Burkhalter, 2013). Moreover, the PPC is involved in an array of behaviors including perceptual decision making, motor planning and object recognition (Nitz, 2006; Harvey et al., 2012; Tafazoli et al., 2017). The PPC was recently demonstrated to be causally involved in a *stimulus detection task* in which mice faced audio-visual conflicts (Song et al., 2017). Animals were trained to lick upon presentation of an auditory cue, but refrain in case of a visual cue (or vice versa), by administering – respectively – a reward or a punishment. The authors found that the auditory cue determined the behavioral response at the expense of the visual cue, when both were presented simultaneously. Optogenetic blocking of the auditory-driven feedforward inhibition from sensory cortices, mediated by PPC parvalbumin-positive (PV) interneurons, was able to lift the auditory dominance in cue conflict choices. Consistent with this finding is the fact that PV interneurons in area RL (which is often considered to be part of PPC (Olcese et al., 2013; Mohan et al., 2017)) were shown to mediate visuo-tactile integration in layer 2–3 pyramidal neurons (Olcese et al., 2013). In contrast, the PPC did not appear to be causally involved in *discriminating* between high and low rate audio-visual stimulus trains, whereas it did show such involvement in discriminating the visual component of these stimuli (Raposo et al., 2012; Licata et al., 2017;). Neuronal responses to visual stimuli in PPC were larger than to auditory stimuli (Licata et al., 2017) and multisensory responses were usually well predicted by a linear combination of auditory and visual responses (Raposo et al., 2014), showing no sign of multisensory interactions. Cross-modal *object recognition,* in which rats use e.g. a tactile object representation to identify a familiar object using visual cues only (see Section 2), was disrupted when PPC function or its communication with the perirhinal cortex was compromised (Winters and Reid, 2010) – Fig. 2c. However, when rats were allowed to form a multisensory object representation before testing, by exploring the objects visually and haptically, object recognition appeared to be solely dependent on the perirhinal cortex whereas PPC was not involved (Jacklin et al., 2016). This indicates that experience may alter the neural substrates and mechanisms associated with multisensory behaviors, i.e. switching from a feedforward-mediated approach in which sensory information is being transferred from PPC to perirhinal cortex to a feedback-mediated strategy in which multisensory representations stored in perirhinal cortex affect sensory representations in earlier cortices. Although the PPC is a heterogeneous structure in terms of efferents, afferents and behavioral involvement, what becomes clear is that it serves multisensory behaviors in differential ways depending on the underlying cognitive process and factors such as experience-dependent plasticity.

As yet, the most direct link between multisensory neuronal correlates and behavior has been shown in the dorsal medial superior temporal area (MSTd) of the non-human primate, which is a node in the network facilitating multisensory heading perception. The population of neurons in this area was shown to code for the relative reliability of visual and vestibular cues and to be able to adjust its representation on the time scale of a few seconds (i.e. on a trial-by-trial basis) in conjunction with *discrimination* behavior (Fetsch et al., 2009, 2012). Although these studies elegantly show the dependence of neural coding on the specific cue combination used, a causal link between MSTd neuronal coding and discrimination behaviors has not been firmly established (Gu et al., 1991, 2012).

Sensory selection (Lakatos et al., 2008; Schroeder and Lakatos, 2009; Wimmer et al., 2015; Ahrens et al., 2015), a form of attention, has long been hypothesized to be dependent on top-down modulation originating in the medial prefrontal cortex and impinging on primary sensory and association cortices (Birrell and Brown, 2000). Recent studies have shown a causal role of the reticular thalamic formation in mediating sensory selection (Wimmer et al., 2015; Ahrens et al., 2015). The reticular thalamic formation acts as a relay station which – based on the inputs received from the medial prefrontal cortex – is able to selectively control thalamic sensory gain and therefore gate the level of stimulation that reaches primary cortices (Wimmer et al., 2015) – Figs. 2d and 3 . This gating operates by suppressing the activity evoked by the non-attended modality at the level of sensory thalamus and primary cortices (Wimmer et al., 2015; Lee et al., 2016), but a mechanistic understanding of the circuit for the different sensory modalities, as well as its impact on association cortices, is still missing. Importantly, top-down cortico-cortical influences from frontal to sensory cortices have been shown to directly enhance processing of the attended modality (Zhang et al., 2014). Therefore, a better understanding of how bottom-up (multi)sensory processing and top-down attention interact is key to develop detailed models of how sensory selection, and consequently MP, operates.

Box 3
Multisensory processes as a window on consciousness.

Although the study of consciousness goes largely beyond the scope of this paper, it is worthwhile to summarize briefly why MP is important to consider when trying to understand how the brain generates conscious experience. We take conscious experience to include perception, dreaming and imagery; it is a state in which we experience the world and our body in a unitary, immersive and qualitatively rich manner. As such, conscious states contrast with unconscious states such as dreamless sleep, anaesthesia or coma. Following Jackendoff (Jackendoff (1987), consciousness is essentially defined by its sensory or sentient nature. Multisensory processing (or sensory processing in general) is not sufficient *per se* for consciousness to arise, because it can also occur non-consciously (Faivre et al., 2014; Mudrik et al., 2014; Tononi et al., 2016). Nonetheless, particular forms of MP are necessary for consciousness (Pennartz, 2009, 2015). This is related to the qualitative richness and complexity of the elements to be integrated: consciousness provides the subject with a multimodal survey of the situation we are in – both the situation in our surrounding world and our body within it (Pennartz, 2018) – whereas low-level forms of integration (e.g. MP in the superior colliculus) represent a form of signal conditioning (Mudrik et al., 2014). This does not mean that a subject's situational survey must represent all modalities at any time, but that each modality can be represented in this survey and that each modality is experientially distinct from others. A multisensory approach to perceptual inference is uniquely positioned to explain such qualitative richness, as it allows to investigate, at the same time, *i)* how a sensory scene composed of different modalities is experienced as unitary yet qualitatively differentiated, and *ii)* why we experience visual inputs as something different than, for instance, auditory or olfactory inputs [the modality identification problem (Pennartz, 2009)]. The conscious, multimodal survey of the situation that impinges on MP requires our brains to generate a quick-and-dirty "best guess" model (Gregory, 1980; von Helmholtz, 1867; Mumford, 1992; Crick and Koch, 1995) of our world-body situation as specified by the subject's collection of qualitatively different senses. For instance, the generation of a model of what we see – which is continuously updated as novel sensory inputs reach the brain – has been proposed to depend on a process of perceptual inference, relying on a hierarchy of visual cortical areas (Gregory, 1980; Marcel, 1983; Mumford, 1992; Friston, 2010) and this concept can also be applied to other sensory modalities as well as multisensory percepts (Olcese et al., 2013; Pennartz, 2015). The theoretical proposal in (Pennartz, 2009, 2015) holds that the brain contains a neural machinery consisting of distributed cortical areas which collectively sustain an interactive multimodal topology, by which it infers not only which objects and properties are at play in distinct individual senses, but also what the best-guess construction of one's multimodal situation as a whole is. This proposal thus amounts to characterizing conscious experience as a higher-order form of multisensory integration which supervenes on lower-level forms such as multisensory cue integration, object recognition and sensory selection (which are by themselves insufficient to generate conscious experience). In other words, consciousness is proposed to depend on a higher-level process of multisensory integration that results in the multimodal situational survey characteristic of conscious experience, whereas lower-level forms can occur without consciousness. Current studies are helping to pave the way for further investigations of the role of MP in conscious processing.

Addressing MP in a behavioral context allows revealing neuronal mechanisms associated with MP that are not a simple function of stimulus-driven processing, such as object recognition and sensory modality selection. In addition, even for cue integration-based stimulus detection and discrimination paradigms, including behavioral context expands the understanding of the underlying neural mechanisms. Taking the presented neuronal MP evidence together, we may conclude that various multisensory functions heavily depend on a common anatomical network that is situated especially in the sensory and parietal cortices but also involves higher-order cortical areas, which mediate specific functions. At the same time, many different micro-circuits co-exist both within and between brain areas belonging to this common anatomical network, each mediating a specific type of multisensory operation (Fig. 3). The picture of available micro-circuitries starts emerging from recent research but is far from being unambiguous and complete. A closely related experimental challenge is to link specific behaviors to supporting micro-circuitries and to unravel which factors determine the selection of the neuronal substrate to be used.

5. A broader picture of multisensory processing

Which factors determine whether cross-modal stimuli will be integrated or segregated? How is this operation performed? And which neuronal substrates will be called upon? Until now, most studies addressing these questions focused on the behavioral and neural correlates of multisensory cue integration. This is, however, only one form – albeit a major one – of a class of sensory computations that can be broadly defined as multisensory processing. Our main goal with this review is to expand the framework on MP, and its underlying neuron-level substrates, including any operation in which information across sensory modalities is combined to make a perceptual inference. We pursued this objective by investigating the existing body of results, gathered primarily in rodents, on how multisensory contributions to cognitive functions expressed in the behavior of an individual can be explained by basic neuronal, local-circuit and population mechanisms (cf. Figs. 1 and 3).

The central V-shaped module in the scheme that we have outlined (Fig. 1) represents the key outcome of multisensory processing: information conveyed by the different sensory organs can be integrated and treated as a unified representation (*integration*) or separated into two or more distinct representations (*segregation*). These representations are feeding into higher functions such as decision making and memory formation. Multisensory processes such as stimulus detection, discrimination, recognition and selection, leading to inferences based on sensory information, all operate within this central module. Depending on the ongoing cognitive processes, behaviors and situational constraints (top-down factors), and on the specific stimulus features (bottom-up influences), the behavioral benefit that individuals experience using multi- versus unisensory cues may be accounted for by correct integration (i.e. detection, discrimination and recognition of a multisensory object) or segregation (i.e. selection, conflicts) of the cues. Understanding how the brain can dynamically implement all of these different functions is thus imperative to understand the neuronal substrates of MP.

Achieving such a comprehensive understanding of the neuronal substrates of MP implies an elaborate and challenging experimental effort. Two paradigm shifts in the study of MP have recently accelerated the gain of its understanding. First, whereas MP has been mostly studied in anesthetized or passively observing awake animals which were presented with cross-modal stimulus sets, in recent years the number of behavioral paradigms probing MP has rapidly increased (see Section 2, Fig. 1). This effort created the circumstances needed to assess the top-down influences of behavioral constraints and cognitive processes on MP, especially when the behavioral paradigms are designed in a way

that is compatible with neurophysiological recording techniques. Such paradigms should include high trial counts for reliable tracking of behavioral and neuronal read outs, and a system by which the amplitudes of all stimulus constituents can be adjusted relative to their detection thresholds for each individual subject to standardize testing (Meijer et al., 2018). Second, whereas the experimental focus has traditionally been on the single-neuron computations underlying multisensory cue integration (Meredith and Stein, 1983; Stein and Stanford, 2008), in the last decade this was expanded with approaches able to investigate micro-circuitry and population activity codes (Fig. 3) (Lakatos et al., 2008; Iurilli et al., 2012; Olcese et al., 2013; Wimmer et al., 2015; Ibrahim et al., 2016; Meijer et al., 2017; Deneux et al., 2018; Nikbakht et al., 2018). This provides the opportunity, among others, to assess the interactions not only within, but also between brain areas.

New experimental strategies investigating MP should take into account several factors which have been extensively investigated in the context of unisensory processing, and at the mesoscale in humans, but comprise an important new element in the framework of circuit-level MP. First, the choice of cortical area and animal species is a key factor which will influence the results that can be obtained. As an example, recordings performed in the PPC of mice and rats have yielded contrasting results as regards the use of multisensory information in decision making (cf. (Raposo et al., 2014; Song et al., 2017)). The dominance of auditory over visual stimuli reported in mice (Song et al., 2017) could be ascribed to the fact that rodents rely more on audition than on vision (cf. (Campi and Krubitzer, 2010)). Alternatively, the lack of a behavioral role for auditory signals in PPC reported in rats by (Raposo et al., 2014; Licata et al., 2017) might be explained by a recording position more proximal to the visual cortex when compared to (Song et al., 2017), or possibly also to species-specific differences. It is thus crucial not only to select the most appropriate animal model, but also to take into account subtle yet crucial differences in the anatomical location that is investigated.

Second, cross-modal stimuli used to study MP in passively observing or behaving subjects have been usually very coarse and simple, such as light flashes, visual gratings or broadband white noise bursts (but see (Kayser et al., 2010; Hwang and Romanski, 2015) in non-human primates). However, the level of stimulus complexity, as well as the MP requirements of the task, together define the cortical areas and neuronal mechanisms involved in the processing of multisensory cues. Mice, for example, are able to discriminate between both simple, first order and more complex visual cues. Neurons in V1 and the extrastriate region LM were less activated by complex (i.e., second order) compared to simple stimuli (Khastkhodaei et al., 2016). Furthermore, in monkeys the perirhinal cortex is causally involved in the processing of complex but not of simple stimuli (Bussey et al., 2002). Thus, using very simple cues in cross-modal paradigms may lead to a biased view on the involved brain areas and neuronal mechanisms, because involvement of the primary sensory cortices may be over-represented whereas potential involvement of association cortices may remain unnoticed. Indeed, human studies suggest that simple features (e.g. orientation) are already integrated in classical unisensory areas (van Kemenade et al., 2014), while more complex ones (e.g. limb position) are processed in a supramodal manner in association areas (Limanowski and Blankenburg, 2016). Finally, whereas stimulus properties influence multisensory processing via bottom-up mechanisms, task demands, behavioral contexts and cognitive resources such as attention and memory involve top-down mechanisms and require experience with the stimuli in order to be engaged (Fig. 1). These top-down mechanisms influence the way in which even primary sensory cortices operate (Convento et al., 2018). The first few circuit-level studies directly assessing prioritizing or selecting sensory modalities at the expense of others convincingly showed that these processes are mediated by neuronal networks not considered in multisensory research before (including the medial prefrontal cortex and the reticular nucleus of the thalamus (Wimmer et al., 2015; Ahrens et al., 2015)). For instance, a cognitive resource that has not received any attention to date in the context of multisensory processing is reward expectancy. Although neural correlates of reward and reward prediction are observed in the primary sensory cortices (Shuler and Bear, 2006; Goltstein et al., 2013; Poort et al., 2015; Goltstein et al., 2018) it is unclear how (expectation of) reward influences multisensory behavior and the supporting neuronal mechanisms. The broader picture of MP, sketched above, opens up new avenues for research on the role of MP in higher cognitive processes, including what is arguably the most complex function of the brain: consciousness. Box 3, Figure Box includes a brief introduction and discussion on the link between MP and consciousness.

It could be argued that a broadening of the scope of MP and the considerations described above are complicating an already multi-faceted and intricate phenomenon and its investigation. We would like to stress, however, that generalization of these matters may render important MP aspects or processes invisible. In a recent review, for example, van Atteveld et al. (2014) proposed that two major population-level mechanisms (divisive normalization and phase resetting) might explain most types of cortical MP. However, while some forms of MP (e.g. cue integration) might be fully explained by the proposed two mechanisms, we believe that a complex, diverse phenomenon such as MP can only be explained by considering a larger set of computational rules. For example, the complex cortico-thalamic-cortical loop for sensory selection outlined in Wimmer et al. (2015) has to interact with sensory factors (e.g. relative salience, timing), other forms of top-down modulation (e.g. selective attention (Zhang et al., 2014)) and additional factors such as arousal and locomotion (McGinley et al., 2015). These interactions occur along different stages of sensory processing (e.g. early sensory cortices vs. association regions). While general operations might be present along all these stages and contexts, the cellular implementations are likely to be highly flexible and diverse. In addition, at the level of local neural circuits, divisive normalization has been hypothesized as a possible mechanism able to explain both Bayesian cue integration and SC-derived principles of multisensory integration (Ohshiro et al., 2011). Importantly, this study showed how the two frameworks are totally compatible. However, studies showing differential effects of auditory influences on primary visual cortex (Iurilli et al., 2012; Ibrahim et al., 2016) challenge the general applicability of this scheme. Indeed, while sound-induced hyperpolarizations are mediated by a circuit involving deep cortical layers (Iurilli et al., 2012), a network limited to superficial layers implements audiovisual facilitation (Fig. 3b). Whether such specific circuit implementations are compatible with a general mechanism for MP, and can be modeled in a hierarchical Bayesian fashion (Box 1) remains to be addressed.

How do rodent – and, more generally animal – studies of MP relate to mechanisms in the human brain? The need to combine information from different sensory modalities is an essential aspect of our lives, and has been extensively investigated at the psychophysical level and in terms of mesoscopic neural activity. Yet, even the most low-level multisensory influences on cognitive processes – for example the sound-induced flash illusion (Shams et al., 2000) – lack a clearly described neuron-level mechanism at present. The range of techniques for circuit-level investigation available for rodent studies, combined with the advances that have been made in developing behavioral tasks for mice and rats (Carandini and Churchland, 2013), now open up the possibility to address this question. Importantly, unveiling the mechanism of MP is not purely of academic interest. Impaired sensory and multi-sensory processing may lie at the basis of various neuropsychiatric disorders, including autism (Marco et al., 2011; Baum et al., 2015) and schizophrenia (de Gelder et al., 2005; Ross et al., 2007; Javitt, 2009). A better understanding of how sensory modalities are processed in healthy brains is a first step to better understand – and eventually address – these disorders.

In conclusion, we have sketched the contours of an emerging framework to study MP. The novel contribution of this scheme is jointly determined by *i)* its broader scope compared with cue integration

(Fig. 1), and *ii)* the availability of new and highly suitable experimental models to investigate different forms of MP. Specifically, a rapid gain in our understanding of the neuron- and systems-level mechanisms underlying multisensory processing is now achievable. This is the result of recent efforts in designing behavioral tasks for rodents, the availability of techniques to probe and intervene with neuronal activity at the population and micro-circuit level, the inclusion of top-down influences in MP, and the emergence of novel links between experimental work and theoretical frameworks (e.g. Box 3). The scheme that we present will make it possible to address several key open questions about MP, including: which single unit and population neuronal processes, and which brain structures are causally involved in multisensory behaviors? Which bottom-up and top-down factors determine the micro-circuitry that is activated for specific MP processes?

Acknowledgments

This study was supported by the FLAG-ERA JTC 2015 project CANON (cofinanced by The Netherlands Organization for Scientific Research) to U.O., by a Amsterdam Brain and Mind Project grant – to C.M.A.P. and U.O., by a Amsterdam Brain and Cognition Project Grant to C.S.L., by the European Commission Horizon 2020 Research and Innovation Programme under grant agreement 720270 (Human Brain Project SGA1) to C.M.A.P. and grant agreement 785907 (Human Brain Project SGA2) to C.M.A.P. and U.O.

References

Ahrens, S., Jaramillo, S., Yu, K., Ghosh, S., Hwang, G.-R., Paik, R., et al., 2015. ErbB4 regulation of a thalamic reticular nucleus circuit for sensory selection. Nat. Neurosci. 18, 104–111. https://doi.org/10.1038/nn.3897.

Alvarado, J.C., Vaughan, J.W., Stanford, T.R., Stein, B.E., 2007. Multisensory versus unisensory integration: contrasting modes in the superior colliculus. J. Neurophysiol. 97, 3193–3205. https://doi.org/10.1152/jn.00018.2007.

Andersen, R.A., Snyder, L.H., Bradley, D.C., Xing, J., 1997. Multimodal representation of space in the posterior parietal cortex and its use in planning movements. Annu. Rev. Neurosci. 20, 303–330. https://doi.org/10.1146/annurev.neuro.20.1.303.

Avillac, M., Denève, S., Olivier, E., Pouget, A., Duhamel, J.-R., 2005. Reference frames for representing visual and tactile locations in parietal cortex. Nat. Neurosci. 8, 941–949. https://doi.org/10.1038/nn1480.

Baum, S.H., Stevenson, R.A., Wallace, M.T., 2015. Behavioral, perceptual, and neural alterations in sensory and multisensory function in autism spectrum disorder. Prog. Neurobiol. 134, 140–160. https://doi.org/10.1016/j.pneurobio.2015.09.007.

Beauchamp, M.S., 2005. See me, hear me, touch me: multisensory integration in lateral occipital-temporal cortex. Curr. Opin. Neurobiol. 15, 145–153. https://doi.org/10.1016/j.conb.2005.03.011.

Bell, A.H., Meredith, M.A., Van Opstal, A.J., Munoz, D.P., 2005. Crossmodal integration in the primate superior colliculus underlying the preparation and initiation of saccadic eye movements. J. Neurophysiol. 93, 3659–3673. https://doi.org/10.1152/jn.01214.2004.

Birrell, J.M., Brown, V.J., 2000. Medial frontal cortex mediates perceptual attentional set shifting in the rat. J. Neurosci. 20, 4320–4324.

Bizley, J.K., King, A.J., 2008. Visual-auditory spatial processing in auditory cortical neurons. Brain Res. 1242, 24–36. https://doi.org/10.1016/j.brainres.2008.02.087.

Bizley, J.K., Nodal, F.R., Bajo, V.M., Nelken, I., King, A.J., 2007. Physiological and anatomical evidence for multisensory interactions in auditory cortex. Cereb. Cortex 17, 2172–2189. https://doi.org/10.1093/cercor/bhl128.

Bizley, J.K., Maddox, R.K., Lee, A.K.C., 2016. Defining auditory-visual objects: behavioral tests and physiological mechanisms. Trends Neurosci. 39, 74–85. https://doi.org/10.1016/j.tins.2015.12.007.

Bremmer, F., Klam, F., Duhamel, J.-R., Ben Hamed, S., Graf, W., 2002. Visual-vestibular interactive responses in the macaque ventral intraparietal area (VIP). Eur. J. Neurosci. 16, 1569–1586.

Bruce, C., Desimone, R., Gross, C.G., 1981. Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. J. Neurophysiol. 46, 369–384.

Brunton, B.W., Botvinick, M.M., Brody, C.D., 2013. Rats and humans can optimally accumulate evidence for decision-making. Science. 340, 95–98. https://doi.org/10.1126/science.1233912.

Bussey, T.J., Saksida, L.M., Murray, E.A., 2002. Perirhinal cortex resolves feature ambiguity in complex visual discriminations. Eur. J. Neurosci. 15, 365–374.

Calvert, G.A., Thesen, T., 2004. Multisensory integration: methodological approaches and emerging principles in the human brain. J. Physiol. Paris 98, 191–205. https://doi.org/10.1016/j.jphysparis.2004.03.018.

Calvert, G.A., Bullmore, E.T., Brammer, M.J., Campbell, R., Williams, S.C., McGuire, P.K., et al., 1997. Activation of auditory cortex during silent lipreading. Science 276, 593–596.

Campi, K.L., Krubitzer, L., 2010. Comparative studies of diurnal and nocturnal rodents: differences in lifestyle result in alterations in cortical field size and number. J. Comp. Neurol. 518, 4491–4512. https://doi.org/10.1002/cne.22466.

Cappe, C., Rouiller, E.M., Barone, P., 2009. Multisensory anatomical pathways. Hear. Res. 258, 28–36. https://doi.org/10.1016/j.heares.2009.04.017.

Cappe, C., Murray, M.M., Barone, P., Rouiller, E.M., 2010. Multisensory facilitation of behavior in monkeys: effects of stimulus intensity. J. Cogn. Neurosci. 22, 2850–2863. https://doi.org/10.1162/jocn.2010.21423.

Carandini, M., Churchland, A.K., 2013. Probing perceptual decisions in rodents. Nat. Neurosci. 16, 824–831. https://doi.org/10.1038/nn.3410.

Carandini, M., Heeger, D.J., 2012. Normalization as a canonical neural computation. Nat. Rev. Neurosci. 13, 51–62. https://doi.org/10.1038/nrn3136.

Carandini, M., Heeger, D.J., Movshon, J.A., 1997. Linearity and normalization in simple cells of the macaque primary visual cortex. J. Neurosci. 17, 8621–8644.

Colonius, H., Diederich, A., 2004. Multisensory interaction in saccadic reaction time: a time-window-of-integration model. J. Cogn. Neurosci. 16, 1000–1009. https://doi.org/10.1162/0898929041502733.

Convento, S., Rahman, M.S., Yau, J.M., 2018. Selective attention gates the interactive crossmodal coupling between perceptual systems. Curr. Biol. CB. 28, 746–752. https://doi.org/10.1016/j.cub.2018.01.021. e5.

Crick, F., Koch, C., 1995. Are we aware of neural activity in primary visual cortex? Nature 375, 121–123. https://doi.org/10.1038/375121a0.

de Gelder, B., Vroomen, J., de Jong, S.J., Masthoff, E.D., Trompenaars, F.J., Hodiamont, P., 2005. Multisensory integration of emotional faces and voices in schizophrenics. Schizophr. Res. 72, 195–203. https://doi.org/10.1016/j.schres.2004.02.013.

De Meo, R., Murray, M.M., Clarke, S., Matusz, P.J., 2015. Top-down control and early multisensory processes: chicken vs. egg. Front. Integr. Neurosci. 9, 17. https://doi.org/10.3389/fnint.2015.00017.

Deneux, T., Kempf, A., Bathellier, B., 2018. Context-dependent signaling of coincident auditory and visual events in primary visual cortex. bioRxiv 258970. https://doi.org/10.1101/258970.

Diederich, A., Colonius, H., 2004. Bimodal and trimodal multisensory enhancement: effects of stimulus onset and intensity on reaction time. Percept. Psychophys. 66, 1388–1404.

Diederich, A., Colonius, H., 2008. Crossmodal interaction in saccadic reaction time: separating multisensory from warning effects in the time window of integration model. Exp. Brain Res. 186, 1–22. https://doi.org/10.1007/s00221-007-1197-4.

Dokka, K., DeAngelis, G.C., Angelaki, D.E., 2015. Multisensory integration of visual and vestibular signals improves heading discrimination in the presence of a moving object. J. Neurosci. 35, 13599–13607. https://doi.org/10.1523/JNEUROSCI.2267-15.2015.

Ernst, M.O., Banks, M.S., 2002. Humans integrate visual and haptic information in a statistically optimal fashion. Nature 415, 429–433. https://doi.org/10.1038/415429a.

Ernst, M.O., Bülthoff, H.H., 2004. Merging the senses into a robust percept. Trends Cogn Sci. 8, 162–169. https://doi.org/10.1016/j.tics.2004.02.002.

Faivre, N., Berthet, V., Kouider, S., 2014. Sustained invisibility through crowding and continuous flash suppression: a comparative review. Front. Psychol. 5, 475. https://doi.org/10.3389/fpsyg.2014.00475.

Falchier, A., Clavagnier, S., Barone, P., Kennedy, H., 2002. Anatomical evidence of multimodal integration in primate striate cortex. J. Neurosci. 22, 5749–5759 doi:20026562.

Feng, W., Störmer, V.S., Martinez, A., McDonald, J.J., Hillyard, S.A., 2014. Sounds activate visual cortex and improve visual discrimination. J. Neurosci. 34, 9817–9824. https://doi.org/10.1523/JNEUROSCI.4869-13.2014.

Fetsch, C.R., Turner, A.H., DeAngelis, G.C., Angelaki, D.E., 2009. Dynamic reweighting of visual and vestibular cues during self-motion perception. J. Neurosci. 29, 15601–15612. https://doi.org/10.1523/JNEUROSCI.2574-09.2009.

Fetsch, C.R., Pouget, A., DeAngelis, G.C., Angelaki, D.E., 2012. Neural correlates of reliability-based cue weighting during multisensory integration. Nat. Neurosci. 15, 146–154. https://doi.org/10.1038/nn.2983.

Fetsch, C.R., DeAngelis, G.C., Angelaki, D.E., 2013. Bridging the gap between theories of sensory cue integration and the physiology of multisensory neurons. Nat. Rev. Neurosci. 14, 429–442. https://doi.org/10.1038/nrn3503.

Foxworthy, W.A., Allman, B.L., Keniston, L.P., Meredith, M.A., 2013a. Multisensory and unisensory neurons in ferret parietal cortex exhibit distinct functional properties. Eur. J. Neurosci. 37, 910–923. https://doi.org/10.1111/ejn.12085.

Foxworthy, W.A., Clemo, H.R., Meredith, M.A., 2013b. Laminar and connectional organization of a multisensory cortex. J. Comp. Neurol. 521, 1867–1890. https://doi.org/10.1002/cne.23264.

Fries, P., 2015. Rhythms for cognition: communication through coherence. Neuron 88, 220–235. https://doi.org/10.1016/j.neuron.2015.09.034.

Friston, K., 2010. The free-energy principle: a unified brain theory? Nat. Rev. Neurosci. 11, 127–138. https://doi.org/10.1038/nrn2787.

Ghazanfar, A.A., Lemus, L., 2010. Multisensory integration: vision boosts information through suppression in auditory cortex. Curr. Biol. CB 20, R22–23. https://doi.org/10.1016/j.cub.2009.11.046.

Ghazanfar, A.A., Schroeder, C.E., 2006. Is neocortex essentially multisensory? Trends Cogn. Sci. 10, 278–285. https://doi.org/10.1016/j.tics.2006.04.008.

Ghazanfar, A.A., Maier, J.X., Hoffman, K.L., Logothetis, N.K., 2005. Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. J. Neurosci 25, 5004–5012. https://doi.org/10.1523/JNEUROSCI.0799-05.2005.

Gielen, S.C., Schmidt, R.A., Van den Heuvel, P.J., 1983. On the nature of intersensory facilitation of reaction time. Percept. Psychophys. 34, 161–168.

Gifford, G.W., Cohen, Y.E., 2004. Effect of a central fixation light on auditory spatial responses in area LIP. J. Neurophysiol. 91, 2929–2933. https://doi.org/10.1152/jn.01117.2003.

Gingras, G., Rowland, B.A., Stein, B.E., 2009. The differing impact of multisensory and unisensory integration on behavior. J. Neurosci. 29, 4897–4902. https://doi.org/10.1523/JNEUROSCI.4120-08.2009.
Gleiss, S., Kayser, C., 2012. Audio-visual detection benefits in the rat. PLoS One 7, e45677. https://doi.org/10.1371/journal.pone.0045677.
Goltstein, P.M., Coffey, E.B.J., Roelfsema, P.R., Pennartz, C.M.A., 2013. In vivo two-photon Ca2+ imaging reveals selective reward effects on stimulus-specific assemblies in mouse visual cortex. J. Neurosci. 33, 11540–11555. https://doi.org/10.1523/JNEUROSCI.1341-12.2013.
Goltstein, P.M., Meijer, G.T., Pennartz, C.M., 2018. Conditioning sharpens the spatial representation of rewarded stimuli in mouse primary visual cortex. eLife 7. https://doi.org/10.7554/eLife.37683.
Gregory, R.L., 1980. Perceptions as hypotheses. Philos. Trans. R. Soc. Lond. B Biol. Sci. 290, 181–197.
Gu, Y., Cheng, Z., Yang, L., DeAngelis, G.C., Angelaki, D.E., 1991. Multisensory convergence of visual and vestibular heading cues in the pursuit area of the frontal eye field. Cereb. Cortex N. Y. N. 2016 (26), 3785–3801. https://doi.org/10.1093/cercor/bhv183.
Gu, Y., Angelaki, D.E., DeAngelis, G.C., 2008. Neural correlates of multisensory cue integration in macaque MSTd. Nat. Neurosci. 11, 1201–1210. https://doi.org/10.1038/nn.2191.
Gu, Y., Deangelis, G.C., Angelaki, D.E., 2012. Causal links between dorsal medial superior temporal area neurons and multisensory heading perception. J. Neurosci. 32, 2299–2313. https://doi.org/10.1523/JNEUROSCI.5154-11.2012.
Hanks, T.D., Kopec, C.D., Brunton, B.W., Duan, C.A., Erlich, J.C., Brody, C.D., 2015. Distinct relationships of parietal and prefrontal cortices to evidence accumulation. Nature 520, 220–223. https://doi.org/10.1038/nature14066.
Harvey, C.D., Coen, P., Tank, D.W., 2012. Choice-specific sequences in parietal cortex during a virtual-navigation decision task. Nature 484, 62–68. https://doi.org/10.1038/nature10918.
Hirokawa, J., Bosch, M., Sakata, S., Sakurai, Y., Yamamori, T., 2008. Functional role of the secondary visual cortex in multisensory facilitation in rats. Neuroscience 153, 1402–1417. https://doi.org/10.1016/j.neuroscience.2008.01.011.
Hollensteiner, K.J., Pieper, F., Engler, G., König, P., Engel, A.K., 2015. Crossmodal Integration Improves Sensory Detection Thresholds in the Ferret. PLoS One 10. https://doi.org/10.1371/journal.pone.0124952.
Holmes, N.P., 2007. The law of inverse effectiveness in neurons and behaviour: multisensory integration versus normal variability. Neuropsychologia 45, 3340–3345. https://doi.org/10.1016/j.neuropsychologia.2007.05.025.
Holmes, N.P., 2009. The principle of inverse effectiveness in multisensory integration: some statistical considerations. Brain Topogr. 21, 168–176. https://doi.org/10.1007/s10548-009-0097-2.
Hwang, J., Romanski, L.M., 2015. Prefrontal neuronal responses during audiovisual mnemonic processing. J. Neurosci. 35, 960–971. https://doi.org/10.1523/JNEUROSCI.1328-14.2015.
Ibrahim, L.A., Mesik, L., Ji, X., Fang, Q., Li, H., Li, Y., et al., 2016. Cross-modality sharpening of visual cortical processing through Layer-1-Mediated inhibition and disinhibition. Neuron 89, 1031–1045. https://doi.org/10.1016/j.neuron.2016.01.027.
Iurilli, G., Ghezzi, D., Olcese, U., Lassi, G., Nazzaro, C., Tonini, R., et al., 2012. Sound-driven synaptic inhibition in primary visual cortex. Neuron 73, 814–828. https://doi.org/10.1016/j.neuron.2011.12.026.
Jackendoff, R.S., 1987. Consciousness and the Computational Mind [Internet]. Available:. MIT Press. https://mitpress.mit.edu/books/consciousness-and-computational-mind.
Jacklin, D.L., Cloke, J.M., Potvin, A., Garrett, I., Winters, B.D., 2016. The Dynamic Multisensory Engram: Neural Circuitry Underlying Crossmodal Object Recognition in Rats Changes with the Nature of Object Experience. J. Neurosci. 36, 1273–1289. https://doi.org/10.1523/JNEUROSCI.3043-15.2016.
Javitt, D.C., 2009. Sensory processing in schizophrenia: neither simple nor intact. Schizophr. Bull. 35, 1059–1064. https://doi.org/10.1093/schbul/sbp110.
Jiang, W., Wallace, M.T., Jiang, H., Vaughan, J.W., Stein, B.E., 2001. Two cortical areas mediate multisensory integration in superior colliculus neurons. J. Neurophysiol. 85, 506–522.
Jiang, W., Jiang, H., Stein, B.E., 2002. Two corticotectal areas facilitate multisensory orientation behavior. J. Cogn. Neurosci. 14, 1240–1255. https://doi.org/10.1162/089892902760807230.
Jones, P.R., 2016. A tutorial on cue combination and Signal Detection Theory: using changes in sensitivity to evaluate how observers integrate sensory information. J. Math. Psychol. 73, 117–139.
Kadunce, D.C., Vaughan, J.W., Wallace, M.T., Stein, B.E., 2001. The influence of visual and auditory receptive field organization on multisensory integration in the superior colliculus. Exp. Brain Res. 139, 303–310.
Kayser, C., Logothetis, N.K., 2007. Do early sensory cortices integrate cross-modal information? Brain Struct. Funct. 212, 121–132. https://doi.org/10.1007/s00429-007-0154-0.
Kayser, C., Petkov, C.I., Logothetis, N.K., 2008. Visual modulation of neurons in auditory cortex. Cereb. Cortex 18, 1560–1574. https://doi.org/10.1093/cercor/bhm187.
Kayser, C., Logothetis, N.K., Panzeri, S., 2010. Visual enhancement of the information representation in auditory cortex. Curr. Biol. 20, 19–24. https://doi.org/10.1016/j.cub.2009.10.068.
Khastkhodaei, Z., Jurjut, O., Katzner, S., Busse, L., 2016. Mice can use second-order, contrast-modulated stimuli to guide visual perception. J. Neurosci. 36, 4457–4469. https://doi.org/10.1523/JNEUROSCI.4595-15.2016.
Knill, D.C., Pouget, A., 2004. The Bayesian brain: the role of uncertainty in neural coding and computation. Trends Neurosci. 27, 712–719. https://doi.org/10.1016/j.tins.2004.10.007.
Knutson, B., Burgdorf, J., Panksepp, J., 2002. Ultrasonic vocalizations as indices of affective states in rats. Psychol. Bull. 128, 961–977.
Lakatos, P., Karmos, G., Mehta, A.D., Ulbert, I., Schroeder, C.E., 2008. Entrainment of neuronal oscillations as a mechanism of attentional selection. Science 320, 110–113. https://doi.org/10.1126/science.1154735.
Lee, C.C.Y., Diamond, M.E., Arabzadeh, E., 2016. Sensory prioritization in rats: behavioral performance and neuronal correlates. J. Neurosci. 36, 3243–3253. https://doi.org/10.1523/JNEUROSCI.3636-15.2016.
Lehmann, S., Murray, M.M., 2005. The role of multisensory memories in unisensory object discrimination. Brain Res. Cogn. Brain Res. 24, 326–334. https://doi.org/10.1016/j.cogbrainres.2005.02.005.
Lemus, L., Hernández, A., Luna, R., Zainos, A., Romo, R., 2010. Do Sensory Cortices Process More than One Sensory Modality during Perceptual Judgments? Neuron 67, 335–348. https://doi.org/10.1016/j.neuron.2010.06.015.
Licata, A.M., Kaufman, M.T., Raposo, D., Ryan, M.B., Sheppard, J.P., Churchland, A.K., 2017. Posterior parietal cortex guides visual decisions in rats. J. Neurosci. 37, 4954–4966. https://doi.org/10.1523/JNEUROSCI.0105-17.2017.
Limanowski, J., Blankenburg, F., 2016. Integration of Visual and Proprioceptive Limb Position Information in Human Posterior Parietal, Premotor, and Extrastriate Cortex. J. Neurosci. 36, 2582–2589. https://doi.org/10.1523/JNEUROSCI.3987-15.2016.
Lippert, M., Logothetis, N.K., Kayser, C., 2007. Improvement of visual contrast detection by a simultaneous sound. Brain Res. 1173, 102–109. https://doi.org/10.1016/j.brainres.2007.07.050.
Lippert, M.T., Takagaki, K., Kayser, C., Ohl, F.W., 2013. Asymmetric multisensory interactions of visual and somatosensory responses in a region of the rat parietal cortex. PLoS One 8, e63631. https://doi.org/10.1371/journal.pone.0063631.
Ma, W.J., Beck, J.M., Latham, P.E., Pouget, A., 2006. Bayesian inference with probabilistic population codes. Nat. Neurosci. 9, 1432–1438. https://doi.org/10.1038/nn1790.
Maier, J.X., Blankenship, M.L., Li, J.X., Katz, D.B., 2015. A multisensory network for olfactory processing. Curr. Biol. CB 25, 2642–2650. https://doi.org/10.1016/j.cub.2015.08.060.
Marcel, A.J., 1983. Conscious and unconscious perception: an approach to the relations between phenomenal experience and perceptual processes. Cognit Psychol. 15, 238–300.
Marco, E.J., Hinkley, L.B.N., Hill, S.S., Nagarajan, S.S., 2011. Sensory processing in autism: a review of neurophysiologic findings. Pediatr. Res. 69, 48R–54R. https://doi.org/10.1203/PDR.0b013e3182130c54.
Matusz, P.J., Wallace, M.T., Murray, M.M., 2017. A multisensory perspective on object memory. Neuropsychologia 105, 243–252. https://doi.org/10.1016/j.neuropsychologia.2017.04.008.
McGinley, M.J., Vinck, M., Reimer, J., Batista-Brito, R., Zagha, E., Cadwell, C.R., et al., 2015. Waking state: rapid variations modulate neural and behavioral responses. Neuron 87, 1143–1161. https://doi.org/10.1016/j.neuron.2015.09.012.
Meijer, G.T., Montijn, J.S., Pennartz, C.M.A., Lansink, C.S., 2017. Audio-visual modulation in mouse V1 depends on cross-modal stimulus configuration and congruency. J. Neurosci. https://doi.org/10.1523/JNEUROSCI.0468-17.2017.
Meijer, G.T., Pie, J.L., Dolman, T.L., Pennartz, C.M.A., Lansink, C.S., 2018. Audiovisual integration enhances stimulus detection performance in mice. Front. Behav. Neurosci. 12, 231. https://doi.org/10.3389/fnbeh.2018.00231.
Meredith, M.A., Stein, B.E., 1983. Interactions among converging sensory inputs in the superior colliculus. Science 221, 389–391.
Meredith, M.A., Stein, B.E., 1986. Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. J. Neurophysiol. 56, 640–662.
Meredith, M.A., Nemitz, J.W., Stein, B.E., 1987. Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. J. Neurosci. 7, 3215–3229.
Meyniel, F., Sigman, M., Mainen, Z.F., 2015. Confidence as bayesian probability: from neural origins to behavior. Neuron 88, 78–92. https://doi.org/10.1016/j.neuron.2015.09.039.
Miller, J., 1982. Divided attention: evidence for coactivation with redundant signals. Cognit. Psychol. 14, 247–279.
Miller, J., 1986. Timecourse of coactivation in bimodal divided attention. Percept. Psychophys. 40, 331–343.
Mohan, H., de Haan, R., Mansvelder, H.D., de Kock, C.P.J., 2017. The posterior parietal cortex as integrative hub for whisker sensorimotor information. Neuroscience. https://doi.org/10.1016/j.neuroscience.2017.06.020.
Molholm, S., Ritter, W., Javitt, D.C., Foxe, J.J., 1991. Multisensory visual-auditory object recognition in humans: a high-density electrical mapping study. Cereb. Cortex N. Y. N. 2004 (14), 452–465.
Moran, Z.D., Bachman, P., Pham, P., Cho, S.H., Cannon, T.D., Shams, L., 2013. Multisensory encoding improves auditory recognition. Multisensory Res. 26, 581–592.
Morgan, M.L., DeAngelis, G.C., Angelaki, D.E., 2008. Multisensory integration in macaque visual cortex depends on cue reliability. Neuron 59, 662–673. https://doi.org/10.1016/j.neuron.2008.06.024.
Mudrik, L., Faivre, N., Koch, C., 2014. Information integration without awareness. Trends Cogn. Sci. 18, 488–496. https://doi.org/10.1016/j.tics.2014.04.009.
Mumford, D., 1992. On the computational architecture of the neocortex. II. The role of cortico-cortical loops. Biol. Cybern. 66, 241–251.
Murray, M.M., Thelen, A., Thut, G., Romei, V., Martuzzi, R., Matusz, P.J., 2016a. The multisensory function of the human primary visual cortex. Neuropsychologia 83, 161–169. https://doi.org/10.1016/j.neuropsychologia.2015.08.011.
Murray, M.M., Lewkowicz, D.J., Amedi, A., Wallace, M.T., 2016b. Multisensory processes: a balancing act across the lifespan. Trends Neurosci. 39, 567–579. https://doi.org/10.1016/j.tins.2016.05.003.

Nikbakht, N., Tafreshiha, A., Zoccolan, D., Diamond, M.E., 2018. Supralinear and supramodal integration of visual and tactile signals in rats: psychophysics and neuronal mechanisms. Neuron 97, 626–639. https://doi.org/10.1016/j.neuron.2018.01.003. e8.

Nitz, D.A., 2006. Tracking route progression in the posterior parietal cortex. Neuron 49, 747–756. https://doi.org/10.1016/j.neuron.2006.01.037.

Ohshiro, T., Angelaki, D.E., DeAngelis, G.C., 2011. A normalization model of multisensory integration. Nat. Neurosci. 14, 775–782. https://doi.org/10.1038/nn.2815.

Ohshiro, T., Angelaki, D.E., DeAngelis, G.C., 2017. A neural signature of divisive normalization at the level of multisensory integration in primate cortex. Neuron 95, 399–411. https://doi.org/10.1016/j.neuron.2017.06.043. e8.

Olcese, U., Iurilli, G., Medini, P., 2013. Cellular and synaptic architecture of multisensory integration in the mouse neocortex. Neuron 79, 579–593. https://doi.org/10.1016/j.neuron.2013.06.010.

Pennartz, C.M.A., 2009. Identification and integration of sensory modalities: neural basis and relation to consciousness. Conscious. Cogn. 18, 718–739. https://doi.org/10.1016/j.concog.2009.03.003.

Pennartz, C.M.A., 2015. The Brain's Representational Power [Internet]. Available:. The MIT Press. https://mitpress.mit.edu/books/brains-representational-power.

Pennartz, C.M.A., 2018. Consciousness, Representation, Action: The Importance of Being Goal-Directed. Trends Cogn. Sci. 22, 137–153. https://doi.org/10.1016/j.tics.2017.10.006.

Perrault, T.J., Vaughan, J.W., Stein, B.E., Wallace, M.T., 2003. Neuron-specific response characteristics predict the magnitude of multisensory integration. J. Neurophysiol. 90, 4022–4026. https://doi.org/10.1152/jn.00494.2003.

Poort, J., Khan, A.G., Pachitariu, M., Nemri, A., Orsolic, I., Krupic, J., et al., 2015. Learning enhances sensory and multiple non-sensory representations in primary visual cortex. Neuron 86, 1478–1490. https://doi.org/10.1016/j.neuron.2015.05.037.

Portfors, C.V., 2007. Types and functions of ultrasonic vocalizations in laboratory rats and mice. J. Am. Assoc. Lab. Anim. Sci. JAALAS 46, 28–34.

Pouget, A., Deneve, S., Duhamel, J.-R., 2002. A computational perspective on the neural basis of multisensory spatial representations. Nat. Rev. Neurosci. 3, 741–747. https://doi.org/10.1038/nrn914.

Raab, D.H., 1962. Statistical facilitation of simple reaction times. Trans. N. Y. Acad. Sci. 24, 574–590.

Raposo, D., Sheppard, J.P., Schrater, P.R., Churchland, A.K., 2012. Multisensory decision-making in rats and humans. J. Neurosci. 32, 3726–3735. https://doi.org/10.1523/JNEUROSCI.4998-11.2012.

Raposo, D., Kaufman, M.T., Churchland, A.K., 2014. A category-free neural population supports evolving demands during decision-making. Nat. Neurosci. 17, 1784–1792. https://doi.org/10.1038/nn.3865.

Regenbogen, C., Johansson, E., Andersson, P., Olsson, M.J., Lundström, J.N., 2016. Bayesian-based integration of multisensory naturalistic perithreshold stimuli. Neuropsychologia 88, 123–130. https://doi.org/10.1016/j.neuropsychologia.2015.12.017.

Ross, L.A., Saint-Amour, D., Leavitt, V.M., Molholm, S., Javitt, D.C., Foxe, J.J., 2007. Impaired multisensory processing in schizophrenia: deficits in the visual enhancement of speech comprehension under noisy environmental conditions. Schizophr. Res. 97, 173–183. https://doi.org/10.1016/j.schres.2007.08.008.

Rossi-Pool, R., Salinas, E., Zainos, A., Alvarez, M., Vergara, J., Parga, N., et al., 2016. Emergence of an abstract categorical code enabling the discrimination of temporally structured tactile stimuli. Proc. Natl. Acad. Sci. U. S. A. 113, E7966–E7975. https://doi.org/10.1073/pnas.1618196113.

Rowland, B., Stanford, T., Stein, B., 2007. A Bayesian model unifies multisensory spatial localization with the physiological properties of the superior colliculus. Exp. Brain Res. 180, 153–161. https://doi.org/10.1007/s00221-006-0847-2.

Schroeder, C.E., Lakatos, P., 2009. Low-frequency neuronal oscillations as instruments of sensory selection. Trends Neurosci. 32, 9–18. https://doi.org/10.1016/j.tins.2008.09.012.

Seitz, A.R., Kim, R., Shams, L., 2006. Sound facilitates visual learning. Curr. Biol. CB 16, 1422–1427. https://doi.org/10.1016/j.cub.2006.05.048.

Shams, L., Beierholm, U.R., 2010. Causal inference in perception. Trends Cogn. Sci. 14, 425–432. https://doi.org/10.1016/j.tics.2010.07.001.

Shams, L., Seitz, A.R., 2008. Benefits of multisensory learning. Trends Cogn. Sci. 12, 411–417. https://doi.org/10.1016/j.tics.2008.07.006.

Shams, L., Kamitani, Y., Shimojo, S., 2000. Illusions. What you see is what you hear. Nature 408, 788. https://doi.org/10.1038/35048669.

Shams, L., Wozny, D.R., Kim, R., Seitz, A., 2011. Influences of multisensory experience on subsequent unisensory processing. Front. Psychol. 2, 264. https://doi.org/10.3389/fpsyg.2011.00264.

Sheppard, J.P., Raposo, D., Churchland, A.K., 2013. Dynamic weighting of multisensory stimuli shapes decision-making in rats and humans. J. Vis. 13. https://doi.org/10.1167/13.6.4.

Shuler, M.G., Bear, M.F., 2006. Reward timing in the primary visual cortex. Science. 311, 1606–1609. https://doi.org/10.1126/science.1123513.

Siemann, J.K., Muller, C.L., Bamberger, G., Allison, J.D., Veenstra-VanderWeele, J., Wallace, M.T., 2015. A novel behavioral paradigm to assess multisensory processing in mice. Front. Behav. Neurosci. 8, 456. https://doi.org/10.3389/fnbeh.2014.00456.

Song, Y.-H., Kim, J.-H., Jeong, H.-W., Choi, I., Jeong, D., Kim, K., et al., 2017. A neural circuit for auditory dominance over visual perception. Neuron 93, 940–954. https://doi.org/10.1016/j.neuron.2017.01.006. e6.

Spence, C., 2011. Crossmodal correspondences: a tutorial review. Atten. Percept. Psychophys. 73, 971–995. https://doi.org/10.3758/s13414-010-0073-7.

Stein, B.E., Stanford, T.R., 2008. Multisensory integration: current issues from the perspective of the single neuron. Nat. Rev. Neurosci. 9, 255–266. https://doi.org/10.1038/nrn2331.

Stein, B.E., Stanford, T.R., Rowland, B.A., 2009a. The neural basis of multisensory integration in the midbrain: its organization and maturation. Hear. Res. 258, 4–15. https://doi.org/10.1016/j.heares.2009.03.012.

Stein, B.E., Stanford, T.R., Ramachandran, R., Perrault, T.J., Rowland, B.A., 2009b. Challenges in quantifying multisensory integration: alternative criteria, models, and inverse effectiveness. Exp. Brain Res. 198, 113–126. https://doi.org/10.1007/s00221-009-1880-8.

Stevenson, R.A., Ghose, D., Fister, J.K., Sarko, D.K., Altieri, N.A., Nidiffer, A.R., et al., 2014. Identifying and quantifying multisensory integration: a tutorial review. Brain Topogr. https://doi.org/10.1007/s10548-014-0365-7.

Tafazoli, S., Safaai, H., De Franceschi, G., Rosselli, F.B., Vanzella, W., Riggi, M., et al., 2017. Emergence of transformation-tolerant representations of visual objects in rat lateral extrastriate cortex. eLife 6. https://doi.org/10.7554/eLife.22794.

Talsma, D., Senkowski, D., Soto-Faraco, S., Woldorff, M.G., 2010. The multifaceted interplay between attention and multisensory integration. Trends Cogn. Sci. 14, 400–410. https://doi.org/10.1016/j.tics.2010.06.008.

ten Oever, S., Romei, V., van Atteveldt, N., Soto-Faraco, S., Murray, M.M., Matusz, P.J., 2016. The COGs (context, object, and goals) in multisensory processing. Exp. Brain Res. 234, 1307–1323. https://doi.org/10.1007/s00221-016-4590-z.

Todd, J.W., 1912. Reaction to Multiple Stimuli. The Science Presshttps://doi.org/10.1037/13053-000.

Tononi, G., Boly, M., Massimini, M., Koch, C., 2016. Integrated information theory: from consciousness to its physical substrate. Nat. Rev. Neurosci. 17, 450–461. https://doi.org/10.1038/nrn.2016.44.

Ursino, M., Cuppini, C., Magosso, E., 2014. Neurocomputational approaches to modelling multisensory integration in the brain: a review. Neural Netw. Off. J. Int. Neural Netw. Soc. 60, 141–165. https://doi.org/10.1016/j.neunet.2014.08.003.

van Atteveldt, N., Murray, M.M., Thut, G., Schroeder, C.E., 2014. Multisensory integration: flexible use of general operations. Neuron 81, 1240–1253. https://doi.org/10.1016/j.neuron.2014.02.044.

van Kemenade, B.M., Seymour, K., Wacker, E., Spitzer, B., Blankenburg, F., Sterzer, P., 2014. Tactile and visual motion direction processing in hMT+/V5. NeuroImage 84, 420–427. https://doi.org/10.1016/j.neuroimage.2013.09.004.

Vasconcelos, N., Pantoja, J., Belchior, H., Caixeta, F.V., Faber, J., Freire, M.A.M., et al., 2011. Cross-modal responses in the primary visual cortex encode complex objects and correlate with tactile discrimination. Proc. Natl. Acad. Sci. 108, 15408–15413. https://doi.org/10.1073/pnas.1102780108.

Vergara, J., Rivera, N., Rossi-Pool, R., Romo, R., 2016. A neural parametric code for storing information of more than one sensory modality in working memory. Neuron 89, 54–62. https://doi.org/10.1016/j.neuron.2015.11.026.

Vincis, R., Fontanini, A., 2016. Associative learning changes cross-modal representations in the gustatory cortex. eLife 5. https://doi.org/10.7554/eLife.16420.

von, Helmholtz H., 1867. Handbuch der physiologischen Optik. Voss.

Wallace, M.T., Meredith, M.A., Stein, B.E., 1993. Converging influences from visual, auditory, and somatosensory cortices onto output neurons of the superior colliculus. J. Neurophysiol. 69, 1797–1809.

Wallace, M.T., Ramachandran, R., Stein, B.E., 2004. A revised view of sensory cortical parcellation. Proc. Natl. Acad. Sci. U. S. A. 101, 2167–2172. https://doi.org/10.1073/pnas.0305697101.

Wang, Q., Burkhalter, A., 2013. Stream-related preferences of inputs to the superior colliculus from areas of dorsal and ventral streams of mouse visual cortex. J. Neurosci. 33, 1696–1705. https://doi.org/10.1523/JNEUROSCI.3067-12.2013.

Wimmer, R.D., Schmitt, L.I., Davidson, T.J., Nakajima, M., Deisseroth, K., Halassa, M.M., 2015. Thalamic control of sensory selection in divided attention. Nature 526, 705–709. https://doi.org/10.1038/nature15398.

Winters, B.D., Reid, J.M.A., 2010. Distributed cortical representation underlies cross-modal object recognition in rats. J. Neurosci. 30, 6253–6261. https://doi.org/10.1523/JNEUROSCI.6073-09.2010.

Zhang, S., Xu, M., Kamigaki, T., Hoang Do, J.P., Chang, W.-C., Jenvay, S., et al., 2014. Selective attention. Long-range and local circuits for top-down modulation of visual cortex processing. Science 345, 660–665. https://doi.org/10.1126/science.1254126.