



Published in final edited form as:

J Comp Neurol. 2013 December 15; 521(18): . doi:10.1002/cne.23408.

Neural convergence and divergence in the mammalian cerebral cortex: from experimental neuroanatomy to functional neuroimaging

Kingson Man, Jonas Kaplan, Hanna Damasio, and Antonio Damasio

Brain and Creativity Institute, University of Southern California, Los Angeles, CA, 90089

Abstract

A development essential for understanding the neural basis of complex behavior and cognition is the description, during the last quarter of the twentieth century, of detailed patterns of neuronal circuitry in the mammalian cerebral cortex. This effort established that sensory pathways exhibit successive levels of convergence, from the early sensory cortices to sensory-specific association cortices and to multisensory association cortices, culminating in maximally integrative regions; and that this convergence is reciprocated by successive levels of divergence, from the maximally integrative areas all the way back to the early sensory cortices. This article first provides a brief historical review of these neuroanatomical findings, which were relevant to the study of brain and mind-behavior relationships using a variety of approaches and to the proposal of heuristic anatomo-functional frameworks. In a second part, the article reviews new evidence that has accumulated from studies of functional neuroimaging, employing both univariate and multivariate analyses, as well as electrophysiology, in humans and other mammals, that the integration of information across the auditory, visual, and somatosensory-motor modalities proceeds in a content-rich manner. Behaviorally and cognitively relevant information is extracted from and conserved across the different modalities, both in higher-order association cortices and in early sensory cortices. Such stimulus-specific information is plausibly relayed along the neuroanatomical pathways alluded to above. The evidence reviewed here suggests the need for further in-depth exploration of the intricate connectivity of the mammalian cerebral cortex in experimental neuroanatomical studies.

Keywords

Van Hoesen; tract-tracing; convergence; divergence; fMRI; multivariate; MVPA; multisensory; crossmodal

1. Introduction

In the last quarter of the twentieth century, a critical development occurred in the history of understanding complex behavior and cognition. This development was the systematic description of detailed patterns of neuronal circuitry in the mammalian cerebral cortex. In this section, we first provide a brief review of these studies of experimental neuroanatomy.

Conflict of interest statement:

The authors declare no conflicts of interest.

Statement of authors' contributions:

A.D. and H.D. conceptualized the article and wrote the introduction. K.M. reviewed the neuroanatomical and cognitive neuroscience evidence and wrote the related text. A.D., K.M., J.K., and H.D. decided on the conclusions to draw from the review and participated in the final writing and editing of the manuscript.

Of particular interest are a number of landmark neuroanatomical studies, conducted in non-human primates and authored by, among others, E.G. Jones, T.P. Powell, Deepak Pandya, Gary W. Van Hoesen, Kathleen Rockland, and Marsel Mesulam, that revealed intriguing sets of connections projecting from primary sensory regions to successive regions of association cortex. The connections were organized according to an ordered and hierarchical architecture whose likely functional result was a convergence of diverse sensory signals into certain cortical areas, which, of necessity, became multisensory. The studies also revealed that, in turn, the convergent projections were usually reciprocated, diverging in succession back to the originating primary sensory regions. For example, Jones and Powell (1970) found that the primary sensory areas project unidirectionally to their adjacent sensory-specific association areas. The initial strict topography of the primary sensory areas was progressively lost along the stream of associational processing. The sensory-specific association areas further converged onto multimodal association areas, most notably in the depths of the superior temporal sulcus, and these multimodal association areas originated reciprocal, divergent projections back to the sensory-specific association cortices. The grand design of this sort of architecture was made quite transparent by the manner in which connections from varied sensory regions, such as visual and auditory, gradually converged into the hippocampal formation via the entorhinal cortex, a critical gateway into the hippocampus; and by how, in turn, the entorhinal cortex initiated diverging projections that reciprocated the converging pathways all the way back to the originating cortices. In this regard, the work of Gary W. Van Hoesen and of his colleagues stands out by offering remarkable evidence for sensory convergence and divergence. Specifically, Van Hoesen, Pandya, and Butters (1972) revealed the sources of afferents to the entorhinal cortex. Three cortical regions, themselves the recipients of already highly convergent multimodal input, were found to be the main inputs to entorhinal cortex. They were the parahippocampal cortex; prepiriform cortex; and ventral frontal cortex (Fig. 1). In 1975 this same research group published a more comprehensive description of the cortical afferents and efferents of the entorhinal cortex describing (a) afferents from ventral temporal cortex (Van Hoesen and Pandya 1975a); (b) afferents from orbitofrontal cortex (Van Hoesen et al., 1975); and (c) efferents that form the perforant pathway of the hippocampus (Van Hoesen and Pandya 1975b).

The connectional patterns that emerged from these studies were novel and functionally suggestive. Through a series of convergent steps, the hippocampus was provided with modality-specific and multisensory input hailing from association areas throughout the cerebral cortex. The entorhinal and perirhinal cortices, which are in of themselves divisible into subregions with distinct afferent/efferent connectivity profiles, acted as final waystations, funneling sensory information into the hippocampus proper.

Work from Seltzer and Pandya (1976) confirmed that the parahippocampal area — which projects to entorhinal cortex — receives projections from each of the association cortices of the auditory, visual, and somatosensory modalities, and that the cytoarchitectonic subdivisions of the superior temporal sulcus (STS) possess distinctive profiles of afferents from auditory, visual, and somatosensory associational cortices (Seltzer and Pandya 1994). Also multisensory, but perhaps to a lesser extent, the inferior parietal cortices were found to receive inputs from the visual and somatosensory cortices (Seltzer and Pandya 1980). Later, making use of a neuropathological signature of Alzheimer's disease — the presence of neurofibrillary tangles — and of the fact that Alzheimer's disease pathology can be highly selective and regional, Van Hoesen and his colleagues would make an effort to demonstrate equivalent patterns of connectivity in the human brain (Hyman et al., 1984; Hyman et al., 1987; Van Hoesen and Damasio 1987; Van Hoesen et al., 1991). For example, both the input and output layers of the entorhinal cortex were heavily compromised by tangles, but

not the other layers, in effect disconnecting the sensory cortices from the hippocampus proper (Fig 2).

From a perspective of comparative morphology, the development of higher cognitive functions was enabled by the progressive addition of new association cortices and new levels of convergence. Simpler architectural designs, such as direct connections between primary sensory and limbic regions, only provide a relatively narrow behavioral repertoire. Additional, intermediary regions of association – modality-specific association cortices and then higher-order multimodal association cortices – provide greater behavioral flexibility, complexity, and abstraction, as suggested in prior theoretical work that in part inspired these anatomical studies (Geschwind 1965a,b; reviewed by Catani & ffytche 2005) (Fig. 3).

An attempt to investigate comparable connectional patterns in humans, makes use of a form of noninvasive tractography, diffusion tensor imaging (DTI), based on magnetic resonance scanning. A recent study employing DTI demonstrated convergence of auditory, visual, and somatosensory white-matter fibers in the temporo-parietal cortices (Bonner et al., 2013) (Fig. 4).

The convergent bottom-up projections are reciprocated by top-down projections that cascade from medial temporal lobe structures to multisensory association cortices, and on to parasensory association cortices and then finally to early sensory cortices. The pattern of divergent projections from sensory association cortices back to primary sensory cortices has been confirmed for the auditory modality (Pandya 1995), and for the visual (Rockland and Pandya 1979) and somatosensory modalities (Vogt and Pandya 1978). The back-projections exhibit a distinctive laminar profile, targeting the superficial layers of the primary sensory cortices (reviewed in Meyer 2011). More recent work has also revealed sparse direct projections between the primary and secondary sensory cortices of, for example, audition and vision (Falchier et al., 2002; Rockland and Ojima 2003; Clavagnier et al., 2004; Cappe and Barone 2005; Hackett et al., 2007; Falchier et al., 2010; and reviewed by Falchier et al., 2012) (Fig. 5). We will address the significance of these findings later in the article. We also note that we have focused exclusively on neuroanatomical findings in cerebral cortex. We must, of necessity, omit discussion of the rich pathways of subcortical convergence and cortical-subcortical interactions (reviewed in several chapters of (Stein, ed., 2012)), although, for an excellent example of the latter, see (Jiang et al., 2001).

2. Integrating multisensory information in the mammalian cerebral cortex

2.1. Anatomico-functional framework

The work cited so far established the following facts: (i) sensory pathways exhibit successive levels of convergence, from the early sensory cortices to sensory-specific association cortices and to multisensory association cortices, culminating in maximally integrative regions such as in medial temporal lobe cortices and both lateral and medial prefrontal cortices; and (ii) the convergence of sensory pathways is reciprocated by successive levels of divergence, from the maximally integrative areas to the multisensory association cortices, to the sensory-specific association cortices, and finally to the early sensory cortices.

In an attempt to bridge these anatomical facts and the evidence provided by a variety of approaches to the study of brain and mind-behavior relationships, a number of anatomico-functional frameworks were proposed, for example by Damasio (1989a; 1989b; Meyer and Damasio 2009), and by Mesulam (1998) (Fig. 6). Briefly, the Damasio framework proposes an *architecture* of convergence-divergence zones (CDZ) and a *mechanism* of time-locked retroactivation. Convergence-divergence zones are arranged in a multi-level hierarchy, with

higher-level CDZs being both sensitive to, and capable of reinstating, specific patterns of activity in lower-level CDZs. Successive levels of CDZs are tuned to detect increasingly complex features. Each more-complex feature is defined by the conjunction and configuration of multiple less-complex features detected by the preceding level. CDZs at the highest levels of the hierarchy achieve the highest level of semantic and contextual integration, across all sensory modalities. At the foundations of the hierarchy lie the early sensory cortices, each containing a mapped (i.e., retinotopic, tonotopic, or somatotopic) representation of sensory space. When a CDZ is activated by an input pattern that resembles the template for which it has been tuned, it retro-activates the template pattern of lower-level CDZs. This continues down the hierarchy of CDZs, resulting in an ensemble of well-specified and time-locked activity extending to the early sensory cortices. The mid- and high-level CDZs that span multiple sensory modalities share much in common with Mesulam's (1998) account of "transmodal nodes", as well as with "hubs", or nodes with high centrality, from a graph-theoretic approach (Bullmore and Sporns, 2009).

The overall framework allows for regions in which a strict processing hierarchy is not maintained. Certain sub-sectors may have particularly strong internal connections, both "vertical" and "horizontal", forming somewhat independent functional complexes. Two examples are the dorsal and ventral pathways, whose internal connectivities are related with, respectively, visually guided action (Kravitz et al., 2011) and the representation of object qualities (Kravitz et al., 2013). Contrary to other models that posit nonspecific, or modulatory, feedback mechanisms, time-locked retroactivation provides a mechanism for the global reconstruction of specific neural states. In the next section we review recent findings that are compatible and possibly supportive of this neuroarchitectural framework.

2.2. Investigating sensory convergence and divergence in functional neuroimaging experiments from humans

In this section, we turn to evidence from functional neuroimaging studies in humans. Specifically, we focus on the integration of multisensory information in the creation of representations of the objects of perception. These representations are content-rich, in that they contain cognitively and behaviorally relevant information *about* the stimuli. The information is abstracted across different sensory modalities. Recent work from our laboratory has provided evidence for content-rich multisensory information integration both at the level of the early, sensory-specific cortices, which, we argue, rely on divergent projections, and at the level of the multisensory association cortices, which depend on convergence.

Our studies were conducted using functional magnetic resonance imaging (fMRI), which noninvasively measures blood-oxygenation level dependent signals in the brain as a proxy measure of neural activity. This generates three-dimensional images of neural activity with good spatial resolution (around two cubic mm voxels), permitting the fractionation of discrete brain regions (e.g., primary visual cortex) into many independently measured voxels. By performing multivariate (or multivoxel) pattern analysis (MVPA) of these human functional brain images, the representational content may be decoded from distributed patterns of activity (Mur et al., 2009). In a common implementation of MVPA, machine-learning algorithms are first trained to recognize the association between given classes of stimuli and certain spatial patterns of brain activity. Next, the algorithm is tested by having it assign class labels to new sets of data based on the recognition of diagnostic spatial patterns learned from the training set. We used this technique to make a strong case that stimuli of one modality could orchestrate content-specific activity in the early sensory cortices of another modality. Visual stimuli in the form of silent *but sound-implying* movies were found to evoke content-specific representations in auditory cortices (Meyer et al., 2010). For

example, silent video clips of a violin being played and of a dog barking were reliably distinguished based *solely* on activity in the anatomically defined primary auditory cortices. The nine individual objects as well as the three semantic categories to which they belonged (“animals”, “musical instruments”, and “objects”) were successfully decoded. Intriguingly, the subjects’ ratings of the vividness of the video-cued auditory imagery significantly correlated with decoding performance (Fig. 7).

Extending this approach to the somatosensory cortices, Meyer et al. (2011) demonstrated that visual stimuli could also orchestrate content-specific activity in somatosensory cortices. Touch-implicating videos that depicted haptic exploration of common objects, such as rubbing a skein of yarn or handling a set of keys, were shown to subjects. These video stimuli were then successfully decoded from anatomically defined primary somatosensory cortices.

We next reasoned that information from the visual stimuli would likely have travelled up the convergence hierarchy and then back down into a different sensory sector, either the auditory or the somatosensory. Over the course of that journey, information specifying the modality-invariant *identity* of the stimulus would have been abstracted from the raw energy pattern transduced by the sensory organ. The recovery of this modality-invariant information would have allowed a higher-order convergence zone to specify the manner in which to retro-activate a cortex of a different modality. Once again, performing machine learning analyses of fMRI data, we searched for modality-invariant neural representations across audition and vision (Man et al., 2012). We tested the hypothesis that the brain contains representations of objects that are similar both upon seeing the object and hearing the object. Subjects were shown video clips and audio clips that depicted various common objects. We then performed a *crossmodal* classification, by training an algorithm to distinguish between sound clips and then testing it to decode video clips (and vice versa). Out of several *a priori* defined “multisensory” regions of interest that became active both to audio and video stimuli, only a region near the posterior superior temporal sulcus was found to contain content-specific and modality-invariant representations. A whole-brain searchlight analysis confirmed that the pSTS uniquely and consistently contained object representations that were invariant across vision and audition (Fig. 8).

Earlier fMRI studies on the extraction of semantic information across vision and audition typically followed one of two designs: (i) a comparison of brain activations for congruent vs. incongruent audiovisual stimuli, or (ii) an investigation of crossmodal carry-over effects when an auditory stimulus was followed by a congruent or incongruent visual stimulus (or vice versa). A comprehensive review of these studies (Doehrmann and Naumer 2008) identified a common pattern: regions in the lateral temporal cortices, including pSTS and STG, are more activated by audiovisual semantic congruency, whereas regions in the inferior frontal cortices are more activated by audiovisual semantic *incongruency*. The authors reasoned that the lateral temporal cortices organize stable multisensory object representations, whereas the inferior frontal cortices operate for the more cognitively demanding incongruent stimuli. However, somewhat challenging to this pattern was the finding of Taylor and colleagues (2006) that perirhinal cortex, but not the pSTS, is sensitive to audiovisual semantic congruency. Doehrmann and Naumer (2008) suggest that these and other discrepancies may be due to variation in the stimuli used across studies.

The importance of stimulus control was emphasized in a review of crossmodal object identification by Amedi and colleagues (2005). There is an important distinction between “naturally” and “arbitrarily” associated crossmodal stimuli. In the former, the dynamics of sight and sound inhere within the same source object. The facial movements that accompany speech sounds are a good example of such naturally crossmodal stimuli. In the latter, sight and sound are associated by convention, such as found when speech sounds accompany

printed words. One study highlighting this distinction found elevated functional connectivity between the voice selective areas of auditory cortex and the fusiform face area, for voice/face stimuli but not for voice/printed name stimuli (von Kriegstein and Giraud, 2006). However, it remains unclear whether the elevated functional connectivity was necessarily mediated by an intervening convergence area — as it was during the training phase of the experiment — or if the face and voice areas synchronized their activity through direct connections (as suggested by (Amedi et al., 2005)).

We suggest that content-rich crossmodal patterns of activation require the involvement of supramodal convergence zones. While direct projections between sensory cortices have been identified and possibly have a behavioral function regarding stimuli in the periphery of the visual field (Falchier et al., 2002; Falchier et al., 2010), we assume that they are too sparse to specify the content-specific patterns of activity we have observed. Direct connections between the *primary* sensory cortices of different modalities seem to be even sparser (Borra and Rockland, 2011). Rather, direct connections are likely involved in sub-threshold modulation of ongoing activity (Lakatos et al., 2007) or response latency (Wang et al., 2008; reviewed in Falchier et al., 2012). For the crossmodal orchestration of content-specific representations across multiple brain voxels, we submit that the involvement of convergence zones would be a more plausible account than that of direct connections alone.

Our findings are relevant to the debate over the neural localization of semantic congruency effects. While a modality-invariant representation in pSTS implies crossmodal semantic congruency, the converse is not true: demonstrating a crossmodal semantic congruency effect does not necessarily imply the existence of a modality-invariant representation. Given that our study identified pSTS as the unique location of audiovisual invariant representations, we suggest that it may be the source of the semantic congruency signal. Upon presentation of a bimodal stimulus, pSTS would search for a match in its store of audiovisual invariant representations, and, upon succeeding or failing to find a match, would announce its verdict of “congruent” or “incongruent” downstream to the other regions of the brain that also show semantic congruency effects (Man et al., 2012).

In the following sections we review additional evidence for multisensory integration of information in the audiovisual, visuotactile, and audiotactile domains, from the perspectives of bottom-up convergence and top-down divergence.

2.3. Bottom-up crossmodal integration of information

It has long been known that certain regions and neurons of the mammalian cerebral cortex are multisensory, a fact based on the identification of neurons responsive to auditory, visual, and tactile stimuli (Jung et al., 1963). It is reasonable to assume that nearly all neocortex exhibits some multisensory activity (Calvert, Spence, and Stein, 2004; Ghazanfar and Shroeder, 2006; Driver and Noesselt, 2008). However, multisensory activations and modulations permit only a relatively weak and nonspecific characterization of multisensory processing (Kayser and Logothetis, 2007). Our focus here is on multisensory processing that *retains information about the stimulus*, or multisensory integration of information.

At the highest levels of sensory convergence, single neurons in, for example, the medial temporal lobes can selectively respond to the identities of certain people, whether they are seen or heard (Quian Quiroga et al., 2009). A single neuron robustly responded to the voice, printed name, and various pictures of Oprah Winfrey, but not of any other person tested (with the exception of minor responses to Whoopi Goldberg). Furthermore, the progression of anatomical sensory convergence was recapitulated in the progressive increase in the proportion of cells showing visual invariance at each stage (Fig. 9). From parahippocampal

cortex, to entorhinal cortex, to hippocampus, an increasing proportion of cells in each region were selectively activated by pictures of a particular person.

In a review of human neuroimaging studies of audiovisual and visuotactile crossmodal object identification, three multisensory convergence regions stood out prominently: the posterior superior temporal sulcus (STS) and superior temporal gyrus (STG), for audiovisual convergence, and the lateral occipital tactile-visual area (LOtv) as well as the intra-parietal sulcus (IPS) for visuotactile convergence (Amedi et al., 2005). As noted earlier, each of these multisensory regions are bi-directionally connected to the medial temporal cortices.

Audiovisual integration of information, to the level of semantic congruency, was demonstrated in an electrophysiology study in the superior temporal cortex of the macaque. Dahl and colleagues (2010) recorded single and multi-unit activity during the presentation of audiovisual scenes. Congruent scenes (e.g., corresponding video and audio tracks of a conspecific's grunt) were more reliably decoded than incongruent scenes (e.g., the audio of a grunt paired with the video of a cage door).

Convergence of information across the visual and tactile modalities was found in an fMRI study involving the perception of abstract clay objects. Lateral and middle occipital cortices responded more actively to the visual presentation of objects that had previously been touched, than to objects that had not yet been touched (James et al., 2002). There is evidence that the lateral occipital cortices may play a greater role in visuotactile integration for familiar objects, whereas the intraparietal sulcus would be more involved with unfamiliar objects, during which spatial imagery would be recruited (Lacey et al., 2009; but see also Zhang et al., 2004, in which the level of activity in LOC correlated with the vividness of visual imagery induced by tactile exploration). Another study decoded object category related-response patterns across sight and touch in the ventral temporal cortices (Pietrini et al., 2004). A recent study of visuotactile semantic congruency varied the order in which the modalities were presented and found greater responses in lateral occipital cortex, fusiform gyrus, and intraparietal sulcus when visual stimuli were followed by congruent haptic stimuli as opposed to the other way around, leading the authors to conclude that vision may predominate over touch in those regions (Kassuba et al., 2013).

Crossmodal studies of supramodal object representations across hearing and touch are as yet few, although audiotactile semantic congruency effects have been reported in the pSTS and the fusiform gyrus (Kassuba et al., 2012).

2.4. Top-down crossmodal information integration

There is a vast literature on the modulatory effects of arousal, attention, and imagery on the sensory cortices, which are presumably mediated top-down, from higher-order cortices toward earlier sensory cortices. In this section, once again, we focus only on studies demonstrating *informative* retro-activations of sensory cortices. In other words, we concentrate on activations that carry content-specific information from areas of sensory convergence back to early sensory areas.

2.4.1. Decoding retro-activations in early visual cortices—Studies of visual imagery, in particular, provide rich evidence for internally generated, content-specific activity extending into early visual cortices. These activations are content-specific to the degree that they permit distinctions between different imagined forms and between different form locations. Thirion and colleagues (2006) built an inverse model of retinotopic cortex that was able to reconstruct both the forms of visual imagery (simple geometric patterns) and their locations (in the left or right visual field) from fMRI data (Fig. 10).

This leads naturally to the question of whether visual imagery and visual perception share a code. Is the same visual cortical representation evoked when seeing something and when imagining the same thing? Thirion et al. (2006) found that classification across perception and imagery was indeed successful, although only in a minority of their subjects.

Commonality of representation could be inferred by training a classification algorithm to distinguish among seen objects and then testing it to decode imagined objects, or vice versa. Successful performance indicates generalization of learning across the classes of “seen” and “imagined”. Following this approach, several groups have found abundant evidence for common coding across visual perception and imagery. The decoded stimuli include: “X”’s and “O”’s in lateral occipital cortices (Stokes et al., 2009); the object categories “tools”, “food”, “famous faces”, and “buildings” in ventral temporal cortex (Reddy et al., 2010); the categories “objects”, “scenes”, “body parts”, and “faces” in their respective category-selective visual regions (e.g. faces in fusiform face area), as well as the location of the stimulus (in the left or right visual fields) in V1, V2, and V3 (Cichy et al., 2012). Taken together, these studies suggest that visual imagery is a top-down process that reconstructs and uses, to some extent, the patterns of activity that were established during veridical perception. Remarkably, these reconstructions appear to extend all the way to the initial site of cortical visual processing, V1.

The studies of (Stokes et al., 2009, Reddy et al., 2010, and Cichy et al., 2012) used auditory cues can trigger visual imagery. But whether auditory cues are sufficient to trigger content-specific activity in visual cortices without necessarily triggering visual imagery is an important follow-up question. After establishing that various natural sounds could be decoded in V2 and V3, Vetter and colleagues (2011) constrained visual imagery with an orthogonal working memory task. Subjects heard natural sounds while memorizing word lists and performing a delayed match-to-sample task. Despite these constraints on visual imagery, the sounds were still successfully decoded in early visual cortices.

Perceptual expectation is another top-down cognitive process that modulates V1 in a content-specific manner. An fMRI study (Kok et al., 2012) used auditory tones to cue subjects to *expect* that a particular visual stimulus would be shown immediately afterward. High or low tones were associated with right- or left-oriented contrast gratings, respectively. This prediction of visual orientation by auditory tone was valid in 75% of trials (expected condition); the remaining trials violated the prediction (unexpected condition). Performing MVPA in V1, the expected visual stimuli were decoded more accurately than the unexpected ones. (In V2 and V3, however, expectation had no effect on decoding accuracy.) Interestingly, the gross level of activity in V1 was lower in the expected condition than in the unexpected condition. These findings support an account of expectation as producing a *sharpening* effect: the consistency and informativeness of representation increases, even as overall level of activity declines (Fig. 11). This sharpening effect was present even when the expectation was not task-relevant — when subjects performed a contrast judgment task that was unrelated to stimulus orientation — showing it to be independent of task-related attention.

2.4.2. Decoding retro-activations in early auditory cortices—We noted earlier that we decoded sound-implicating visual stimuli from primary auditory cortices (Meyer et al., 2010). However, visual stimuli that did not imply sounds failed to be decoded from early auditory cortices (Hsieh et al., 2012). Other findings from auditory cortex resemble those from visual cortex: there is substantial multisensory information conveyed back to relatively “early” sensory cortices (Schroeder and Foxe 2005 *Curr Op Neurobiol*; but see also Kayser et al., 2009 *Hearing Res*). Kayser and colleagues (2010 *Curr Biol*) recorded local field potentials and spiking activity from monkey auditory cortex during the presentation of various naturalistic stimuli. Compared to presenting sounds alone, the presentation of sounds

paired with congruent videos resulted in a more consistent and informative pattern of neural activity across each presentation of a particular stimulus. This consistency was evident in both firing rates and spike timings. For incongruent pairings of sounds and videos, this consistency across stimulus presentations was reduced.

Another example of higher-level semantic processing influencing crossmodal activations was given by an fMRI study by Calvert et al. (1997; reviewed in Meyer 2011). Primary auditory cortex activity was observed when subjects watched a person silently mouthing numbers but not when they watched nonlinguistic facial movements that did not imply sounds. More recently, an fMRI study of the McGurk effect (Benoit et al., 2010; reviewed in Meyer 2011) showed a stimulus-specific response to visual input in auditory cortices. McGurk and MacDonald (1976) reported a perceptual phenomenon in which the auditory presentation of the syllable /ba/, combined with the visual presentation of the mouth movement for /ga/, resulted in the percept of a third syllable, /da/. Benoit and colleagues (2010), performing fMRI in primary auditory cortices, exploited this effect in a repetition suppression paradigm. They presented a train of three audiovisual congruent /pa/ syllables, followed by the presentation of either a fourth congruent /pa/ or an incongruent McGurk stimulus (visual /ka/, auditory /pa/). Although the auditory stimulus was the same /pa/ syllable in both conditions, the co-presentation of an incongruent /ka/ mouth movement evoked greater activity (release from adaptation) in A1. This demonstrates that primary auditory cortices can respond differently to the same auditory stimulus when it is paired with different visual stimuli.

The question of whether auditory cortices receive content-specific tactile information also deserves attention. Multi-electrode recordings in macaque have shown that tactile input may modulate A1 excitability by resetting phase oscillatory activity (Lakatos et al., 2007), potentially affecting subsequent auditory processing in a stimulus-specific manner. A psychophysical study in humans found that auditory stimuli interfered with a tactile frequency discrimination task only when the frequencies of the auditory and tactile stimuli were similar (Yau et al., 2009). This effect is likely mediated by the caudomedial or caudolateral belt regions of auditory cortex (Fuxe 2009). A high-field fMRI study of the auditory cortex belt area in macaque showed greater enhancement of activity when a sound stimulus was synchronously paired with a tactile stimulus, compared to pairing with an asynchronous tactile stimulus (Kayser et al., 2005). This shows the temporal selectivity of secondary auditory cortex for tactile stimuli. On the other hand, also in macaque, Lemus and colleagues (2010) performed single neuron recordings in A1 and found that it could not distinguish between two tactile flutter stimuli.

2.4.3. Decoding retro-activations in early somatosensory cortices—In the study cited above, Lemus and colleagues (2010) also found that primary somatosensory cortices failed to distinguish between two acoustic flutter stimuli. However, in an fMRI study in humans, Etzel and colleagues (2008) successfully decoded sounds made by either the hands or the mouth from activity patterns in left S1 and bilateral S2. With respect to content-specific activations by visual stimuli, a univariate fMRI study found selectivity for visually presented shapes over visually presented textures in precentral sulcus (Stilla and Sathian 2008), and the multivariate fMRI study of Meyer et al. (2011), cited earlier, decoded the identities of visual objects from somatosensory cortices.

2.4.4. Decoding actions across sensory and motor systems—The discovery of neurons that activate both during action observation and action execution in macaque premotor cortices (Gallese et al., 1996) has led to proposals that they participate in a mirror neuron system that underlies action recognition and comprehension. A pertinent question is whether these activations constitute action-specific representations, and, if so, whether they

represent a given action similarly across observation and execution. Single cell recordings in ventral premotor cortices were indeed found to distinguish between two actions, whether they were seen, heard, or performed (Keysers et al 2003).

In human fMRI studies the evidence for common coding of actions has been somewhat mixed. Adaptation studies test the hypothesis that observing an action after executing the same action (or vice versa) would result in reduced activity compared to observing and executing different actions. Evidence supporting this hypothesis has been found for certain goal-directed motor acts (Kilner et al., 2009) and yet not for other intransitive, non-meaningful motor acts (Lingnau et al., 2009). Evidence from MVPA studies has been similarly mixed (reviewed by Oosterhof et al., 2012). Actions were successfully decoded across hearing and performing them (Etzel et al., 2008) or across seeing and performing them (Oosterhof et al., 2010), however another group reported failed cross-classification between seeing and performing an action (Dinstein et al., 2008). The likely source of disagreement among these research findings, echoing the debate over the localization of audiovisual semantic congruency effects, is the variation in the stimuli employed. However, it seems safe to conclude that there is at least *some* commonality of representation between action observation and execution, with room for disagreement on the degree of fidelity between representations.

3. From percepts to concepts

The evidence reviewed above establishes that (i) the multisensory association cortices extract information regarding the congruence or identity of stimuli across modalities in a content-specific manner; and (ii) the early sensory cortices contain content-specific information regarding heteromodal stimuli (that is, stimuli of a modality different from the sensory cortex surveyed). We suggest that these findings are consonant with a neuroarchitectural framework of convergence-divergence zones, in which content-rich information is integrated across the different sensory modalities, as well as broadcasted back to them.

The systematic linkage of sensory information has long been seen as playing a role in conceptual processes by philosophers (Hume 1739; Prinz 2002) as well as by cognitive neuroscientists (Damasio 1989c; Barsalou 1999; Martin 2007). This position, known as concept empiricism, holds that abstract thought and explicit sensory representation share an intimate and obligate link. For example, the concept of a “bell” is determined by its sensory associations: it is composed by the sound of its ding-dong, the sight of its curved and flared profile, the feel of its cold and rigid surface. Even highly abstract concepts, such as “truth” or “disjunction”, may be decomposed into sensory primitives (Barsalou 1999). Concept empiricism has enjoyed the accumulation of evidence from several lines of inquiry. Studies of experimental neuroanatomy in the final quarter of the 20th century established the existence of pathways of sensory convergence and divergence in the mammalian cerebral cortex. Largely as a result of studies performed in the past decade, these neuroanatomical pathways were then found, by electrophysiological and functional neuroimaging methods, to carry content-rich information both up to integrating centers and back down to sensory cortices. The present review focused on audition, vision, and touch, owing to the relative dearth of studies for the other modalities. However, concepts also make abundant reference to the modalities of taste, smell, vestibular sensation, and interoception, and additional studies are needed on the integration among these modalities. Moreover, while bi-modally invariant representations (across, e.g. vision and audition or vision and touch) have been found, the identification of tri-modally invariant representations would buttress the argument that they participate in conceptual thought, and would establish a yet higher level atop the hierarchy of convergence-divergence.

We hope that the human evidence discussed in this article can lead to further in-depth exploration of the intricate connectivity of the mammalian cerebral cortex using experimental neuroanatomy.

Acknowledgments

Grant sponsor:

This research was supported by the Mathers Foundation and by the National Institutes of Health (NIH/NINDS 5P50NS019632).

References

- Amedi A, Von Kriegstein K, Van Atteveldt NM, Beauchamp MS, Naumer MJ. Functional imaging of human crossmodal identification and object recognition. *Exp Brain Res*. 2005; 166:559–571. [PubMed: 16028028]
- Barsalou LW. Perceptual symbol systems. *Behav Brain Sci*. 1999; 22:577–660. [PubMed: 11301525]
- Borra E, Rockland KS. Projections to early visual areas V1 and V2 in the calcarine fissure from parietal association areas in the macaque. *Front Neuroanat*. 2011; 5:1–12. [PubMed: 21373368]
- Benoit MM, Raji T, Lin F-H, Jääskeläinen IP, Stufflebeam S. Primary and multisensory cortical activity is correlated with audiovisual percepts. *Hum Brain Mapp*. 2010; 31:526–538. [PubMed: 19780040]
- Bonner MF, Peelle JE, Cook PA, Grossman M. Heteromodal conceptual processing in the angular gyrus. *NeuroImage*. 2013; 71:175–186. [PubMed: 23333416]
- Bullmore E, Sporns O. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci*. 2009; 10:186–198. [PubMed: 19190637]
- Calvert GA. Activation of Auditory Cortex During Silent Lipreading. *Science*. 1997; 276:593–596. [PubMed: 9110978]
- Calvert, GA.; Spence, C.; Stein, BE., editors. *The handbook of multisensory processes*. MIT Press; 2004.
- Cappe C, Barone P. Heteromodal connections supporting multisensory integration at low levels of cortical processing in the monkey. *Eur J Neurosci*. 2005; 22:2886–2902. [PubMed: 16324124]
- Catani M, ffytche DH. The rises and falls of disconnection syndromes. *Brain*. 2005; 128:2224–2239. [PubMed: 16141282]
- Cichy RM, Heinze J, Haynes J-D. Imagery and perception share cortical representations of content and location. *Cereb Cortex*. 2012; 22:372–380. [PubMed: 21666128]
- Clavagnier S, Falchier A, Kennedy H. Long-distance feedback projections to area V1: implications for multisensory integration, spatial awareness, and visual consciousness. *Cogn Affect Behav Neurosci*. 2004; 4:117–126. [PubMed: 15460918]
- Dahl CD, Logothetis NK, Kayser C. Modulation of visual responses in the superior temporal sulcus by audio-visual congruency. *Front Int Neurosci*. 2010; 4:10.
- Damasio A. Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition. *Cognition*. 1989a; 33:25–62. [PubMed: 2691184]
- Damasio A. The brain binds entities and events by multiregional activation from convergence zones. *Neural Comput*. 1989b; 1:123–132.
- Damasio A. Concepts in the brain. *Mind Lang*. 1989c; 4:24–28.
- Dinstein I, Gardner JL, Jazayeri M, Heeger DJ. Executed and observed movements have different distributed representations in human aIPS. *J Neurosci*. 2008; 28:11231–11239. [PubMed: 18971465]
- Doehrmann O, Naumer MJ. Semantics and the multisensory brain: how meaning modulates processes of audio-visual integration. *Brain Res*. 2008; 1242:136–150. [PubMed: 18479672]
- Driver J, Noesselt T. Multisensory interplay reveals crossmodal influences on “sensory-specific” brain regions, neural responses, and judgments. *Neuron*. 2008; 57:11–23. [PubMed: 18184561]

- Etzel JA, Gazzola V, Keysers C. Testing simulation theory with cross-modal multivariate classification of fMRI data. *PLoS One*. 2008; 3:e3690. [PubMed: 18997869]
- Falchier, A.; Cappe, C.; Barone, P.; Schroeder, CE. Sensory convergence in low-level cortices. In: Stein, BE., editor. *The New Handbook of Multisensory Processing*. Cambridge, MA: MIT Press; 2012. p. 67
- Falchier A, Clavagnier S, Barone P, Kennedy H. Anatomical evidence of multimodal integration in primate striate cortex. *J Neurosci*. 2002; 22:5749–5759. [PubMed: 12097528]
- Falchier A, Schroeder CE, Hackett TA, Lakatos P, Nascimento-Silva S, Ulbert I, Karmos G, Smiley JF. Projection from visual areas V2 and prostriata to caudal auditory cortex in the monkey. *Cereb Cortex*. 2010; 20:1529–1538. [PubMed: 19875677]
- Foxe JJ. Multisensory integration: frequency tuning of audio-tactile integration. *Curr Biol*. 2009; 19:R373–R375. [PubMed: 19439261]
- Gallese V, Fadiga L, Fogassi L, Rizzolatti G. Action recognition in the premotor cortex. *Brain*. 1996; 119:593–609. [PubMed: 8800951]
- Geschwind N. Disconnexion syndromes in animals and man. I. *Brain*. 1965a; 88:237–294. [PubMed: 5318481]
- Geschwind N. Disconnexion syndromes in animals and man. II. *Brain*. 1965b; 88:585–644. [PubMed: 5318824]
- Ghazanfar AA, Schroeder CE. Is neocortex essentially multisensory? *Trends Cogn Sci*. 2006; 10:278–285. [PubMed: 16713325]
- Hackett TA, Smiley JF, Ulbert I, Karmos G, Lakatos P, De la Mothe LA, Schroeder CE. Sources of somatosensory input to the caudal belt areas of auditory cortex. *Perception*. 2007; 36:1419–1430. [PubMed: 18265825]
- Hsieh P-J, Colas JT, Kanwisher N. Spatial pattern of BOLD fMRI activation reveals cross-modal information in auditory cortex. *J Neurophysiol*. 2012; 107:3428–3432. [PubMed: 22514287]
- Hume, D. A treatise of human nature. Project Gutenberg, eBook Collection. 1739. <http://www.gutenberg.org/files/4705/4705-h/4705-h.htm>
- Hyman BT, Van Hoesen GW, Damasio A, Barnes CL. Alzheimer's Disease: Cell-Specific Pathology Isolates the Hippocampal Formation. *Science*. 1984; 225:1168–1170. [PubMed: 6474172]
- Hyman BT, Van Hoesen GW, Damasio A. Alzheimer's disease: glutamate depletion in the hippocampal perforant pathway zone. *Ann Neurol*. 1987; 22:37–40. [PubMed: 2443073]
- James TW, Humphrey GK, Gati JS, Servos P, Menon RS, Goodale MA. Haptic study of three-dimensional objects activates extrastriate visual areas. *Neuropsychologia*. 2002; 40:1706–1714. [PubMed: 11992658]
- Jiang W, Wallace MT, Jiang H, Vaughan JW, Stein BE. Two cortical areas mediate multisensory integration in superior colliculus neurons. *J Neurophysiol*. 2001; 85:506–522. [PubMed: 11160489]
- Jones EG, Powell TP. An anatomical study of converging sensory pathways within the cerebral cortex of the monkey. *Brain*. 1970; 93:793–820. [PubMed: 4992433]
- Jung, R.; Kornhuber, H.; Da Fonseca, J. Multisensory convergence on cortical neurons. In: Moruzzi, G.; Fessard, A.; Jasper, HH., editors. *Progress in Brain Research*. New York, NY: Elsevier; 1963. p. 207-240.
- Kassuba T, Klinge C, Hölig C, Röder B, Siebner HR. Vision holds a greater share in visuo-haptic object recognition than touch. *NeuroImage*. 2013; 65:59–68. [PubMed: 23032487]
- Kassuba T, Menz MM, Röder B, Siebner HR. Multisensory Interactions between Auditory and Haptic Object Recognition. *Cereb Cortex*. 2012; 23:1097–1107. [PubMed: 22518017]
- Kayser C, Logothetis NK. Do early sensory cortices integrate cross-modal information? *Brain Struct Func*. 2007; 212:121–132.
- Kayser C, Logothetis NK, Panzeri S. Visual enhancement of the information representation in auditory cortex. *Curr Biol*. 2010; 20:19–24. [PubMed: 20036538]
- Kayser C, Petkov CI, Augath M, Logothetis NK. Integration of touch and sound in auditory cortex. *Neuron*. 2005; 48:373–384. [PubMed: 16242415]

- Kayser C, Petkov CI, Logothetis NK. Multisensory interactions in primate auditory cortex: fMRI and electrophysiology. *Hearing Res.* 2009; 258:80–88.
- Keysers C, Kohler E, Umiltà MA, Nanetti L, Fogassi L, Gallese V. Audiovisual mirror neurons and action recognition. *Exp Brain Res.* 2003; 153:628–636. [PubMed: 12937876]
- Kilner JM, Neal A, Weiskopf N, Friston KJ, Frith CD. Evidence of mirror neurons in human inferior frontal gyrus. *J Neurosci.* 2009; 29:10153–10159. [PubMed: 19675249]
- Kok P, Jehee J, de Lange F. Less Is More: Expectation Sharpens Representations in the Primary Visual Cortex. *Neuron.* 2012; 75:265–270. [PubMed: 22841311]
- Kravitz DJ, Saleem KS, Baker CI, Mishkin M. A new neural framework for visuospatial processing. *Nat Rev Neurosci.* 2011; 12:217–230. [PubMed: 21415848]
- Kravitz DJ, Saleem KS, Baker CI, Ungerleider LG, Mishkin M. The ventral visual pathway : an expanded neural framework for the processing of object quality. *Trends Cog Sci.* 2013; 17:26–49.
- Lacey S, Tal N, Amedi A, Sathian K. A putative model of multisensory object representation. *Brain Topogr.* 2009; 21:269–274. [PubMed: 19330441]
- Lakatos P, Chen C-M, O’Connell MN, Mills A, Schroeder CE. Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron.* 2007; 53:279–292. [PubMed: 17224408]
- Lemus L, Hernández A, Luna R, Zainos A, Romo R. Do sensory cortices process more than one sensory modality during perceptual judgments? *Neuron.* 2010; 67:335–348. [PubMed: 20670839]
- Lingnau A, Gesierich B, Caramazza A. Asymmetric fMRI adaptation reveals no evidence for mirror neurons in humans. *Proc Natl Acad Sci U S A.* 2009; 106:9925–9930. [PubMed: 19497880]
- Man K, Kaplan JT, Damasio A, Meyer K. Sight and sound converge to form modality-invariant representations in temporoparietal cortex. *J Neurosci.* 2012; 32:16629–16636. [PubMed: 23175818]
- Martin A. The representation of object concepts in the brain. *Annu Rev Psychol.* 2007; 58:25–45. [PubMed: 16968210]
- McGurk H, MacDonald J. Hearing lips and seeing voices. *Nature.* 1976; 264:746–748. [PubMed: 1012311]
- Mesulam M. From sensation to cognition. *Brain.* 1998; 121:1013–1052. [PubMed: 9648540]
- Meyer K. Primary sensory cortices, top-down projections and conscious experience. *Prog Neurobiol.* 2011; 94:408–417. [PubMed: 21683755]
- Meyer K, Damasio A. Convergence and divergence in a neural architecture for recognition and memory. *Trends Neurosci.* 2009; 32:376–382. [PubMed: 19520438]
- Meyer K, Kaplan JT, Essex R, Damasio H, Damasio A. Seeing touch is correlated with content-specific activity in primary somatosensory cortex. *Cereb Cortex.* 2011; 21:2113–2121. [PubMed: 21330469]
- Meyer K, Kaplan JT, Essex R, Webber C, Damasio H, Damasio A. Predicting visual stimuli on the basis of activity in auditory cortices. *Nat Neurosci.* 2010; 13:1–26. [PubMed: 20033076]
- Mur M, Bandettini PA, Kriegeskorte N. Revealing representational content with pattern-information fMRI—an introductory guide. *Soc Cogn Affect Neur.* 2009; 4:101–109.
- Oosterhof NN, Tipper SP, Downing PE. Viewpoint (In)dependence of Action Representations: An MVPA Study. *J Cognitive Neurosci.* 2012; 24:975–989.
- Oosterhof NN, Wiggett AJ, Diedrichsen J, Tipper SP, Downing PE. Surface-based information mapping reveals crossmodal vision-action representations in human parietal and occipitotemporal cortex. *J Neurophysiol.* 2010; 104:1077–1089. [PubMed: 20538772]
- Pandya D. Anatomy of the auditory cortex. *Rev Neurol.* 1995; 151:486–494. [PubMed: 8578069]
- Pietrini P, Furey ML, Ricciardi E, Gobbini MI, Wu W-HC, Cohen L, Guazzelli M, Haxby JV. Beyond sensory images: Object-based representation in the human ventral pathway. *Proc Natl Acad Sci U S A.* 2004; 101:5658–5663. [PubMed: 15064396]
- Prinz, J. *Furnishing the Mind: Concepts and Their Perceptual Basis.* MIT Press; 2002.
- Quian Quiroga R, Kraskov A, Koch C, Fried I. Explicit encoding of multimodal percepts by single neurons in the human brain. *Curr Biol.* 2009; 19:1308–1313. [PubMed: 19631538]
- Reddy L, Tsuchiya N, Serre T. Reading the mind’s eye: decoding category information during mental imagery. *NeuroImage.* 2010; 50:818–825. [PubMed: 20004247]

- Rockland KS, Ojima H. Multisensory convergence in calcarine visual areas in macaque monkey. *Int J Psychophysiol.* 2003; 50:19–26. [PubMed: 14511833]
- Rockland KS, Pandya D. Laminar origins and terminations of cortical connections of the occipital lobe in the rhesus monkey. *Brain Res.* 1979; 179:3–20. [PubMed: 116716]
- Schroeder CE, Foxe J. Multisensory contributions to low-level, “unisensory” processing. *Curr Op Neurobiol.* 2005; 15:454–458. [PubMed: 16019202]
- Seltzer B, Pandya D. Some Cortical Projections to the Parahippocampal Area in the Rhesus Monkey. *Exp Neurol.* 1976; 160:146–160. [PubMed: 814012]
- Seltzer B, Pandya D. Converging visual and somatic sensory cortical input to the intraparietal sulcus of the rhesus monkey. *Brain Res.* 1980; 192:339–351. [PubMed: 6769545]
- Seltzer B, Pandya D. Parietal, temporal, and occipital projections to cortex of the superior temporal sulcus in the rhesus monkey: a retrograde tracer study. *J Comp Neurol.* 1994; 343:445–463. [PubMed: 8027452]
- Stein, BE., editor. *The New Handbook of Multisensory Processing.* MIT Press; 2012.
- Stilla R, Sathian K. Selective visuo-haptic processing of shape and texture. *Hum Brain Mapp.* 2008; 29:1123–1138. [PubMed: 17924535]
- Stokes M, Thompson R, Cusack R, Duncan J. Top-down activation of shape-specific population codes in visual cortex during mental imagery. *J Neurosci.* 2009; 29:1565–1572. [PubMed: 19193903]
- Taylor KI, Moss HE, Stamatakis EA, Tyler LK. Binding crossmodal object features in perirhinal cortex. *Proc Natl Acad Sci U S A.* 2006; 103:8239–8244. [PubMed: 16702554]
- Thirion B, Duchesnay E, Hubbard E, Dubois J, Poline J-B, Lebihan D, Dehaene S. Inverse retinotopy: inferring the visual content of images from brain activation patterns. *NeuroImage.* 2006; 33:1104–1116. [PubMed: 17029988]
- Van Hoesen, GW.; Damasio, A. Neural correlates of cognitive impairment in Alzheimer’s disease. In: Mountcastle, VB.; Plum, F.; Geiger, SR., editors. *Handbook of Physiology.* Bethesda, Maryland: American Physiological Society; 1987. p. 871-898.
- Van Hoesen GW, Hyman BT, Damasio A. Entorhinal cortex pathology in Alzheimer’s disease. *Hippocampus.* 1991; 1:1–8. [PubMed: 1669339]
- Van Hoesen GW, Pandya D. Some connections of the entorhinal (area 28) and perirhinal (area 35) cortices of the rhesus monkey I temporal lobe afferents. *Brain Res.* 1975a; 95:1–24. [PubMed: 1156859]
- Van Hoesen GW, Pandya D. Some connections of the entorhinal (area 28) and perirhinal (area 35) cortices of the rhesus monkey III efferent connections. *Brain Res.* 1975b; 95:39–59. [PubMed: 1156868]
- Van Hoesen GW, Pandya DN, Butters N. Cortical afferents to the entorhinal cortex of the Rhesus monkey. *Science.* 1972; 175:1471–1473. [PubMed: 4622430]
- Van Hoesen GW, Pandya DN, Butters N. Some connections of the entorhinal (area 28) and perirhinal (area 35) cortices of the rhesus monkey. II. Frontal lobe afferents. *Brain Res.* 1975; 95:25–38. [PubMed: 1156867]
- Vetter P, Smith FW, Muckli L. Decoding natural sounds in early visual cortex. *J Vis.* 2011; 11:779–779.
- Vogt BA, Pandya D. Cortico-cortical connections of somatic sensory cortex (areas 3, 1 and 2) in the rhesus monkey. *J Comp Neurol.* 1978; 177:179–191. [PubMed: 413844]
- Von Kriegstein K, Giraud A-L. Implicit multisensory associations influence voice recognition. *PLoS Biol.* 2006; 4:e326. [PubMed: 17002519]
- Wang Y, Celebrini S, Trotter Y, Barone P. Visuo-auditory interactions in the primary visual cortex of the behaving monkey: electrophysiological evidence. *BMC Neurosci.* 2008; 9:79. [PubMed: 18699988]
- Yau JM, Olenczak JB, Dammann JF, Bensmaia SJ. Temporal frequency channels are linked across audition and touch. *Curr Biol.* 2009; 19:561–566. [PubMed: 19268591]
- Zhang M, Weisser VD, Stilla R, Prather SC, Sathian K. Multisensory cortical processing of object shape and its relation to mental imagery. *Cogn Aff Behav Neurosci.* 2004; 4:251–259.

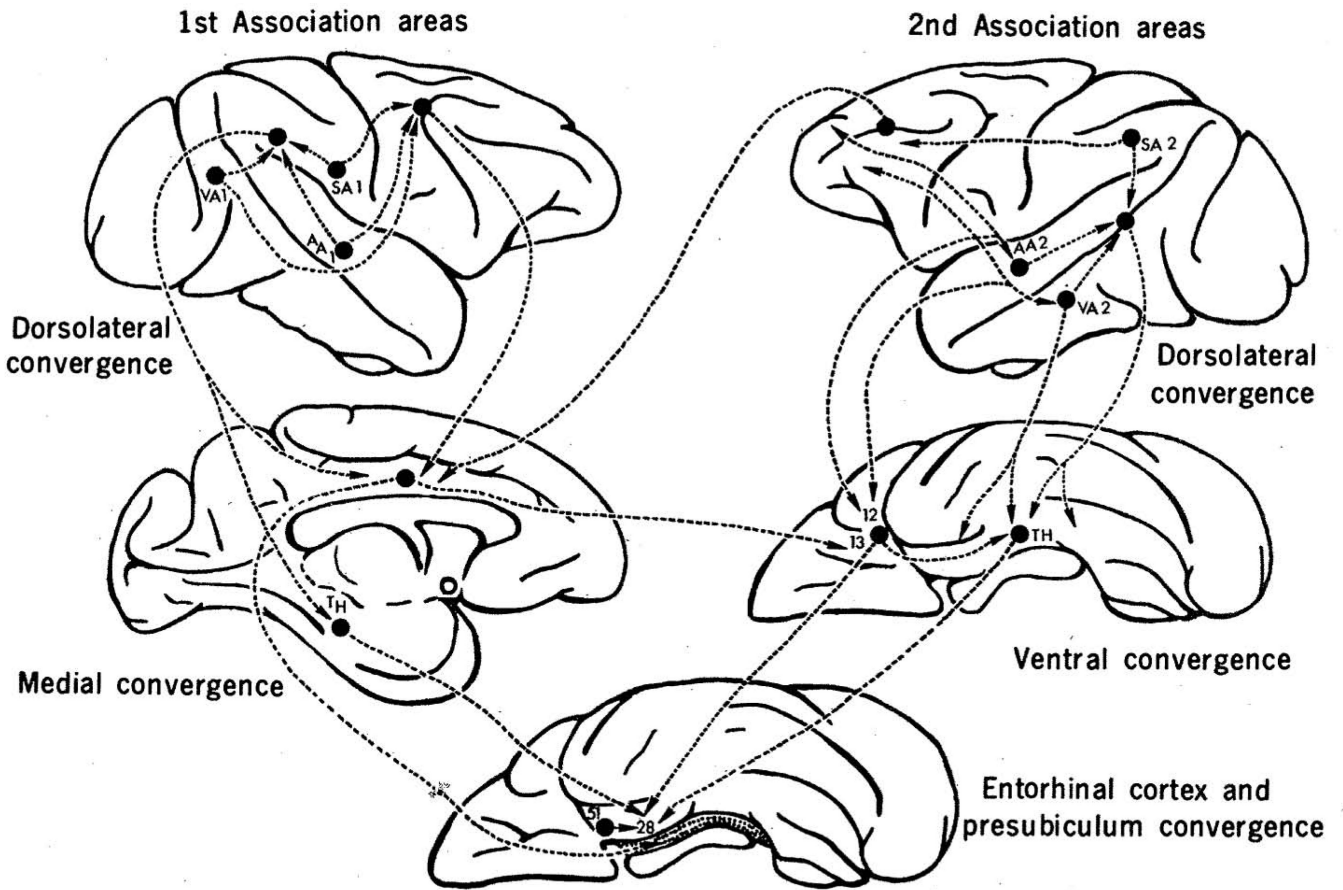


Fig. 1. The multisynaptic pathways by which sensory information converges upon the hippocampus, depicted on the rhesus monkey brain. Primary association areas for visual, auditory, and somatosensory areas are labeled VA1, AA1, and SA1, respectively. Secondary association areas are labeled similarly, as VA2, AA2, and SA2. Reproduced, with permission, from (Van Hoesen et al., 1972).

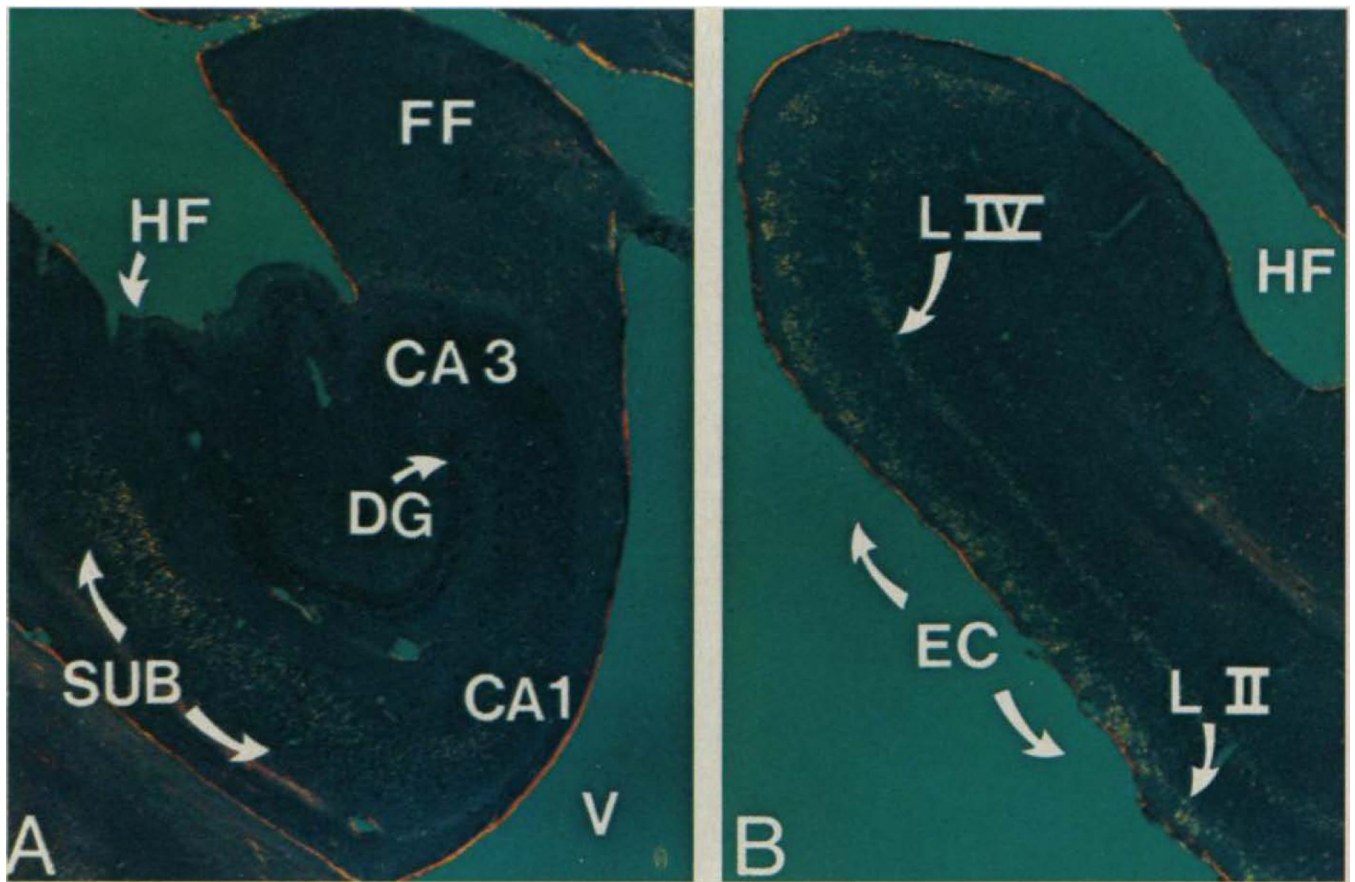


Fig. 2. Distribution of neurofibrillary tangles in the hippocampal formation in Alzheimer's disease. Neurofibrillary tangles appear yellow in these Congo red-stained sections from the brain of a human patient with Alzheimer's disease. Tangles are selectively apparent in (A) subicular CA1 fields and in (B) layers II and IV of entorhinal cortex. Reproduced, with permission, from (Hyman et al 1984).

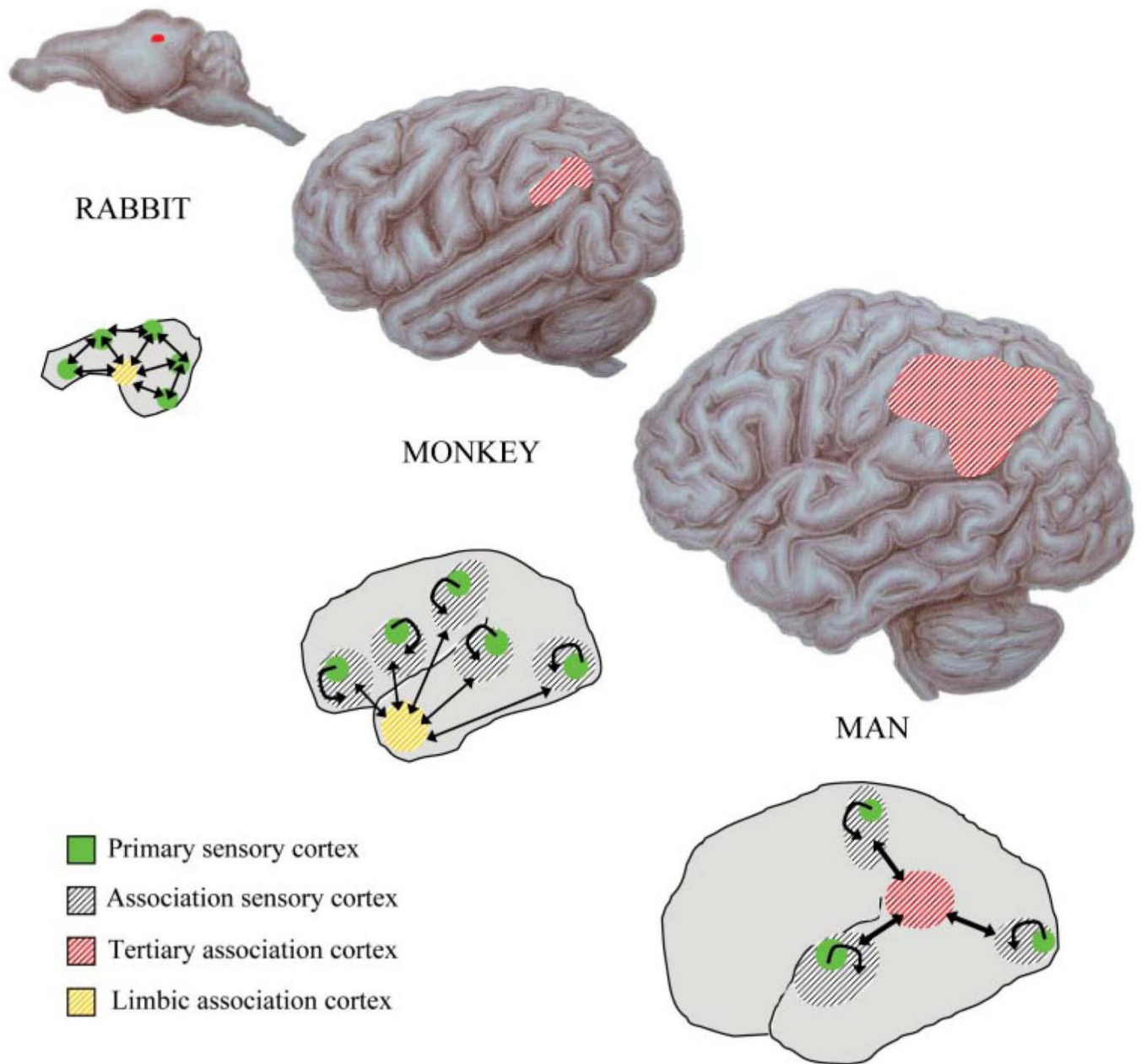


Fig. 3. Phylogenetic expansion of higher-order association cortices and increased complexity in sensory convergence architectures. The top sequence shows the expansion of inferior parietal cortex across species. The bottom sequence shows the variation in connectivity patterns across species, with cross-sensory interactions mediated by higher levels of association cortices in man. Reproduced, with permission, from (Catani & ffytche 2005).

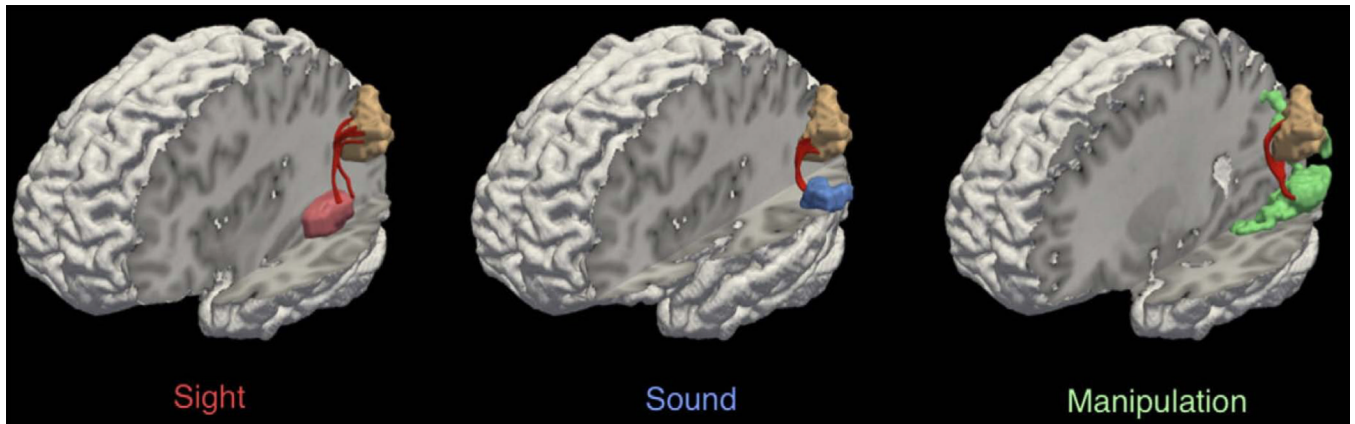


Fig. 4. DTI of a multisensory association region. Tractography of a multisensory activation cluster in the temporo-parietal cortices (angular gyrus, tan) with regions responsive to sight-, sound-, and manipulation-implying words. Reproduced, with permission, from (Bonner et al., 2013).

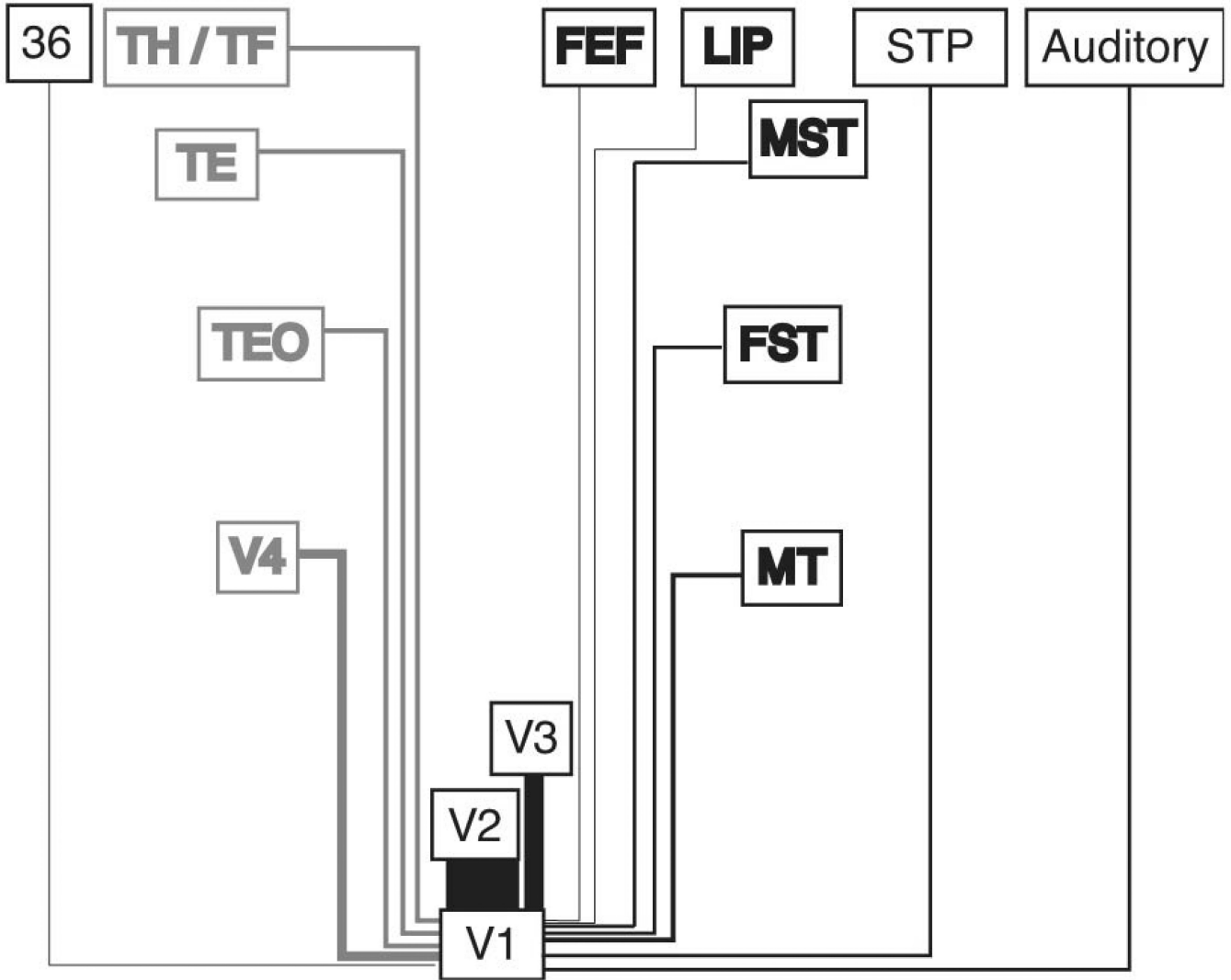


Fig. 5. Direct feedback connections to V1. A retrograde tracer study revealed long-distance feedback projections to V1 from auditory cortices, multisensory cortices, and a perirhinal area. Lines in black represent the dorsal stream; lines in gray, the ventral stream. The thickness of lines represents the strength of connection. Reproduced, with permission, from (Clavagnier et al. 2004).

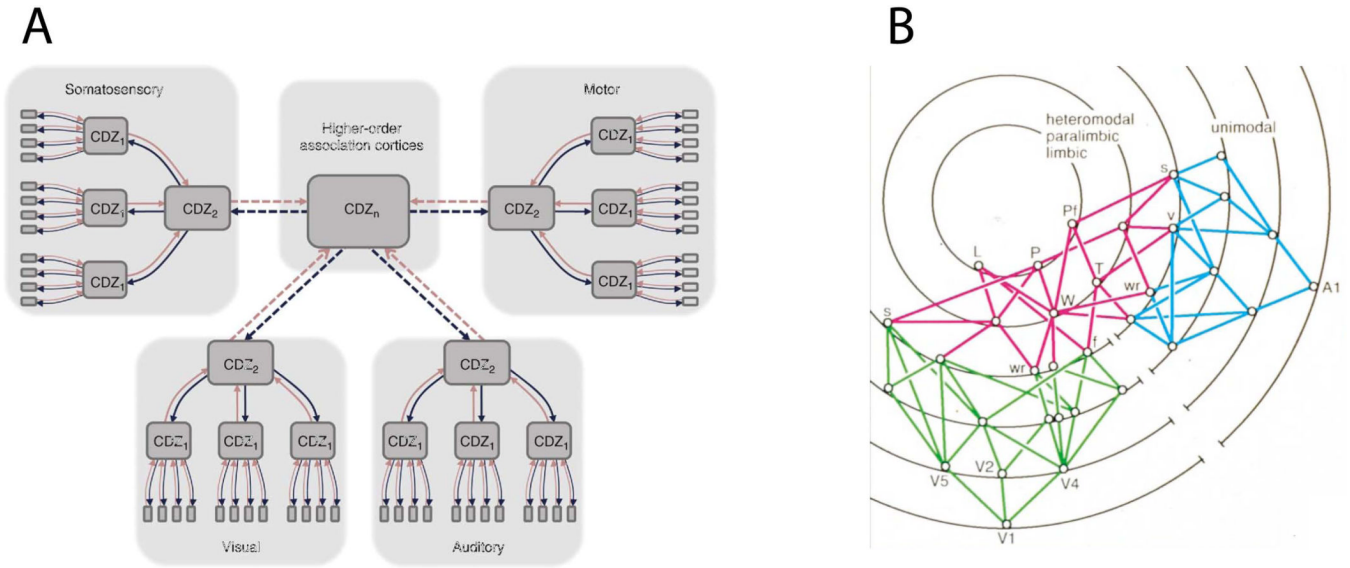


Fig. 6. Large-scale neural frameworks of convergence and divergence. A) A schematic illustration of the convergence-divergence zone framework. Red lines indicate bottom-up connections, blue lines, top-down. From (Meyer & Damasio 2009). B) An illustration of transmodal nodes, in red, connecting visual regions, in green, with auditory regions, in blue. Each concentric ring represents a different synaptic level. Reproduced, with permission, from (Mesulam 1998).

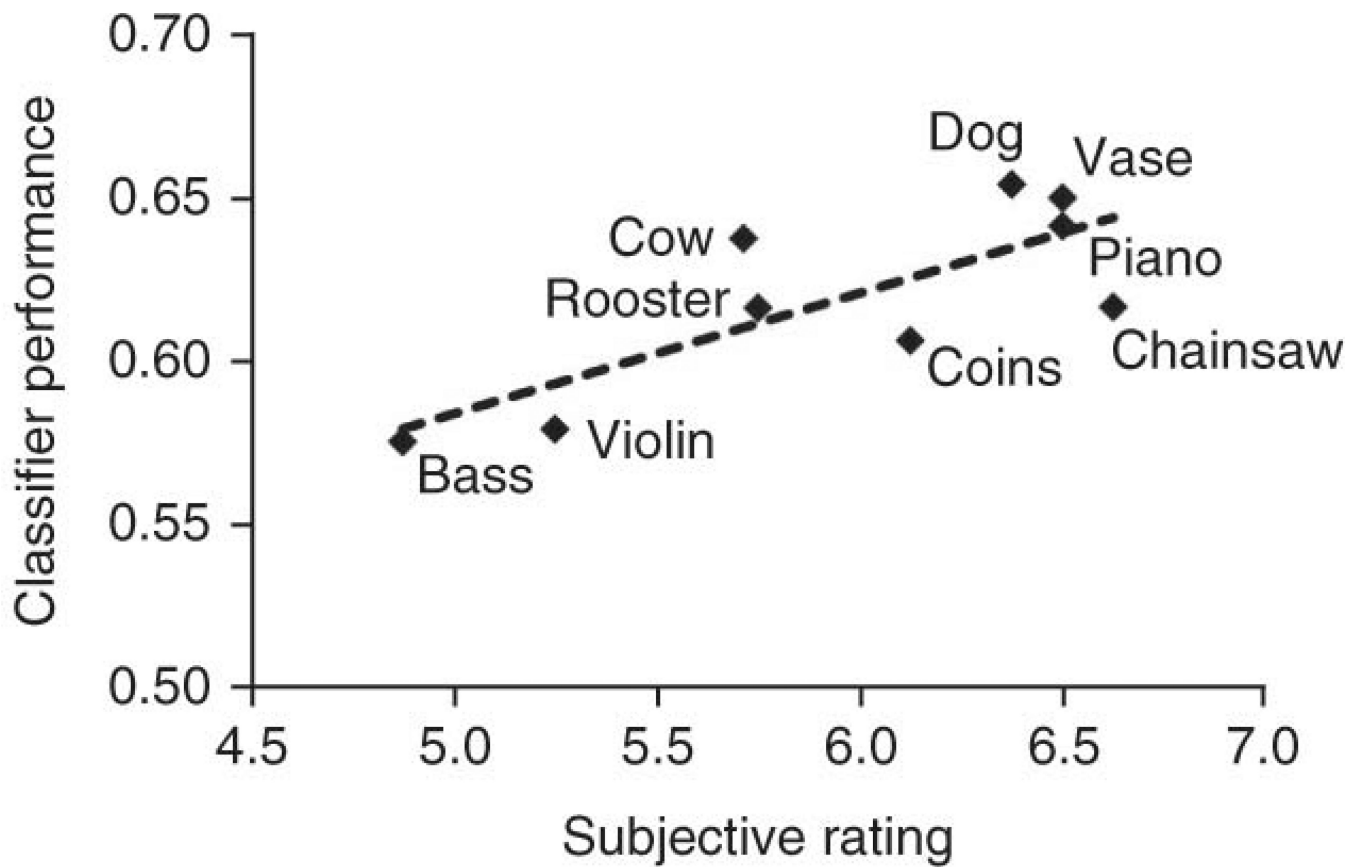


Fig. 7. The decoding accuracy of silent videos from auditory cortex activity correlated with subjects' ratings of the vividness of auditory imagery. Reproduced, with permission, from (Meyer et al. 2010).

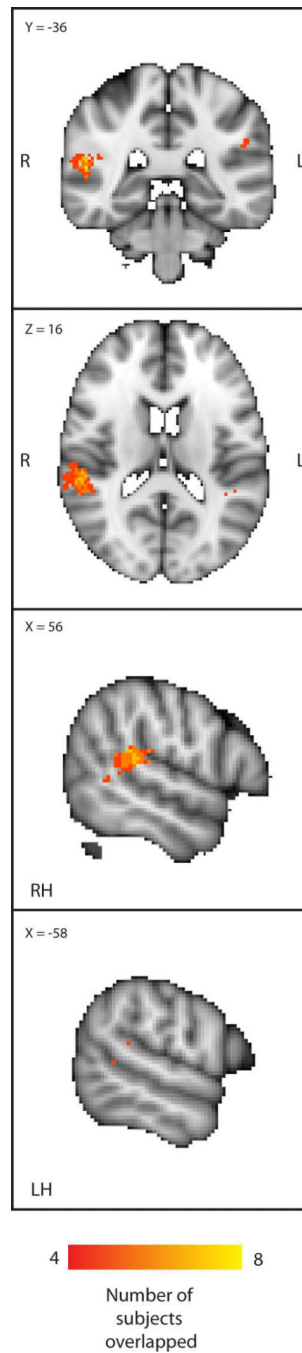


Fig. 8. Audiovisual-invariant representations were most consistently found in the temporoparietal cortices, near the pSTS. Reproduced, with permission, from (Man et al. 2012).

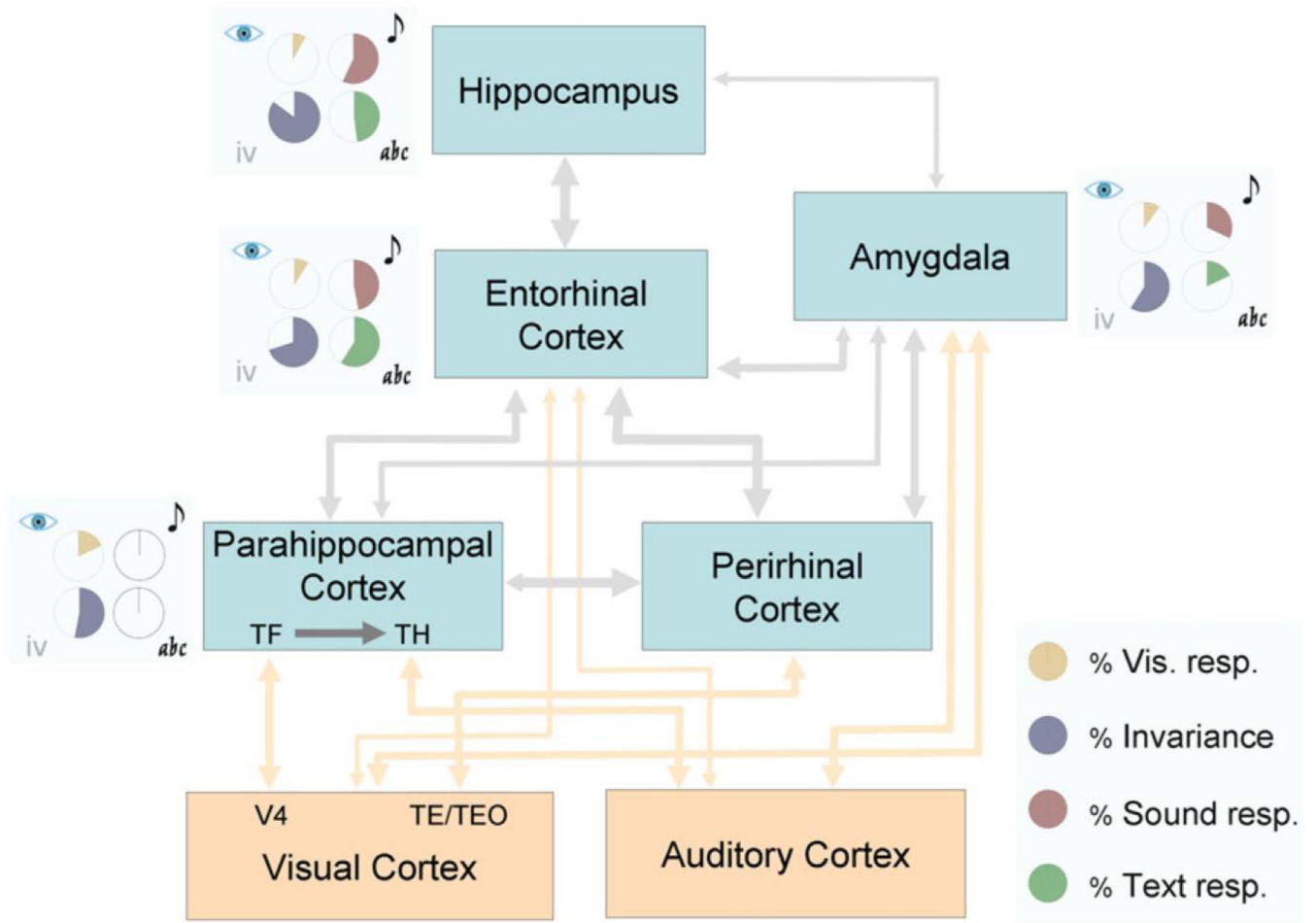


Fig. 9. The proportion of visually-responding neurons that were selective for a particular person across several different pictures, shown as the portion of the circles filled with purple, increased in the progression from parahippocampal cortex to entorhinal cortex and finally to hippocampus. Reproduced, with permission, from (Quian Quiroga et al. 2009).

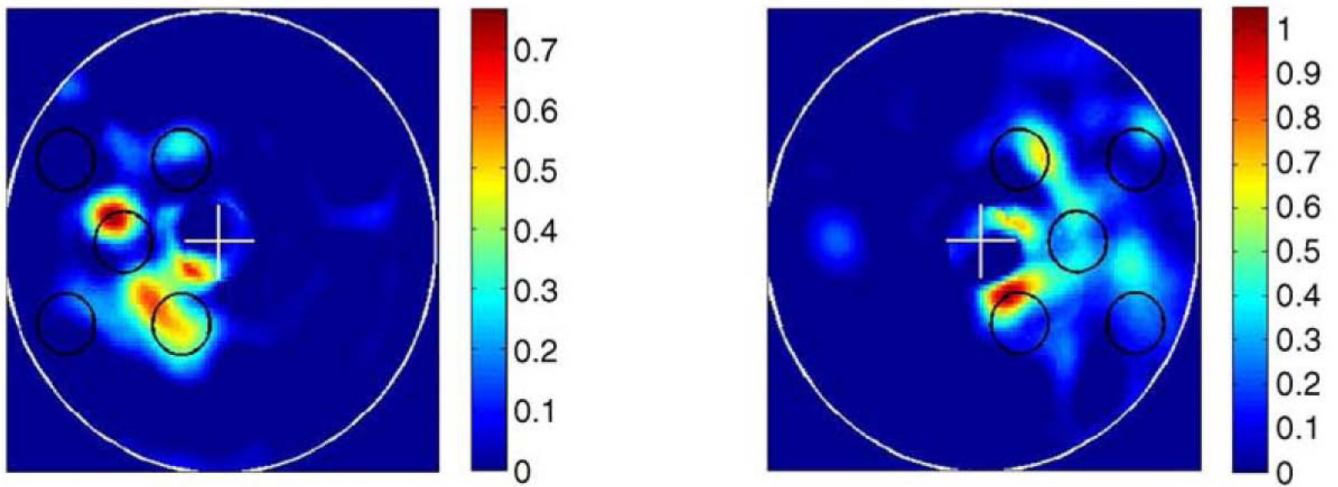


Figure 10.

There is a rough likeness between the target visual pattern, shown in black circles, and the pattern reconstructed from brain activations in early visual cortex during imagery of the target pattern, shown as colored blobs. Reproduced, with permission, from (Thirion et al. 2006).

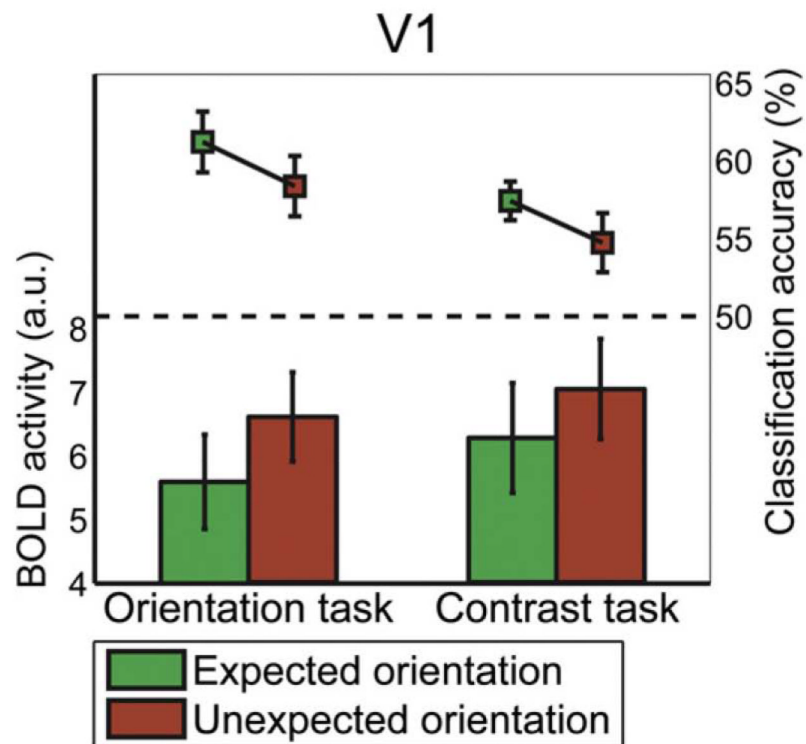
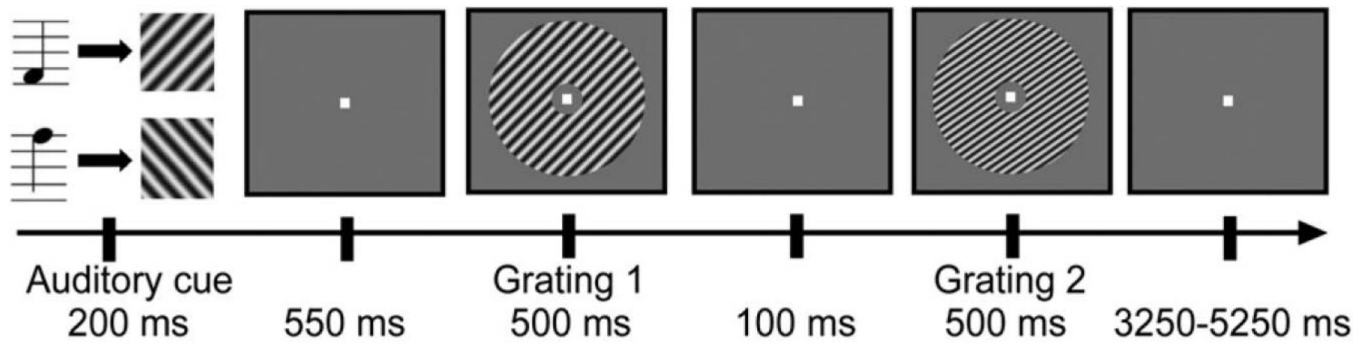


Figure 11. The expectation of a certain visual stimulus resulted in reduced BOLD activity, but higher decoding accuracy for that stimulus, in V1. Adapted and reproduced, with permission, from (Kok et al. 2012).