



HHS Public Access

Author manuscript

Trends Neurosci. Author manuscript; available in PMC 2017 August 01.

Published in final edited form as:

Trends Neurosci. 2016 August ; 39(8): 567–579. doi:10.1016/j.tins.2016.05.003.

Multisensory Processes: A Balancing Act across the Lifespan

Micah M. Murray^{1,2,3,4,*}, David J. Lewkowicz⁵, Amir Amedi^{6,7}, and Mark T. Wallace^{4,8,9,10,11,*}

¹The Laboratory for Investigative Neurophysiology (The LINE), Department of Clinical Neurosciences and Department of Radiology, University Hospital Centre and University of Lausanne, Lausanne, Switzerland ²Electroencephalography Brain Mapping Core, Centre for Biomedical Imaging (CIBM), Lausanne, Switzerland ³Department of Ophthalmology, University of Lausanne, Jules-Gonin Eye Hospital, Lausanne, Switzerland ⁴Department of Hearing and Speech Sciences, Vanderbilt University Medical Center, Nashville, TN, USA ⁵Department of Communication Sciences and Disorders, Northeastern University, Boston, MA, USA ⁶Department of Medical Neurobiology, Institute for Medical Research Israel-Canada (IMRIC), Hadassah Medical School, Hebrew University of Jerusalem, Jerusalem, Israel ⁷Interdisciplinary Center for Neural Computation, The Edmond & Lily Safra Center for Brain Sciences (ELSC), Hebrew University of Jerusalem, Jerusalem, Israel ⁸Department of Psychology, Vanderbilt University, Nashville, TN, USA ⁹Kennedy Center for Research on Human Development, Vanderbilt University, Nashville, TN, USA ¹⁰Department of Psychiatry, Vanderbilt University, Nashville, TN, USA ¹¹Vanderbilt Brain Institute, Vanderbilt University, Nashville, TN, USA

Abstract

Multisensory processes are fundamental in scaffolding perception, cognition, learning and behaviour. How and when stimuli from different sensory modalities are integrated rather than treated as separate entities is poorly understood. We review how the relative reliance on stimulus characteristics versus learned associations dynamically shapes multisensory processes. We illustrate the dynamism in multisensory function across two timescales: one long-term that operates across the lifespan and one short-term that operates during the learning of new multisensory relations. In addition, we highlight the importance of task contingencies. We conclude that these highly dynamic multisensory processes, based on the relative weighting of stimulus characteristics and learned associations, provide both stability and flexibility to brain functions over a wide range of temporal scales.

Keywords

multisensory; crossmodal; development; plasticity; learning; aging

*Correspondence: micah.murray@chuv.ch (M.M. Murray) or mark.wallace@vanderbilt.edu (M.T. Wallace).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Advantages of a multisensory brain

Our world is comprised of a vast array of information that is coded in different sensory modalities. This presents a fundamental challenge for cognitive, perceptual, and motor systems because the multitude of multisensory inputs needs to be seamlessly integrated and appropriately segregated to form coherent perceptual representations and drive adaptive behaviours. Fortunately, the nervous system possesses specialized architectures and processing mechanisms that enable the combination and integration of multisensory information. These mechanisms and brain networks solve fundamental issues associated with sensory “**binding**” and consequently provide marked behavioural and perceptual benefits (see **Glossary**). Take as an example our ability to comprehend a single speaker in a noisy room. Seeing the speaker’s mouth can increase the intelligibility of the auditory signal by 15 dB [1]. This ability to extract speech information from visual cues begins in infancy when infants start babbling and need to learn native speech forms, when they have to disambiguate unfamiliar speech, and when they are learning two languages [2,3]. Extending this further, deaf individuals can learn to watch mouth movements to understand speech by lip reading, which results in extensive **cross-modal plasticity** within regions of the cerebral cortex [4]. Other examples of how **multisensory interactions** facilitate behaviour and perception abound and include faster and more accurate detection, localization and discrimination (see [5–7] for reviews).

The neural architecture and functional rules for multisensory convergence and integration have been described in a number of species and structures. Historically, multisensory functions were considered the bastion of higher-level “association” cortical regions as well as premotor cortices and sensorimotor subcortical regions (Figure 1). Recent research has changed this view by revealing that sensory systems have the capacity to influence one another even at very early processing stages (reviewed in [8–10]). The functional implications of such “early” multisensory influences are likely substantial in that they suggest a previously unrealized level of interaction that could confer tremendous flexibility on sensory function [11].

To date, much research effort has gone into providing mechanistic and circuit-based descriptions of multisensory functions. Substantially less effort has been directed to unpacking the relative contributions of lower-level versus higher level factors in determining the final product of a given multisensory interaction. For purposes of clarity, we consider these lower-level factors as the physical characteristics of the stimuli themselves (e.g., location, timing, intensity) and the higher-level factors the learned associations that are built between the different sensory modalities. Using these operational definitions of stimulus characteristics and learned associations, we seek to highlight the dynamic interplay between them and that are taking place across a host of time scales ranging from the immediate to lifespan.

It is well established that physical stimulus characteristics play a central role in shaping multisensory interactions at the neural, behavioural and perceptual levels. The environmental statistics of these characteristics include the spatial and temporal relationships of the paired stimuli to one another, as well as their relative effectiveness in eliciting a response (Box 1).

Despite their importance, these stimulus characteristics alone are unlikely to be sufficient to specify fully the outcome of a given multisensory interaction. For one, we now know that these stimulus characteristics do not function in isolation, but rather that they interact with one another [12–15]. For example, changing the location of stimuli changes their relative effectiveness in eliciting a response. In addition, we do not know how these stimulus-related factors and their interactions are affected by higher-level factors. For example, task-dependent, goal-dependent, and context-dependent factors have been shown to dramatically alter the nature of certain types of multisensory interactions, even when low-level stimulus features are held constant [11,16–19]. Collectively, this evidence argues that a complex set of interactions between the physical characteristics (including the environmental statistics) of the stimuli and the various learned associations that are acquired through short-term and long-term experience ultimately determine the end product of a given multisensory interaction. Unfortunately, despite the fact that there is now a wealth of evidence for the importance of the separate influences of lower- and higher-level factors on multisensory processing, surprisingly little work has focused on their interplay.

In this review we describe some of the relative contributions of physical stimulus characteristics and learned associations to multisensory processing (Figure 2). We focus on two very different temporal scales as a means to highlight the dynamic interplay that is continually taking place between these factors. The first temporal scale considers multisensory processing across the lifespan. Evidence indicates that there is a developmental shift from an initial emphasis on stimulus characteristics to a greater emphasis on learned associations (see [20] for a review of such evidence from studies of human infants). The second temporal scale is much shorter and refers to the processing of information according to immediate behavioural and perceptual contingencies that are constrained by prior experiences. We argue that multisensory processes taking place within this immediate time scale are highly flexible and dynamic, strongly illustrating the power of learned associations and task contingencies in shaping multisensory processing. At its extreme, such as in the case of using a **sensory substitution device**, individuals can be trained to partially or completely adapt to using an intact sensory channel to replace functions of the impaired sense. Here, we suggest that multisensory processes can be dictated almost completely by task demands and contingencies under certain circumstances (reviewed in [21]).

The young multisensory brain: Reliance on physical stimulus characteristics

Although the development of multisensory function begins prior to birth [22], studies of its development in animals and humans have focused on the postnatal period [23–26]. Collectively, this work has illustrated that multisensory processing and the brain circuits that support it mature over the course of postnatal life, and that they depend critically on early experience. Moreover, there appears to be a developmental re-weighting as maturation progresses, such that low-level physical characteristics are initially weighted more heavily and later increasingly more complex and learned attributes are prioritized. A common principle that emerges from such work is that multisensory systems pass through three developmental stages: immature, broadly tuned, and narrowly tuned (Figure 3).

Studies in human infants have typically manipulated the physical stimulus characteristics of the sensory world and examined how such stimulus changes affect behaviours such as orienting and looking. Preverbal babies will attend preferentially to (multisensory) events that are spatially and/or temporally congruent and thus associated with a common source (e.g., the sight and sound of a bouncing ball). From a developmental perspective, low-level stimulus characteristics such as spatial proximity and temporal coincidence are available prior to the construction of learned perceptual representations. Because young organisms can perceive low-level stimulus characteristics, they provide the scaffold upon which higher-level representations can be built and constitute the essential first ingredients in the assembly of more complex and sophisticated processing architectures. The early dependence on these low-level stimulus characteristics also fits within the developmental chronology of the brain itself. For example, subcortical regions responsible for simple sensorimotor behaviours mature prior to cortical regions responsible for more complex perceptual and cognitive functions. Furthermore, different brain regions and circuits mature at different rates and with different sensitivities to physical characteristics versus learned associations (this is exemplified below in our discussion of the **temporal binding window** (TBW)).

The general fact that multisensory processes are founded on an initial sensitivity to physical stimulus features and that multisensory systems become sensitive and “tuned” to increasingly sophisticated stimulus characteristics is illustrated by a variety of findings. For instance, starting at birth, infants detect low-level, rudimentary multisensory correspondences based on simple cues as intensity [27] and temporal synchrony [28,29] and fail to detect the specific multisensory identity of objects and events. This sort of low-level processing is characterized by an initial state of very broad perceptual tuning that enables young infants to bind multisensory attributes based on their shared statistical features (i.e., their spatiotemporal correlations). For example, young infants not only bind human faces and vocalizations and native-language audible and visible speech sounds, but also monkey faces and vocalizations and non-native audible and visible speech sounds. The advantage of such a more broadly tuned multisensory perceptual system is that it allows young infants to bind a larger set of auditory and visual categories of information and, in the process, bootstrap the development of more sophisticated multisensory representations. Indeed, as infants grow and as they acquire increasingly greater experience with their native world, they gradually cease binding non-native multisensory inputs and only bind native ones [30,31] (Figure 2A). This process is known as **multisensory perceptual narrowing** [20,22] and it reflects the cumulative effects of early and selective experience. As narrowing proceeds, infants’ tendency to bind objects and sounds that have shared statistical features (i.e., that are spatiotemporally congruent) but that are not part of their typical ecology declines. Crucially, during this same time period, infants’ ability to perceive increasingly more complex multisensory associations grow, likely as a result of their experience with the world.

This general developmental pattern – multisensory perceptual narrowing and the parallel gradual emergence of increasingly more sophisticated and specialized multisensory perceptual skills and expertise - is illustrated by a number of findings. For instance, starting at birth and continuing into the first several months of postnatal life, infants exhibit the ability to detect the synchrony of auditory and visual non-speech and speech information

[32–34] as well as the equivalence of isolated visible and audible speech syllables and non-native faces and vocalizations [20,33,35,36]. By five months, infants become susceptible to the multisensory speech illusion known as the **McGurk effect** [37], which indexes the binding of auditory and visual speech elements. By 6–9 months of age infants begin to detect gender [38] and affect [39] as bound multisensory perceptual constructs, by 8 months they start to selectively attend to the source of fluent audiovisual speech located in a talker's mouth [2], and by 12 months they begin to perceive the multisensory coherence of fluent audiovisual speech and the multisensory identity of their native speech [40,41] (Figure 2B). This general developmental picture is the same for pairings beyond vision and audition (e.g. olfactory modulations of vision [42] and tactile-to-visual transfer of shape information [43]). Collectively, these findings illustrate a general developmental progression in which the ability to bind increasingly more complex, higher-level, multisensory characteristics is scaffolded upon an initial ability to only bind low-level characteristics (Figure 3).

The developing multisensory brain: Increasing emphasis on learned associations

In the prior section, we showed that the earliest aspects of multisensory integration and binding depend in large measure on low-level physical stimulus characteristics. In this section, we show that as development progresses and as higher-level cortical regions and circuits mature, experience-related (i.e., learned) factors begin to take on an expanded role. Some of the strongest evidence for this developmental transition can be seen in the shift in the relative weighting of these factors in response to the temporal structure of auditory and visual cues. Although there are a number of measures of multisensory temporal function and acuity, one of the more widely employed of these is based on the concept of a **temporal binding window** (TBW), the epoch of time within which it is highly likely that stimuli from two different sensory modalities will be bound into a single perceptual entity [33,44–46]. Such a window has a great deal of adaptive value because it enables observers to integrate visual and auditory signals coming from common sources at different distances. This is done despite the fact that visual and auditory energies propagate at different speeds as they travel to the sensory receptors and despite the fact that neural conduction speeds differ across the senses (reviewed in [47]).

Recent work focused on characterizing the developmental trajectory of the TBW has found that it does not fully mature until well into adolescence [48]. Furthermore, it has been shown that the TBW differs for different stimuli and tasks in infants [32,33], children [49], and adults [45], suggesting a dependence not only on physical stimulus characteristics but also on learned associations. Most germane for the current argument is that this dependence is differentially weighted at different stages of development. For example, the TBW for the most ethologically relevant stimuli (i.e., auditory-visual speech) matures prior to the TBW for simple and arbitrary auditory-visual pairings (such as flashes and beeps) [49]. This is opposite to the pattern one would expect if the developmental progression were dictated purely by physical stimulus characteristics. In this case, the prediction would be that the integration of simple stimuli would develop earlier and bootstrap the integration of more complex stimuli. Although not direct proof, these developmental changes suggest that the

importance of early communicative signals can prioritize which features are learned first and can play an integral role in driving differences in the maturation of multisensory integrative processes and the circuits that underlie them.

The adult multisensory brain: dynamic and flexible weighting of physical characteristics and learned associations

The work discussed above underscores the remarkable malleability inherent in developing multisensory systems as well as the increasingly greater reliance on higher-level factors to shape multisensory interactions as one moves toward adulthood. When considering the role of these higher-level factors, it is also important to take into account the timeframe over which they operate. One way of doing so is to examine multisensory plasticity in adult systems. Such plasticity can provide an important window into the competition between physical characteristics and learned associations and the timeframe over which this dynamism operates. Although this work has taken a variety of forms (e.g. see the next section), we will emphasize in this section work that has focused on multisensory temporal acuity because it is likely to be exemplary of more general multisensory phenomena (and because it builds on what was introduced in the earlier section). Using classic perceptual plasticity-based approaches, it has been found that training on a simple task involving the judgment of the relative timing of an auditory and visual stimulus results in a substantial improvement in the ability to judge their timing [50,51]. More specifically, when subjects were provided with performance feedback on a trial-by-trial basis (i.e., did they get a simultaneity judgment right or wrong), their multisensory temporal acuity (as indexed by the TBW) was found to improve by approximately 40% after training (Figure 2C). These training-induced changes occurred rapidly (most of the changes happened after the first day of training), were quite durable (lasting for more than a week after the cessation of training), required feedback (simple exposure to the same stimuli had no effect), and perhaps most importantly, generalized to tasks beyond those used during training [50]. For example, training on a simultaneity judgment task with simple flashes and beeps was found to narrow the TBW for sensitivity to the speech-related McGurk illusion (reviewed in [52]). Such training-related improvement on a simple judgment that is grounded in the physical characteristics of the paired stimuli (i.e., their temporal relationship) illustrates the power of learned associations (i.e., the training based on response feedback) in influencing judgments based purely on low-level factors. Most importantly, the absence of an effect on simultaneity judgements after passive exposure to the same stimuli provides strong evidence that higher-level influences are necessary to engender the change in multisensory temporal acuity. Interestingly, changes are seen in other situations following simple passive exposure, such as shifts in the point of subjective simultaneity following overexposure to a specific asynchrony or set of asynchronies [4,53,54]. Such results highlight that different perceptual and cognitive operations can be differentially affected by physical characteristics versus learned associations.

These adult plasticity effects provide an important window into the malleability possible in multisensory processes, and show that over the course of a day or several days, learned associations can strongly modulate integrative processes that are founded in physical

stimulus characteristics (i.e., temporal relationships). Such work raises the important question – can these influences modify multisensory functions on an even shorter time scale? Intuition would suggest that they do given that we bring strong cognitive biases to any evaluation of any stimulus. Recent experiments have begun to shed important light on this issue by focusing on fundamental processes such as learning and memory.

One set of experiments has focused on introducing task-irrelevant multisensory experiences to a task that requires a discrimination within a single sense [55–60]. This allowed for deciphering whether an initial stimulus presentation in a multisensory context (e.g., a bird together with the sound that it usually makes) impacts the ability to recall this same stimulus when specified later in only a single modality. Discrimination of repeated (visual or auditory) objects is improved for items initially encountered in a semantically congruent multisensory context (i.e., the sight and sound of the same bird) versus in a strictly unisensory context. In contrast, discrimination performance is unaffected or impaired if the initial multisensory context was semantically incongruent (e.g., the bird together with the sound of an ambulance siren) or involved an otherwise meaningless stimulus in the task-irrelevant modality (Figure 2D). These data indicate that the initial encoding of information in a multisensory context during the course of a single-trial of exposure has effects that extend to later unisensory information processing and retrieval. Sensory information is differentiated according to semantics of prior multisensory experiences, even if the multisensory nature of these experiences was task-irrelevant (see [61] for a review of psychophysical findings). This finding provides strong evidence for a relatively immediate (i.e., on the time scale of seconds to minutes) impact of a learned multisensory association (i.e., the semantic relationship between auditory-visual cues) versus the physical stimulus characteristics (i.e., the simultaneity of auditory-visual cues) on subsequent sensory processing and retrieval (Figure 2D). Brain mapping and brain imaging studies have furthermore shown that these multisensory experiences first exert their effects during the initial stages of stimulus processing and within lower-level visual and auditory cortices, providing insights into how rapidly semantic information and context-related memories can affect sensory processing [55–58,62,63].

Most recently, important individual differences have been described in these same types of experiments. This work has found that the extent to which an individual integrates auditory and visual information can predict their performance on the old/new discrimination task. Furthermore, the strength of an individual's brain responses to multisensory (but not unisensory) stimuli on their initial encounter predicted whether they would show a benefit or impairment during later object discrimination. Such findings provide direct links between multisensory brain activity at one point in time and subsequent object discrimination abilities based within a single modality [59].

In addition to their relevance for furthering our understanding of multisensory processing and the time scales within which multisensory encoding and retrieval function, these findings challenge established conventions concerning memory performance, which have suggested that performance is best when encoding and recognition contexts remain constant [64]. Such findings document an intuitive but previously unrecognized advantage associated with multisensory processes; they are not limited to the immediate processing of stimuli. A

near-term extension of this work is to apply these findings across the lifespan to improve current teaching, training, and rehabilitation methods. Indeed, and as touched on below, recent research is revealing the potential of utilizing the knowledge we are gaining about multisensory processes to create more powerful remediation tools that can be used clinically in patients. This spans many types of rehabilitation, including for patients with stroke [65], sensory loss [66] (see next section) or in the adaptation to new prosthetic devices like bionic eyes, cochlear implants and prosthesis limbs, as these devices are becoming increasingly widespread in modern clinical practice [67].

The brain as a multisensory task-relevant machine

The above sections demonstrate: i) how the development and plasticity of multisensory processes across the lifespan accommodate the physical characteristics (i.e., statistics) of stimuli in the environment, ii) how a dynamic reweighting between these stimulus characteristics and learned associations ultimately shapes multisensory processing as development progresses toward adulthood, and iii) how multisensory encoding can take advantage of recent experiences to improve learning and memory. The next logical question is how heavily the balance can be tipped toward learned associations as one examines multisensory interactions. Stated in the extreme, can information processing be completely dominated by learned associations and the task contingencies accompanying them in order to completely override the low-level characteristics of the sensory inputs? If such is the case, the underlying neural architecture can no longer be conceived of as being constrained to respond to the physical characteristics of sensory information. Instead, the underlying principle governing such neural architecture would be that task and higher-level factors largely determine responsiveness (reviewed in [21]).

As detailed above, it was traditionally held that the brain is organized into distinct unisensory or sensory-specific domains, as well as into higher-level multisensory areas responsible for integrating information from the different senses. Lately, this view has undergone substantial revision (e.g., [8,9]). For example, activity in “visual” cortices can be elicited in congenitally blind individuals by other (i.e., auditory, tactile) sensory inputs; the response profiles in this case are similar to those observed in response to visual stimulation [67–70]. Such findings need to be reconciled, however, with recent observations of preserved retinotopic organization in low-level visual cortices of congenitally-blind individuals [71–73] as well as preserved inter-regional functional connectivity patterns between higher-level visual cortices independent of visual experience [74,75]. Such findings raise the question of whether the topographic organization principles governing visual, auditory and touch processing will also govern the organization in case of sensory deprivation combined with training in sensory substitution devices (SSDs) using another sense. In the extreme, one might posit that the entire brain, including primary cortices, has the capacity to be activated by any sensory modality [76], although it is prudent to retain distinctions based on the “driving” versus “modulatory” character of the sensory inputs (e.g., [11,77]).

The specificity issue has been addressed by examining the extent to which sensory cortical regions are functionally specialized, such that the given brain region exhibits preferential

responses to one class of stimuli, regardless of input modality. One higher-level “visual” region that has been intensively studied for its role as a sensory-independent task operator is the lateral-occipital (LO) cortex (Figure 2E). A great deal of convergent evidence suggests that LO plays an instrumental role in deciphering the geometrical shape of objects. This evidence comes from observations that both congenitally blind and sighted individuals activate the LO cortices in response to sounds, but only if these individuals learn to extract geometrical shapes from these sounds such as when using a visual-to-auditory SSD or when using touch to recognize objects [67,78]. It worth nothing the using novel crossmodal adaptation techniques one study even showed that such geometrical shape information is shared between touch and vision within the same neurons [79]. Furthermore, individuals do not activate LO cortices when shape rendering is not required such as when listening to the same sounds and performing an arbitrary auditory association naming task or performing an auditory object recognition task [66]. Most critically, the same picture emerges even in the sighted brain. One possible origin of this sensory-independent, task-specific organization was found in the unique connectivity pattern between this (traditionally visual) object area with the hand representation in primary somatosensory cortex [80]. Other examples of such task-dependent processes include the recruitment of the ‘visual’ word form area (vWFA) during reading via either Braille or visual-to-auditory SSDs [74], and the recruitment of the extrastriate body area (EBA) during the extraction of the shape of bodies as conveyed via visual-to-auditory SSDs [75]. Similar examples of task-specific and sensory-independent areas have since been described in other “visual” areas, even in the extreme case of congenital blindness wherein brain development took place in the absence of any visual input or experience (e.g. when the eyes never developed in anophthalmic patients). How the task-specificity of brain regions is established and maintained remains largely unresolved. One possibility stems from findings like the abovementioned resting-state and functional connectivity pattern across long-range neural networks. These patterns might drive plastic changes even in response to extreme changes in the sensory input following complete or partial sensory deprivation. More generally, such findings demonstrate that some brain regions (or maybe even most networks) are perhaps better characterized as not sensory-specific for the inputs that they ordinarily receive. Rather, they are perhaps better characterized as task-specific based on the tasks that they must accomplish or the properties of inputs (e.g. their information regarding shape) that they must extract, regardless of sensory modality. Interestingly, this interplay can be demonstrated even after very short training using SSDs in the non-deprived and normally-developed sighted brain [81]. Overall, these findings illustrate how the weighting of stimulus characteristics vs. learned associations can be tipped to a point where function (i.e., task or context) can completely drive multisensory function independent of the specific stimulus characteristics (discussed in [82]).

Concluding Remarks

Multisensory processes are ubiquitous and essential for the building and maintenance of accurate perceptual and cognitive representations. Much is known about multisensory convergence and integration at the neural level and about the behavioural and perceptual consequences of these processes in adults. However, substantially less is known about their

development and about their maintenance in adulthood. Furthermore, little is known about how the physical characteristics of the stimuli and the learned associations between them work in a coordinated (and at times competitive) fashion to determine the final multisensory product (see **Outstanding Questions**). Our review demonstrates that multisensory processing emerges gradually in development and that it is remarkably plastic and dynamic, not only in early life, but also in adulthood. It also shows that multisensory networks permit rapid adaptation not only to the changing statistics of the sensory world, but also to the changing nature of cognitive and behavioural task requirements and their context. Finally, we show that multisensory processes can not only improve learning and memory under “normal” circumstances, but that they also create opportunities for remediation in cases of sensory loss via their highly plastic and dynamic representational abilities.

Acknowledgments

Pawel J. Matusz provided helpful comments on an earlier version of the manuscript. MMM receives financial support from the Swiss National Science Foundation (grants 320030-149982 as well as the National Centre of Competence in Research project “SYNAPSY, The Synaptic Bases of Mental Disease” [project 51AU40-125759]) and the Swiss Brain League (2014 Research Prize). DJL receives support from the Eunice Kennedy Shriver National Institute of Child Health & Human Development (grant R01HD057116). AA is a European Research Council fellow and is supported by ERC-ITG grant (310809) as well as the James S. McDonnell Foundation scholar award for understanding human cognition (grant number 220020284). MTW receives support from the National Institutes of Health (grants MH063861, CA183492, DC014114, HD083211), the Simons Foundation, and the Wallace Research Foundation.

Appendix I. Glossary of terms

- **Binding:** Processes whereby different information is coded as originating from the same object/event (i.e., bound). Here, different information refers to stimulus elements conveyed via distinct senses, such as audition and vision.
- **Cross-modal plasticity:** Reorganizational processes whereby information from one sense can modulate or even drive activity in brain regions canonically associated with another sensory modality. An extreme example is the case of sensory loss.
- **McGurk effect:** Illusory percept generated when a subject fuses an incongruent auditory and visual speech token, generating a novel percept. A typical example is when an auditory /ba/ and visual /ga/ result in the perception of /da/.
- **Multisensory interaction:** The processes invoked in response to convergent inputs from multiple senses.
- **Multisensory perceptual narrowing (MPN):** An early developmental process that, due to early and typically exclusive experience with the native ecology, leads to narrowing of perceptual tuning and to increasing perceptual differentiation of an initially broadly tuned and poorly differentiated perceptual system. Ultimately, this process leads to the emergence of expertise for native multisensory inputs and a concurrent decline in responsiveness to non-native inputs.

- **Sensory substitution device (SSD):** Non-invasive human-machine interfaces that convey information usually accessed via one sense through another (e.g., when visual input is conveyed via touch or sound for a blind individual).
- **Temporal binding window (TBW):** The epoch of time within which there is a high likelihood that paired stimuli will be actively integrated or bound as indexed by some behavioural or perceptual report. It is considered a proxy measure for multisensory temporal acuity. The TBW is further predicated on the temporal principle of multisensory integration, wherein the largest gains in neural response and behaviour are typically seen when the paired stimuli are in close temporal alignment.

References

1. Sumbly WH, Pollack I. Visual contribution to speech intelligibility in noise. *J. Acoust. Soc. Am.* 1954; 26:212–215.
2. Lewkowicz DJ, Hansen-Tift AM. Infants deploy selective attention to the mouth of a talking face when learning speech. *Proc. Natl. Acad. Sci. U. S. A.* 2012; 109:1431–1436. [PubMed: 22307596]
3. Pons F, et al. Bilingualism modulates infants' selective attention to the mouth of a talking face. *Psychol. Sci.* 2015; 26:490–498. [PubMed: 25767208]
4. Bavelier D, Neville HJ. Cross-modal plasticity: where and how? *Nat. Rev. Neurosci.* 2002; 3:443–452. [PubMed: 12042879]
5. Stein, BE., editor. *The New Handbook of Multisensory Processes.* MIT Press; 2012.
6. Stein, BE.; Meredith, MA. *The Merging of the Senses.* MIT Press; 1993.
7. Murray, MM.; Wallace, MT., editors. *The Neural Bases of Multisensory Processes.* CRC Press; 2012.
8. Murray MM, et al. The multisensory function of the human primary visual cortex. *Neuropsychologia.* 2016; 83:161–169. [PubMed: 26275965]
9. Ghazanfar AA, Schroeder CE. Is neocortex essentially multisensory? *Trends Cogn Sci.* 2006; 10:278–285. [PubMed: 16713325]
10. Kayser C, Shams L. Multisensory causal inference in the brain. *PLoS Biol.* 2015; 13:e1002075. [PubMed: 25710476]
11. van Atteveldt N, et al. Multisensory integration: flexible use of general operations. *Neuron.* 2014; 81:1240–1253. [PubMed: 24656248]
12. Cappe C, et al. Looming signals reveal synergistic principles of multisensory integration. *J. Neurosci.* 2012; 32:1171–1182. [PubMed: 22279203]
13. Krueger J, et al. Spatial receptive field organization of multisensory neurons and its impact on multisensory interactions. *Hear. Res.* 2009; 258:47–54. [PubMed: 19698773]
14. Royal DW, et al. Spatiotemporal architecture of cortical receptive fields and its impact on multisensory interactions. *Exp. Brain Res.* 2009; 198:127–136. [PubMed: 19308362]
15. Carriere BN, et al. Spatial heterogeneity of cortical receptive fields and its impact on multisensory interactions. *J. Neurophysiol.* 2008; 99:2357–68. [PubMed: 18287544]
16. Talsma D, et al. The multifaceted interplay between attention and multisensory integration. *Trends Cogn. Sci.* 2010; 14:400–410. [PubMed: 20675182]
17. De Meo R, et al. Top-down control and early multisensory processes: chicken vs. egg. *Front. Integr. Neurosci.* 2015; 9:1–6. [PubMed: 25698945]
18. ten Oever S, et al. The COGs (Context-Object-Goals) in multisensory processing. *Exp. Brain Res.* 2016
19. Sarmiento BR, et al. Contextual factors multiplex to control multisensory processes. *Hum. Brain Mapp.* 2016; 37:273–288. [PubMed: 26466522]

20. Lewkowicz DJ, Ghazanfar AA. The emergence of multisensory systems through perceptual narrowing. *Trends Cogn. Sci.* 2009; 13:470–478. [PubMed: 19748305]
21. Heimler B, et al. Origins of task-specific sensory-independent organization in the visual and auditory brain: neuroscience evidence, open questions and clinical implications. *Curr. Opin. Neurobiol.* 2015; 35:169–177. [PubMed: 26469211]
22. Lewkowicz DJ. Early experience and multisensory perceptual narrowing. *Dev. Psychobiol.* 2014; 56:292–315. [PubMed: 24435505]
23. Lewkowicz, DJ.; King, AJ. The developmental and evolutionary emergence of multisensory processing: From single cells to behavior. In: Stein, BE., editor. *The New Handbook of Multisensory Processing*. MIT Press; 2012.
24. Bremner, AJ., et al. *Multisensory Development*. Oxford University Press; 2012.
25. Röder B, Wallace M. Development and plasticity of multisensory functions. *Restor. Neurol. Neurosci.* 2010; 28:141–142. [PubMed: 20404403]
26. Polley DB, et al. Development and Plasticity of Intra- and Intersensory Information Processing. *J. Am. Acad. Audiol.* 2008; 19:780–798. [PubMed: 19358458]
27. Lewkowicz DJ, Turkewitz G. Cross-modal equivalence in early infancy: Auditory-visual intensity matching. *Dev. Psychol.* 1980; 16:597–607.
28. Lewkowicz DJ, et al. Intersensory Perception at Birth: Newborns Match Nonhuman Primate Faces and Voices. *Infancy.* 2010; 15:46–60.
29. Bahrick LE, Lickliter R. Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Dev. Psychol.* 2000; 36:190–201. [PubMed: 10749076]
30. Pons F, et al. Narrowing of intersensory speech perception in infancy. *Proc. Natl. Acad. Sci. USA.* 2009; 106:10598–10602. [PubMed: 19541648]
31. Lewkowicz DJ, Ghazanfar AA. The decline of cross-species intersensory perception in human infants. *Proc. Natl. Acad. Sci. U. S. A.* 2006; 103:6771–6774. [PubMed: 16618919]
32. Lewkowicz DJ. Infant perception of audio-visual speech synchrony. *Dev. Psychol.* 2010; 46:66–77. [PubMed: 20053007]
33. Lewkowicz DJ. Perception of auditory-visual temporal synchrony in human infants. *J. Exp. Psychol. Hum. Percept. Perform.* 1996; 22:1094–1106. [PubMed: 8865617]
34. Scheier C, et al. Sound induces perceptual reorganization of an ambiguous motion display in human infants. *Dev. Sci.* 2003; 6:233–241.
35. Kuhl PK, Meltzoff AN. The bimodal perception of speech in infancy. *Science.* 1982; 218:1138–1141. [PubMed: 7146899]
36. Patterson ML, Werker JF. Matching phonetic information in lips and voice is robust in 4.5-month-old infants. *Infant Behav. Dev.* 1999; 22:237–247.
37. Rosenblum LD, et al. The McGurk effect in infants. *Percept. Psychophys.* 1997; 59:347–357. [PubMed: 9136265]
38. Hillairet de Boisferon A, et al. Perception of Multisensory Gender Coherence in 6- and 9-Month-Old Infants. *Infancy.* 2015; 20:661–674. [PubMed: 26561475]
39. Walker-Andrews AS. Infants' perception of expressive behaviors: differentiation of multimodal information. *Psychol. Bull.* 1997; 121:437–456. [PubMed: 9136644]
40. Lewkowicz DJ, et al. Perception of the multisensory coherence of fluent audiovisual speech in infancy: Its emergence and the role of experience. *J. Exp. Child Psychol.* 2015; 130:147–162. [PubMed: 25462038]
41. Lewkowicz DJ, Pons F. Recognition of Amodal Language Identity Emerges in Infancy. *Int. J. Behav. Dev.* 2013; 37:90–94. [PubMed: 24648601]
42. Durand K, et al. Eye-catching odors: olfaction elicits sustained gazing to faces and eyes in 4-month-old infants. *PLoS One.* 2013; 8:e70677. [PubMed: 24015175]
43. Gottfried AW, et al. Cross-modal transfer in human infants. *Child Dev.* 1977; 48:118–123. [PubMed: 844349]
44. Colonius H, Diederich A. Multisensory interaction in saccadic reaction time: a time-window-of-integration model. *J Cogn Neurosci.* 2004; 16:1000–1009. [PubMed: 15298787]

45. Stevenson, Ra; Wallace, MT. Multisensory temporal integration: task and stimulus dependencies. *Exp. brain Res.* 2013; 227:249–261. [PubMed: 23604624]
46. Stevenson, Ra, et al. Interactions between the spatial and temporal stimulus factors that influence multisensory integration in human performance. *Exp. Brain Res.* 2012; 219:121–137. [PubMed: 22447249]
47. Musacchia G, Schroeder CE. Neuronal mechanisms, response dynamics and perceptual functions of multisensory interactions in auditory cortex. *Hear. Res.* 2009; 258:72–79. [PubMed: 19595755]
48. Hillock-Dunn A, Wallace MT. Developmental changes in the multisensory temporal binding window persist into adolescence. *Dev. Sci.* 2012; 15:688–696. [PubMed: 22925516]
49. Hillock-Dunn A, et al. The temporal binding window for audiovisual speech: Children are like little adults. *Neuropsychologia.* 2016
50. Powers AR, et al. Perceptual training narrows the temporal window of multisensory binding. *J. Neurosci.* 2009; 29:12265–12274. [PubMed: 19793985]
51. Powers AR, et al. Neural correlates of multisensory perceptual learning. *J. Neurosci.* 2012; 32:6263–6274. [PubMed: 22553032]
52. Wallace MT, Stevenson RA. The construct of the multisensory temporal binding window and its Dysregulation in developmental Disabilities. *Neuropsychologia.* 2014
53. Fiehler K, Rösler F. Plasticity of multisensory dorsal stream functions: evidence from congenitally blind and sighted adults. *Restor. Neurol. Neurosci.* 2010; 28:193–205. [PubMed: 20404408]
54. Ricciardi E, et al. Mind the blind brain to understand the sighted one! Is there a supramodal cortical functional architecture? *Neurosci. Biobehav. Rev.* 2014; 41:64–77. [PubMed: 24157726]
55. Murray MM, et al. Rapid discrimination of visual and multisensory memories revealed by electrical neuroimaging. *Neuroimage.* 2004; 21:125–135. [PubMed: 14741649]
56. Murray MM, et al. The brain uses single-trial multisensory memories to discriminate without awareness. *Neuroimage.* 2005; 27:473–478. [PubMed: 15894494]
57. Lehmann S, Murray MM. The role of multisensory memories in unisensory object discrimination. *Brain Res. Cogn. Brain Res.* 2005; 24:326–334. [PubMed: 15993770]
58. Thelen A, et al. Electrical neuroimaging of memory discrimination based on single-trial multisensory learning. *Neuroimage.* 2012; 62:1478–1488. [PubMed: 22609795]
59. Thelen A, et al. Multisensory context portends object memory. *Curr. Biol.* 2014; 24:R734–R735. [PubMed: 25137580]
60. Thelen A, et al. Single-trial multisensory memories affect later auditory and visual object discrimination. *Cognition.* 2015; 138:148–160. [PubMed: 25743256]
61. Thelen A, Murray MM. The efficacy of single-trial multisensory memories. *Multisens. Res.* 2013; 26:483–502. [PubMed: 24649531]
62. Schall S, et al. Early auditory sensory processing of voices is facilitated by visual mechanisms. *Neuroimage.* 2013; 77:237–245. [PubMed: 23563227]
63. Matusz PJ, et al. The role of auditory cortices in the retrieval of single-trial auditory-visual object memories. *Eur. J. Neurosci.* 2015; 41:699–708. [PubMed: 25728186]
64. Baddeley, A., et al. *Memory.* Psychology Press; 2009.
65. Johansson BB. Multisensory stimulation in stroke rehabilitation. *Front. Hum. Neurosci.* 2012; 6:60. [PubMed: 22509159]
66. Reich L, et al. The brain as a flexible task machine: implications for visual rehabilitation using noninvasive vs. invasive approaches. *Curr. Opin. Neurol.* 2012; 25:86–95. [PubMed: 22157107]
67. Amedi A, et al. Visuo-haptic object-related activation in the ventral visual pathway. *Nat. Neurosci.* 2001; 4:324–330. [PubMed: 11224551]
68. Amedi A, et al. Shape conveyed by visual-to-auditory sensory substitution activates the lateral occipital complex. *Nat. Neurosci.* 2007; 10:687–689. [PubMed: 17515898]
69. Merabet LB, Pascual-Leone A. Neural reorganization following sensory loss: the opportunity of change. *Nat. Rev. Neurosci.* 2010; 11:44–52. [PubMed: 19935836]
70. Reich L, et al. A ventral visual stream reading center independent of visual experience. *Curr. Biol.* 2011; 21:363–368. [PubMed: 21333539]

71. Striem-Amit E, et al. Functional connectivity of visual cortex in the blind follows retinotopic organization principles. *Brain*. 2015; 138:1679–1695. [PubMed: 25869851]
72. Wang X, et al. How Visual Is the Visual Cortex? Comparing Connectional and Functional Fingerprints between Congenitally Blind and Sighted Individuals. *J. Neurosci*. 2015; 35:12545–12559. [PubMed: 26354920]
73. Bock AS, et al. Resting-State Retinotopic Organization in the Absence of Retinal Input and Visual Experience. *J. Neurosci*. 2015; 35:12366–12382. [PubMed: 26354906]
74. Striem-Amit E, et al. Reading with sounds: sensory substitution selectively activates the visual word form area in the blind. *Neuron*. 2012; 76:640–652. [PubMed: 23141074]
75. Striem-Amit E, Amedi A. Visual cortex extrastriate body-selective area activation in congenitally blind people ‘seeing’ by using sounds. *Curr. Biol*. 2014; 24:687–692. [PubMed: 24613309]
76. Kayser C, Logothetis NK. Do early sensory cortices integrate cross-modal information? *Brain Struct. Funct*. 2007; 212:121–132. [PubMed: 17717687]
77. Meredith MA, Allman BL. Subthreshold multisensory processing in cat auditory cortex. *Neuroreport*. 2009; 20:126–131. [PubMed: 19057421]
78. Amedi A, et al. Cortical activity during tactile exploration of objects in blind and sighted humans. *Restor. Neurol. Neurosci*. 2010; 28:143–156. [PubMed: 20404404]
79. Tal N, Amedi A. Multisensory visual-tactile object related network in humans: insights gained using a novel crossmodal adaptation approach. *Exp. Brain Res*. 2009; 198:165–182. [PubMed: 19652959]
80. Tal Z, et al. The origins of metamodality in visual object area LO: Bodily topographical biases and increased functional connectivity to S1. *Neuroimage*. 2015
81. Hertz U, Amedi A. Flexibility and Stability in Sensory Processing Revealed Using Visual-to-Auditory Sensory Substitution. *Cereb. Cortex*. 2015; 25:2049–2064. [PubMed: 24518756]
82. Murray MM, et al. Neuroplasticity: Unexpected Consequences of Early Blindness. *Curr. Biol*. 2015; 25:R998–R1001. [PubMed: 26485377]
83. Falchier A, et al. Anatomical evidence of multimodal integration in primate striate cortex. *J. Neurosci*. 2002; 22:5749–5759. [PubMed: 12097528]
84. Falchier A, et al. Projection from visual areas V2 and prostriata to caudal auditory cortex in the monkey. *Cereb. Cortex*. 2010; 20:1529–1538. [PubMed: 19875677]
85. Rockland KS, Ojima H. Multisensory convergence in calcarine visual areas in macaque monkey. *Int. J. Psychophysiol*. 2003; 50:19–26. [PubMed: 14511833]
86. Beer AL, et al. Diffusion tensor imaging shows white matter tracts between human auditory and visual cortex. *Exp. Brain Res*. 2011; 213:299–308. [PubMed: 21573953]
87. Beer AL, et al. Combined diffusion-weighted and functional magnetic resonance imaging reveals a temporal-occipital network involved in auditory-visual object processing. *Front. Integr. Neurosci*. 2013; 7:5. [PubMed: 23407860]
88. Ghazanfar, Aa; Schroeder, CE. Is neocortex essentially multisensory? *Trends Cogn. Sci*. 2006; 10:278–285. [PubMed: 16713325]
89. Cappe C, et al. Selective integration of auditory-visual looming cues by humans. *Neuropsychologia*. 2009; 47:1045–1052. [PubMed: 19041883]
90. Wallace MT, et al. Visual experience is necessary for the development of multisensory integration. *J. Neurosci*. 2004; 24:9580–9584. [PubMed: 15509745]
91. Wallace MT, et al. The development of cortical multisensory integration. *J. Neurosci*. 2006; 26:11844–11849. [PubMed: 17108157]
92. Wallace MT, Stein BE. Development of multisensory neurons and multisensory integration in cat superior colliculus. *J. Neurosci*. 1997; 17:2429–2444. [PubMed: 9065504]
93. Stevenson, Ra, et al. Identifying and quantifying multisensory integration: a tutorial review. *Brain Topogr*. 2014; 27:707–730. [PubMed: 24722880]
94. Panzeri S, et al. Neural population coding: combining insights from microscopic and mass signals. *Trends Cogn. Sci*. 2015; 19:162–172. [PubMed: 25670005]

Trends Box

- Multisensory processes are the product of a dynamic reweighting of physical stimulus characteristics and learned associations.
- This reweighting happens across multiple timescales, ranging from long-term (i.e., developmental and lifespan) through short-term (i.e., during the learning and encoding of multisensory relations).
- We propose a novel theoretical framework that combines traditional principles associated with stimulus characteristics (i.e., space, time, effectiveness) together with a new principle of dynamic reweighting of stimulus characteristics and learned associations across different timescales.
- The novel theoretical framework emphasizes the plastic and dynamic nature of multisensory processing across the lifespan and, thus, accounts for improvements in multisensory perception and behaviour under “normal” circumstances and offers plausible remediation strategies for treating patients in which sensory and multisensory function is compromised.

Box 1**A primer on principles of multisensory integration**

Convergence of inputs from multiple sensory modalities has been long-known for classic regions of association cortex as well as for subcortical structures (reviewed in [6]). This traditional view has recently been complemented by evidence of a surprising degree of connectivity between areas traditionally deemed “unisensory” (Figure 1)([83–87]; reviewed in [8,88]). The functional organization and behavioural and perceptual relevance of these cross-modal connectivity patterns are a focus of ongoing research. Nonetheless, functional studies are describing the neurobiological consequences of multisensory convergence spanning levels of analysis ranging from single-units in animal models to human imaging studies of large neural network populations [5,7]. A recurring finding is that multisensory stimulation often results in response or activation changes different from those predicted based on the addition of the component unisensory responses.

Several fundamental processes appear to characterize the integrative features of multisensory neurons. One process revolves around the physical properties of the stimuli to be integrated, including their spatial and temporal proximity. Generally, stimuli presented close together in space or time enhance both neural and behavioural responses (the spatial and temporal principles, respectively). Separating stimuli in these dimensions generally reduces the magnitude of these interactions. Both the response gains and response reductions are often strikingly non-linear (e.g. [12,15,89]). Likewise, pairing weakly effective stimuli generally results in the largest proportionate response gains (the inverse effectiveness principle). These sorts of stimulus-dependent principles make ecological sense; spatially and temporally proximate cues generally originate from a common source, and highly effective stimuli need little amplification [6].

The basic patterns of multisensory development look very similar across species and in many different brain structures. Moreover, the global developmental progression is one in which inputs from the individual sensory systems first appear in structures that will eventually become multisensory, and that as inputs from the different senses converge, neurons that are activated or influenced by inputs from multiple senses (i.e., multisensory neurons) appear and then increase gradually in incidence over postnatal life [90–92]. Although the same types of invasive recording procedures cannot be (readily) used in humans, the overall developmental pattern obtained from behavioural measures appears remarkably similar and consistent with the pattern derived from animal studies.

It has now become evident that interactions based on the physical attributes of multisensory inputs do not fully account for the products of multisensory interactions [93,94]. Low-level (i.e., stimulus-related) factors only partly account for the multi-faceted nature of multisensory processes. Instead, it is now clear that stimulus characteristics work dynamically and in concert with a series of higher-level processes including semantic congruence, attentional allocation, and task demands [11,16–18].

Box 2**Outstanding Questions**

- How does prenatal and early postnatal experience (both normal and abnormal) shape the development of multisensory processes? How do multisensory processes change across the human lifespan, particularly during the aging process?
- What are the mechanistic underpinnings supporting multisensory integration at the cellular level, and are these shared across brain structures? What are the neural network mechanisms that support the binding of information across the senses? Is there a universal code or mechanism that underlies this binding across levels?
- What are the limits on multisensory plasticity?
- What are the qualia of perceptions conveyed via SSDs? What is the optimal age to train sensory-deprived (blind or deaf) on the use of SSDs? Will such training facilitate or interfere with sensory restoration, such as in the case of retinal or cochlear implants?
- What accounts for functional and task specialization of a given brain region? How does such develop, and how is such maintained? How can the same brain region readily execute the same task regardless of sensory input and even following just hours of training with a SSD? Why is training a given brain region to perform a new task (e.g. following damage or stroke) based on the same sensory input is so challenging?
- Can multisensory training protocols remediate (multi)sensory functions? Two findings in healthy adults are particularly encouraging. First, the largest benefits in terms of changes in the size of the TBW are observed in those individuals for whom the TBW was the widest prior to training [49,89]. By extension, clinical populations with particularly enlarged multisensory TBWs, such as those living along the autism spectrum [90], may thus be most prone to benefiting from such training approaches. Second, individuals with particularly strong responses to multisensory (but not unisensory) stimuli show greater benefits of multisensory learning on later unisensory discrimination tasks [55].
- What are the brain bases for the substantial individual differences in multisensory processes? How do cognitive factors influence multisensory interactions, and how are cognitive processes correlated with individual differences in multisensory capacity?
- What role does variability (inter-trial, inter-subject, etc.) play in shaping the final product of a multisensory interaction?

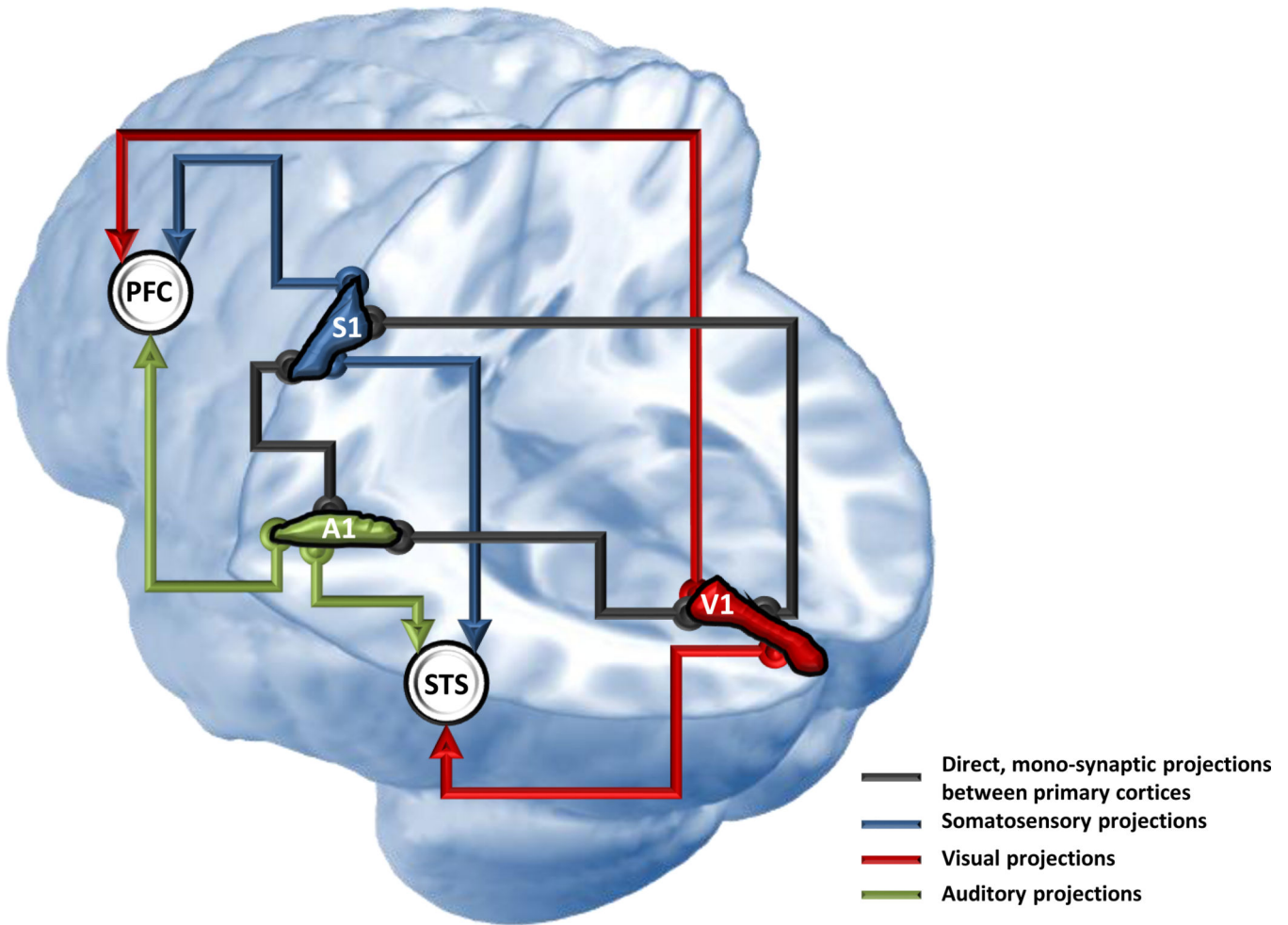


Figure 1.

Schema of cortical loci of multisensory processes. The schema is depicted on a left hemisphere, with the occipital lobe on the right side of the image and the frontal lobe on the left side. Primary visual (V1), auditory (A1), and somatosensory (S1) cortices are indicated by the red, green, and blue blobs, respectively. The correspondingly coloured arrows depict a schema where interactions are restricted to higher-level association cortices, such as the prefrontal cortices (PFC) and superior temporal sulcus (STS) (indicated by white disks). The black lines depict a schema where interactions occur directly between low-level cortices. There is now evidence in support of both schemas. Therefore, multisensory processes undoubtedly involve a dynamic combination of these schemas, which emerge as a consequence of both sensory stimulus- and experience-dependent processes.

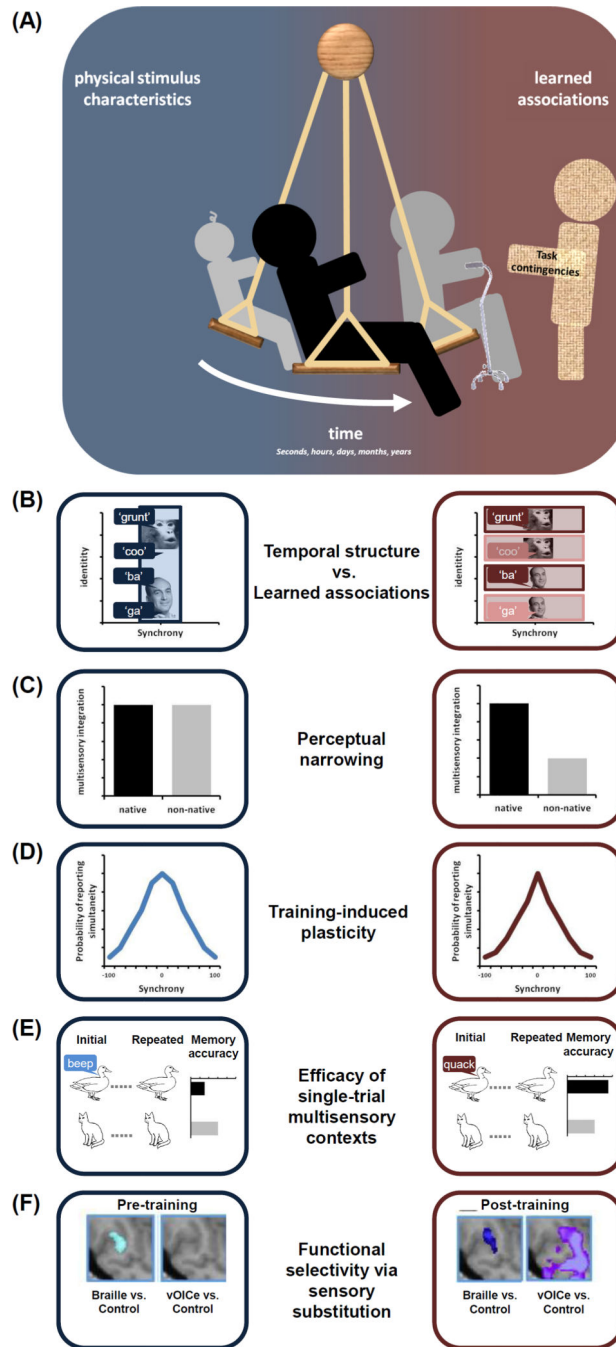


Figure 2. Schema and examples of the interplay between sensitivity to physical stimulus characteristics and sensitivity to learned associations between stimuli in shaping multisensory functions across various time scales. Sensitivity to physical stimulus characteristics is represented in blue, and sensitivity to learned associations is represented in red. **(A)** The interplay between these factors takes place over multiple time scales, ranging from seconds to years (i.e., lifespan). Task contingencies can influence the relative weight attributed to a given factor. **(B)** Illustration of the shift in the sensitivity from a heavy early

dependence on stimulus timing (i.e., synchrony) to a greater dependence on stimulus identity at later time periods. Boxes reflect the “binding” process and the shift from low-level temporal factors to higher-level learned associations. In some cases, including infancy, this shift is itself linked with perceptual narrowing, as depicted in panel (C). **(C)** Illustration of perceptual narrowing characterized by early broad multisensory tuning and later narrower tuning. The broad tuning leads infants to bind non-native as well as native auditory and visual stimuli largely because responsiveness is mediated by low-level synchrony cues, whereas the later narrower tuning leads infants to only bind native auditory and visual stimuli because responsiveness is now mediated by higher-level identity cues. **(D)** An example of how training can modify multisensory processes. Here the left and right panels represent the audiovisual temporal binding window as measured by synchrony judgments before and after feedback training. Note the post-training narrowing of the distribution, reflective of the malleability in multisensory temporal acuity. Importantly, it should be noted that these training effects must be interpreted in the larger developmental context where plasticity generally decreases from infancy into adulthood. Thus, these results indicate that there is a great deal of latent plasticity in adulthood. **(E)** An example of how memory function (e.g. assayed by repetition discrimination with unisensory visual stimuli) is affected when prior stimuli are presented in a multisensory context. When this context entails meaningless sounds such as an image of a duck paired with a pure tone, recall accuracy is impaired versus that for images only presented visually. By contrast, when the prior context entails semantically congruent meaningful sounds, such as an image of a duck paired with a quacking sound, recall accuracy is improved versus that for images only presented visually. **(F)** An example of how training with a sensory substitution device can modify responses within higher-level visual cortices (here the vWFA). Prior to training, differential responses are observed only during Braille-reading of words, but not when hearing transformations of written words via a sensory substitution device. After training, differential responses are observed for both Braille-reading and sensory substitution.

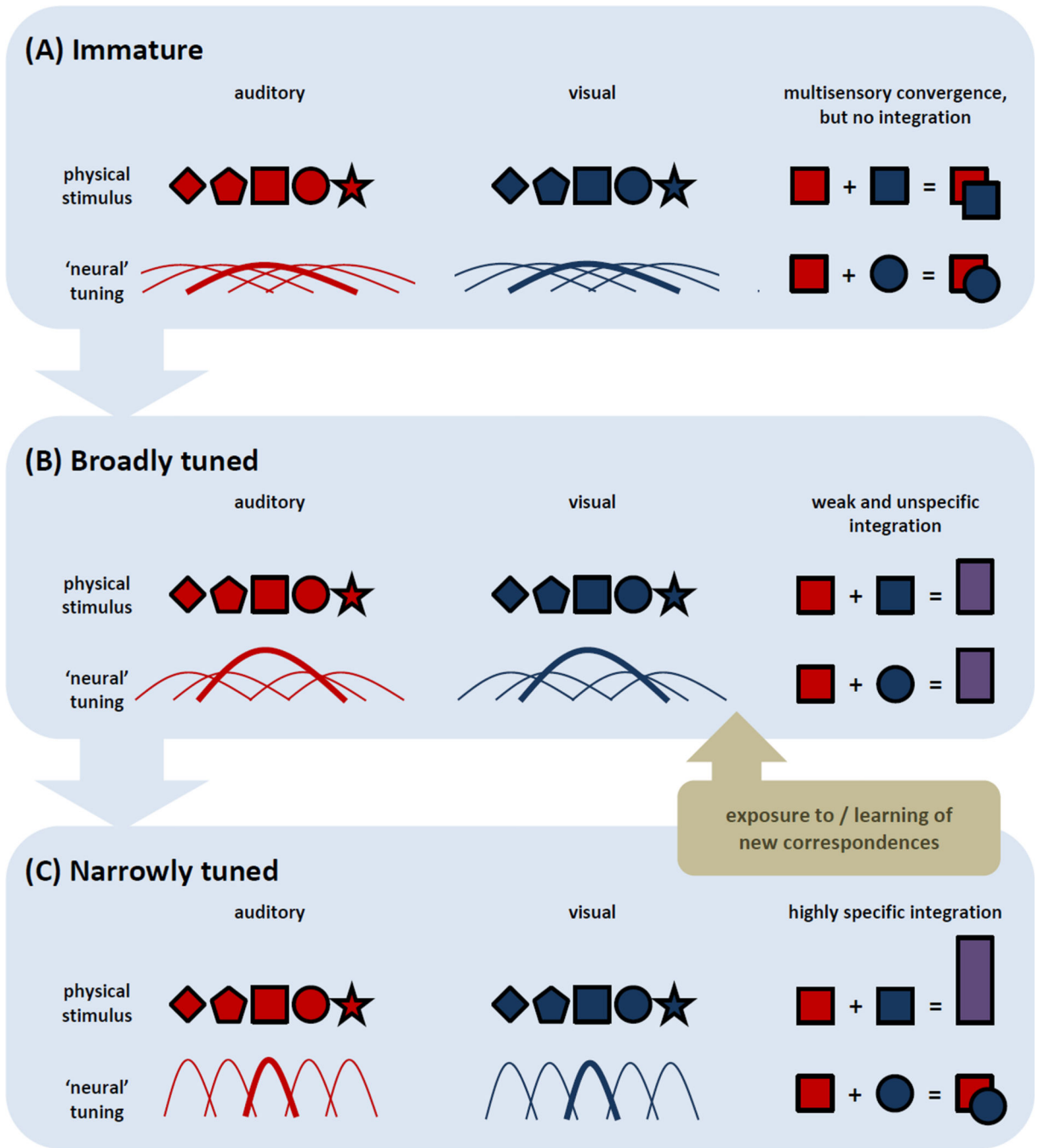


Figure 3. Schemas of the consequences for multisensory integration across three canonical developmental stages. In these schemas, auditory and visual stimulus parameters are denoted by red and blue geometric shapes, respectively, and corresponding shapes refer to features of the same object. The curves in these schemas refer to tuning profiles of neural populations, and the tuning function for an exemplar stimulus parameter (i.e., ‘square’) is highlighted. The right side of these schemas depict putative responses to co-presentation of a given auditory and visual stimulus parameter. **(A)** During an immature stage, neural tuning is

extremely broad and responses are typically of low magnitude (due principally to the immaturity of the sensory systems themselves). Although multisensory convergence is likely, multisensory integration does not occur. **(B)** During an intermediate stage, neural and representational tuning narrows, but still remains relatively broad. Multisensory interactions are now seen (denoted by purple rectangles), and can occur for a broader range of stimulus attributes than seen at later stages. In these initial two stages, low-level physical stimulus characteristics bootstrap the construction of category-general, multisensory representations that become more specific and constrained due to experience with native multisensory inputs. This occurs particularly extensively over the first year of life in humans. **(C)** During a final stage, neural and representational tuning narrows and becomes highly specialised as behaviours become increasingly refined and sophisticated. While integration occurs with co-presentation of stimulus attributes shared across the senses (e.g., same location), no integration occurs with the pairing of unshared attributes. While this absence of integration might resemble what is observed during an immature stage, its cause is instead due to the narrowly tuned representations within each sense (i.e. The 'circle' falls outside of the tuning for the 'square'). Most importantly, in adulthood there can be dynamic shifts between stages **(C)** and **(B)** based on learned associations and task contingencies (denoted by the brown box), which can override low-level stimulus factors in order to promote the integration of a broader class of stimulus attributes. The full mechanisms of these shifts, including any impact on neural tuning and tuning for a given attribute, remain underexplored. Likewise, it is important to note that specific circuits (and their associated functions) mature at different rates, which also remains largely underexplored.