Development of visual category selectivity in ventral visual cortex does not require visual experience

Job van den Hurk[a,b,1], Marc Van Baelen[a,c], and Hans P. Op de Beeck[a]

[a]Laboratory of Biological Psychology, Faculty of Psychology and Educational Sciences, Katholieke Universiteit Leuven, 3000 Leuven, Belgium; [b]Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, 6200 MD Maastricht, The Netherlands; and [c]Department of Engineering Management, University of Antwerp, 2000 Antwerp, Belgium

To what extent does functional brain organization rely on sensory input? Here, we show that for the penultimate visual-processing region, ventral-temporal cortex (VTC), visual experience is not the origin of its fundamental organizational property, category selectivity. In the fMRI study reported here, we presented 14 congenitally blind participants with face-, body-, scene-, and object-related natural sounds and presented 20 healthy controls with both auditory and visual stimuli from these categories. Using macro-anatomical alignment, response mapping, and surface-based multi-voxel pattern analysis, we demonstrated that VTC in blind individuals shows robust discriminatory responses elicited by the four categories and that these patterns of activity in blind subjects could successfully predict the visual categories in sighted controls. These findings were confirmed in a subset of blind participants born without eyes and thus deprived from all light perception since conception. The sounds also could be decoded in primary visual and primary auditory cortex, but these regions did not sustain generalization across modalities. Surprisingly, although not as strong as visual responses, selectivity for auditory stimulation in visual cortex was stronger in blind individuals than in controls. The opposite was observed in primary auditory cortex. Overall, we demonstrated a striking similarity in the cortical response layout of VTC in blind individuals and sighted controls, demonstrating that the overall category-selective map in extrastriate cortex develops independently from visual experience.

blindness | ventral-temporal cortex | category perception | functional MRI | pattern analysis

Category selectivity in human and primate visual cortex is a striking example of how the brain encodes the outer world. Mostly found on the lateral and ventral parts of the temporal lobe, several distinct macroscopic brain regions are known to have a preference for a particular category of visual objects, including faces [fusiform face area, occipital face area (1)], body parts [extrastriate body area (2), fusiform body area (3)], artificial objects [lateral occipital complex (4)], and scenes [parahippocampal place area (5)]. The exact computational mechanisms that drive these regions are still largely unknown: Do these regions operate as distinct functional modules (1, 6), or is category selectivity fully distributed across ventral-temporal cortex (VTC) (7), or—a combination of both—do category-selective regions reflect peaks in a broad tuning map for visual object features (8)?

A prominent question is whether this functional architecture develops independently of visual input or relies on visual experience during early infancy and the following years. As an example of the latter, a key hypothesis indicates that the global functional organization of VTC starts out as a protomap with a retinotopic layout (9–11). Through visual experience, regions within this protomap develop selectivity for the visual categories that often appear in the retinotopic region that is represented. According to this perspective, because faces most likely appear in the fovea, the neural representation of faces develops in the foveal regions of the protomap. In this scenario, the development of higher-level visual processing relies heavily on the shaping of cortical regions through a relatively long period of visual stimulation.

The alternative hypothesis, that the selectivity for behaviorally relevant categories in VTC requires no visual input to develop, is challenging to study scientifically. Our visual system is bombarded by a barrage of visual stimuli from the moment of birth, continuously shaping the highly adaptive visual cortex. A recent imaging study has demonstrated that the large-scale category selectivity of ventral visual regions is already present in 6-mo-old infants (12). Furthermore, it has been suggested that infants show a slow but solid electrophysiological response to face stimuli at posterior temporal sites as early as 3 mo of age (13), but these findings do not rule out the possibility that early visual experience already has shaped these cortical representations.

Arguably the only approach that allows a reliable dissociation between the two hypotheses regarding the functional organization of VTC is to study individuals that have had no visual experience at all: the congenitally blind. If the organization of this part of the brain indeed depends on visual experience, congenitally blind individuals will not show selectivity for face-, body-, object-, and scene categories in VTC. Interestingly, some evidence points toward a large-scale categorical sensitivity and layout of ventral visual cortex in this population. For instance, He and colleagues (14) observed small-vs.-large object selectivity in parahippocampal regions. Moreover, the medial-to-lateral bias that is associated with processing the animate–inanimate distinction is found in blind individuals (15). In addition, studies in blind subjects have demonstrated

Significance

The brain's ability to recognize visual categories is guided by category-selective ventral-temporal cortex (VTC). Whether visual experience is required for the functional organization of VTC into distinct functional subregions remains unknown, hampering our understanding of the mechanisms that drive category recognition. Here, we demonstrate that VTC in individuals who were blind since birth shows robust discriminatory responses to natural sounds representing different categories (faces, scenes, body parts, and objects). These activity patterns in the blind also could predict successfully which category was visually perceived by controls. The functional cortical layout in blind individuals showed remarkable similarity to the well-documented layout observed in sighted controls, suggesting that visual functional brain organization does not rely on visual input.

that the visual word form area responds to both braille-reading (16) and letter soundscapes created with a visual-to-auditory sensory substitution device (17), indicating that this region does not require visual experience for its functional relevance in reading tasks. On a more specific level, some studies have investigated categorical representations in VTC of blind individuals directly. Pietrini and colleagues (18) used tactile stimulation to show that object and face representations in ventral visual cortex are present in the blind, suggesting that these representations are based on abstract features rather than on strictly visual features. Moreover, Kitada and colleagues (19) demonstrated that haptic recognition of basic facial expressions in persons blind since early childhood (hereafter "the early blind") involves the same frontal and temporal regions as in sighted controls during a visual expression-recognition task. Furthermore, the extrastriate body area in the early blind showed a stronger response to haptic identification of hands than to man-made objects (20). These findings suggest that visual experience might not be necessary for the appearance of face and body representations in the human "visual" system.

Still, this body of evidence has at least three limitations. First, these studies have not, or at most have only partially, focused upon the well-known macroscopic functional layout of VTC. To what extent do the cortical locations and boundaries of the face-, body-, place-, and object-sensitive regions in VTC in blind individuals match those regions in the brains of sighted individuals? Previous work has shown that the cortical locations of some categorical representations partially overlap in the brains of blind and sighted persons; here we quantify the relative strength of these representations in the two subject groups and the degree of spatial overlap in multiple categories. Second, the majority of studies have investigated the activation of visual regions by linguistic sounds: mostly auditory words, and in one case braille and learned letter soundscapes [with the notable exception of haptic stimulation (18–20)]. To understand the problem inherent in this limitation, consider trying to study the fusiform face area by presenting written names of people rather than their faces. If we want to investigate the activation of these regions in sensory processing across senses, the investigation should be done primarily with sensory stimuli, not linguistic stimuli. The third and final problem with the existing studies is that they have not fully ruled out the presence of visual input early in life. The typical definition of congenital blindness includes people who have no record of having had patterned vision, but in many so-called "congenitally blind," it is impossible to rule out very limited patterned vision early in life. Eye diseases involved in congenital blindness can be progressive in nature, so vision loss in these persons might be incomplete at birth. Even when patterned vision can be excluded, many patients have light perception, and most still exhibit prenatal retinal waves of spontaneous activity which are known to be crucial for the development of many properties of the visual system before the retina becomes light sensitive (21). It has been demonstrated in vivo in mice that these retinal waves can also modulate extrastriate neural activity (22). If findings are averaged across all blind participants in a study, one or several of these factors could be enough to present evidence for functional selectivity in visual cortex even if the development of such selectivity would critically depend upon visual input or other aspects of retinal function. Studies typically do not have the data quality in individual subjects to restrict at least part of the analyses to the minority (if any) of blind subjects in whom the presence of early rudimentary vision can be completely ruled out. Only patients with anophthalmia—people born without eyes—can satisfy this criterion. In anophthalmia, there is no ocular development, or ocular development is terminated at a very early prenatal stage (23). As a consequence, anophthalmic individuals have no (or only very limited) prenatal endogenous retinal activity and lack all postnatal visual experience. Functional selectivity in visual cortex should be demonstrated in such patients to have certainty about the exact role of (rudimentary) vision in establishing selectivity maps in visual cortex. The important distinction between anophthalmia and other forms of congenital and early blindness has already been put forward by others investigating the effects of visual deprivation on early auditory information processing and language networks (24–26).

We designed this study to overcome the three problems mentioned above. We directly compare the neural responses in VTC of sighted (n = 20) and blind (n = 14) individuals to face-, body-, scene-, and (man-made) object-related sounds. In each blind participant we obtained twice as much data as would normally be sufficient to show functional selectivity for visual stimuli in sighted subjects convincingly. As such, we could also perform targeted analyses on individual subjects, most critically on the three blind participants with anophthalmia. Our main hypothesis is that the categorical selectivity and functional organization of VTC can develop independently of visual experience. We divide this hypothesis into two subsidiary hypotheses: (i) The four categories elicit distinct responses in VTC of our blind volunteers, and (ii) the functional architecture of VTC in the blind shows significant overlap with the layout of this brain region in the sighted controls. Finding confirming evidence for the first contention would demonstrate that the categorical boundaries of our stimuli are picked up and processed by neurons in ventral visual cortex. The second complementary hypothesis provides insights regarding the similarity of the cortical-activation profiles of VTC in blind and sighted individuals. We used multivoxel pattern analysis (MVPA) and functional cortical mapping techniques to help resolve the debate regarding category selectivity in the extrastriate visual cortex.

Results

Univariate Analysis: Contrasts and Selectivity. We first tested the fMRI contrasts between each category individually against the average of the other three categories across subjects. The results are mapped on the average cortex reconstruction in Fig. 1, *Left* for auditory and visual stimuli in the sighted controls and auditory stimuli in blind participants separately. The visually presented trials resulted in solid responses in VTC. On the left and right hemispheres, face-selective vertices are consistently located laterally to the midfusiform sulcus (MFS) (27). More medially, scene-selective vertices are found surrounding the collateral sulcus. The scene- and face-selective patches border an object-selective patch lying in the MFS. On the left hemisphere another object-sensitive region was found on the occipitotemporal sulcus bordering the face-sensitive vertices, whereas the right hemisphere shows a body region ventrally to the face-selective patch. Compared with the visual trials, the auditory trials in the controls yielded little to no significant vertices that survived the statistical threshold. However, the blind participants showed a solid response to face-related sounds on the fusiform gyrus bilaterally. In addition, a few object-sensitive clusters appear, most prominently in the left occipitotemporal sulcus.

The selectivity maps that are depicted in Fig. 1, *Right* show the unthresholded preferred stimulus condition after normalization for global response amplitude. In the visual modality, the selectivity maps closely resemble the statistical contrast maps. However, the auditory trials in the sighted controls yield a less clear outline of functional boundaries. The correlation between the auditory and visual maps was found to be reasonable (r = 0.1775, P < 0.001), probably in part because of the face- and object-selective patches in the left hemisphere. Interestingly, the auditory selectivity maps of the blind individuals show a clear resemblance to the visual maps in controls (thresholded as well as unthresholded). Most pronounced are the face- and scene-selective patches bordering the MFS and the object-selective region near the occipitotemporal sulcus, especially in the left hemisphere. The correlation between the blind map and the visual control map is substantial (r = 0.3415, P < 0.001). We also computed the correlation between the visual control map and the blind map when the latter is based on only the first four runs, resulting in a correlation of r = 0.2671, P < 0.001.
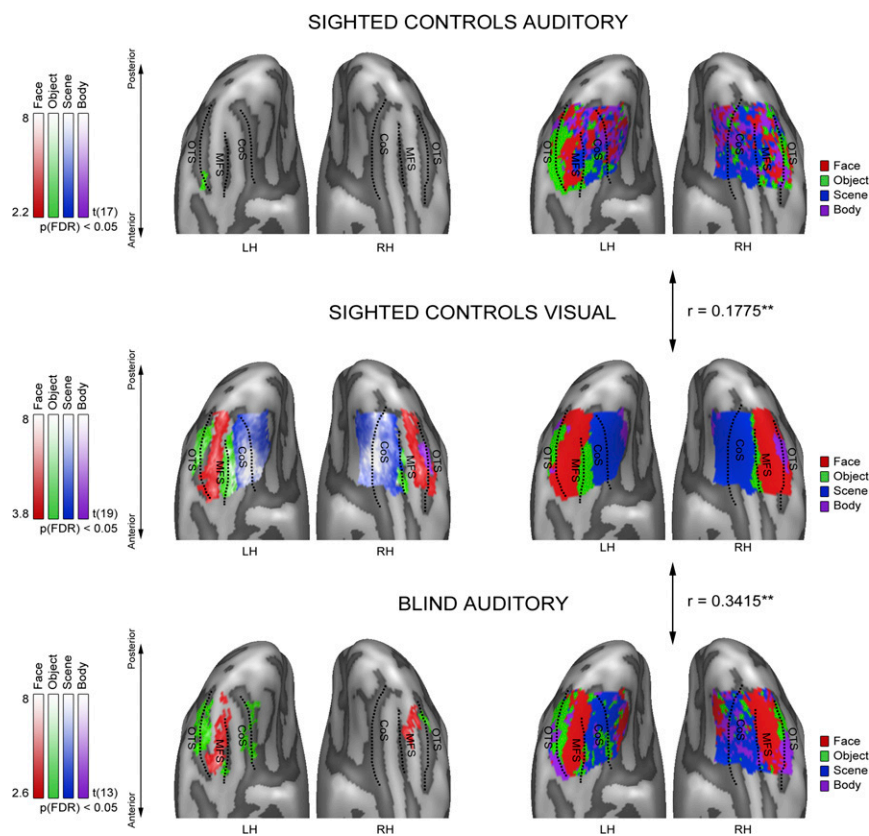
Fig. 1. Statistical contrasts and selectivity mapping in controls and blind subjects. (*Left*) Statistical comparisons (*t* test) between each of the four categories and the average response of the remaining three categories, separately for auditory trials in controls (*Top*), visual trials in controls (*Middle*), and auditory trials in the blind (*Bottom*). For each vertex, the condition that yielded the highest *t* statistic was mapped. Maps were thresholded using an FDR correction, maintaining $q = 0.05$. (*Right*) The (unthresholded) topographical sensitivity maps for the controls and the blind participants. These maps visualize the functional topography of VTC to the four categories. The average visual map was correlated with the average auditory map of the control group ($r = 0.1775$) and of the blind group ($r = 0.3415$). CoS, collateral sulcus; MFS, midfusiform sulcus; OTS, occipito-temporal sulcus. $**P < 0.0001$.

Classification Results: VTC. The average accuracies of MVPA predictions across subjects that resulted from the classification procedure in VTC are shown in Fig. 2. As can be seen, all category pairs can be decoded from neural responses in VTC significantly better than empirical chance level in blind subjects. The classification performances of the auditory and visual trials measured in the sighted controls allow direct comparisons between the two groups of subjects. Unsurprisingly, the visual presentation of the categories yielded nearly perfect results in the control group. The auditory runs resulted in a lower performance, but the classifier still achieved a significant decoding in all but one (body vs. object) category pair.

We also compared the classification performances of the auditory trials in the blind individuals and the controls, averaging across all pairwise comparisons. The prediction accuracies in the blind subjects were substantially higher when all eight runs were included in the classification, yielding a significant difference between the blind and control individuals (permutation test; $P < 0.001$). When we reduced the data to four runs in both the blind

participants and the controls, this difference became smaller but was still significant with a two-tailed test ($P < 0.05$).

We isolated the three category 1 (anophthalmic) blind subjects (for categories of blind subjects, see *Experimental Procedures*) and assessed the accuracy with which the classification algorithm could decode the categories from their neural responses (Fig. 3). Because these subjects are the only participants in our study whom we are certain never had any visual experience, even during a brief period early in life, these results serve as a benchmark for our full pool of blind subjects. Interestingly, the analysis resulted in a robust decoding accuracy in all three anophthalmic subjects.

Classification Results: Primary Visual Cortex and Primary Auditory Cortex. Because primary visual regions serve basic visual processing in sighted individuals, we wondered whether primary visual cortex (V1) shows a discriminatory response to the four auditory conditions in the blind group. Fig. 4 shows the classification results, indicating that V1 indeed discriminates between the face, body, object, and scene categories. Comparing these
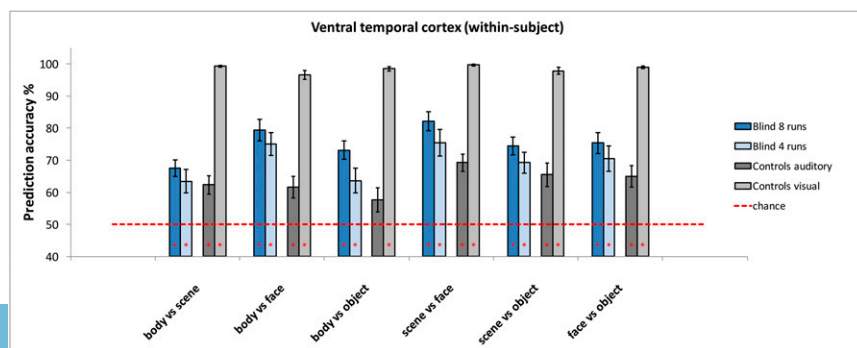

Fig. 2. Classification results for blind subjects and sighted controls in VTC. For all six condition pairs, the average prediction accuracies across subjects are depicted. For a fair comparison between sighted and blind participants, we also report the prediction accuracies when only the first four runs in the blind subjects are used for classification. Error bars represent the SEM. The red dashed line indicates chance level (50%) and is not used for assessment of statistical significance. An asterisk (*) in a bar reflects a significant decoding accuracy compared with the empirically estimated chance level. $*P < 0.05$, FDR-corrected at $q = 0.05$.
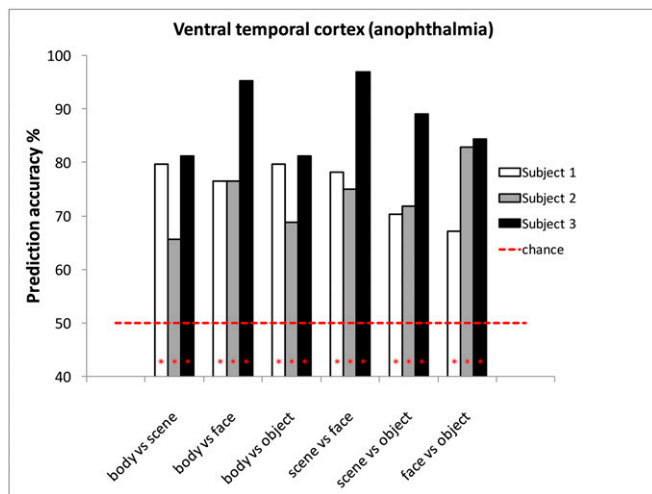
van den Hurk et al.

Fig. 3. Classification results in VTC for three anophthalmic subjects. The red dashed line indicates chance level (50%) and is not used for assessment of statistical significance. An asterisk (*) in a bar reflects a significant decoding accuracy compared with the empirically estimated chance level. *P < 0.05, FDR-corrected at q = 0.05.

results with those in the sighted controls, we found that the auditory trials in the control group yield no statistically significant decoding in any category pair in V1. The direct comparison between the participant groups averaged across categories approaches significance (P = 0.07, two-tailed) when only the first four runs of the blind subjects are considered. However, the visual trials in the controls result in a robust prediction in all category pairs, indicating that the four visual conditions most likely differ in low-level visual features.

In primary auditory cortex (A1), the auditory conditions could be discriminated in both subject groups. Here there was a significantly stronger discrimination in the control group than in the blind group across category pairs (P < 0.001). In addition, we
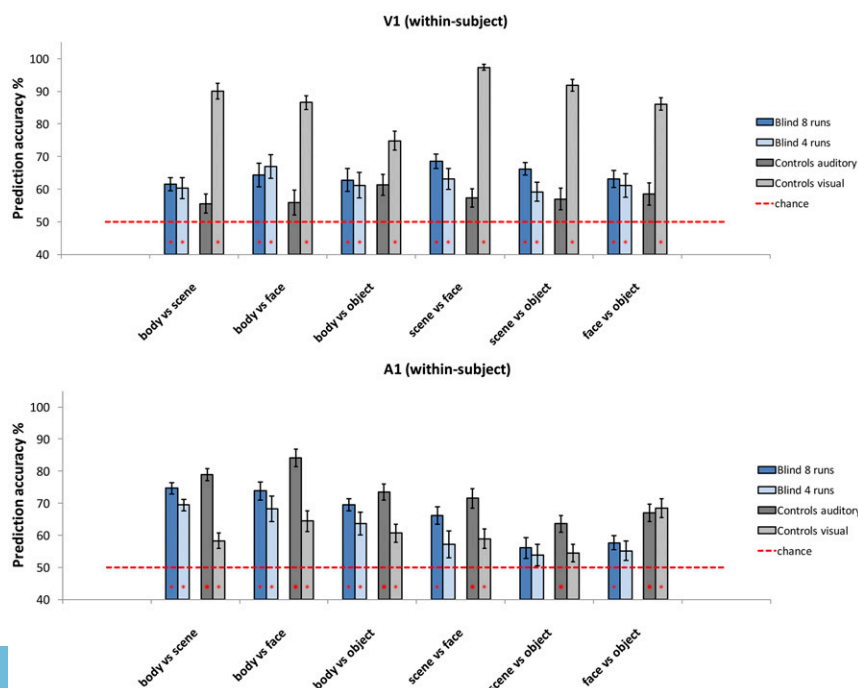
observed significant decoding of the visual categories in A1 in the control group (P < 0.01). When looking at the individual category pairs, we find all but one pair (scene vs. object) to be statistically decodable from response patterns in A1.

We also examined how the neural responses in VTC of the three anophthalmic blind participants differed across categories. Again, we demonstrate that early visual cortex shows a discriminatory response in almost all category pairs in these individuals, see Fig. 5.

Between-Subject Classification Results: VTC. Because a successful decoding in VTC of blind participants does not show whether the same neural populations are at work as in sighted controls, we attempted to cross-decode between blind and sighted individuals. More specifically, we asked if training a classifier on the auditory trials of the blind individuals would allow generalization to the visual and auditory trials in the sighted controls. As can be seen in Fig. 6, the visual-response patterns could be predicted successfully from the auditory responses in the blind participants (P < 0.01). When inspecting the category pairs, we found that all pairs except body vs. object could be decoded using this approach on the visual data. When we tried to predict the auditory trials in the control group, we found the same global outcome, albeit with lower prediction accuracies. Despite the relatively low performance accuracies, the results still are significantly better than our permutation results.

When we invert the generalization direction (Fig. 6) to training on the controls and predicting the classes in the blind subjects, we qualitatively find the same results as when we generalized to the controls, albeit with only marginal prediction accuracies. The asymmetry in decoding direction is significant (P < 0.001). This phenomenon of asymmetric generalization has been reported in other studies (28–31), and the exact source remains unclear (32). One possible interpretation that has been put forward is that it reflects the underlying nature of the representations; for instance, the auditory trials might activate only a subset of voxels that contain categorical information, whereas the visual trials activate the majority of VTC voxels. Therefore, training the classifier on the sparse auditory data ensures that the algorithm captures the relevant information that generalizes well to the well-defined visual

Fig. 4. Classification results for blind subjects and sighted controls in A1 and V1. For all six condition pairs, the average prediction accuracies across subjects are depicted. For a fair comparison between sighted and blind participants, we also report the prediction accuracies when only the first four runs in the blind subjects are used for classification. Error bars represent the SEM. The red dashed line indicates chance level (50%) and is not used for assessment of statistical significance. An asterisk (*) in a bar reflects a significant decoding accuracy compared with the empirically estimated chance level. *P < 0.05, FDR-corrected at q = 0.05.
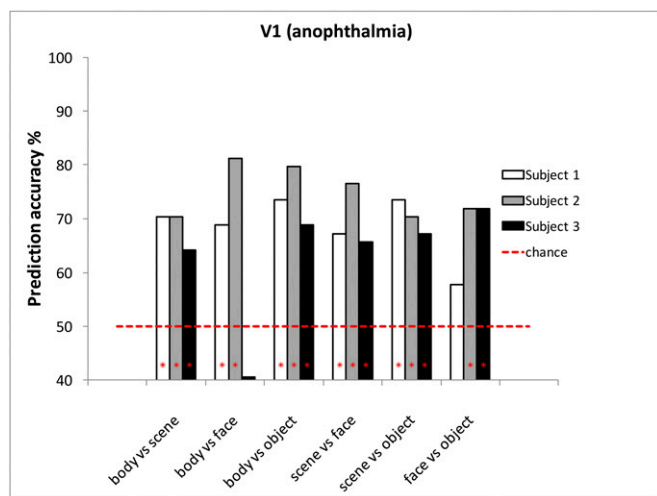
van den Hurk et al.

Fig. 5. Classification results in V1 for three anophthalmic subjects. The red dashed line indicates chance level (50%) and is not used for assessment of statistical significance. An asterisk (*) in a bar reflects a significant decoding accuracy compared with the empirically estimated chance level. *$P < 0.05$, FDR-corrected at $q = 0.05$.

responses, but not vice versa (33). Alternatively, the asymmetry may be based on inherent properties of the data themselves, such as differences in signal-to-noise ratio between the auditory and visual runs. Ultimately, the data themselves do not provide definitive insights into the mechanism that underlies asymmetry.

Between-Subject Classification Results: V1 and A1. To see whether the neural selectivity maps in V1 and A1 observed in the blind group correspond to the maps that process auditory and visual stimuli in the controls, we attempted the cross-subject decoding approach on these brain regions. In both generalization directions, however, the classifier failed to decode the categories from the neural responses in V1 and A1 above empirically estimated chance levels (training on blind, testing on controls: average prediction accuracy V1: $51.3 \pm 2.7\%$, A1: $53.0 \pm 3.4\%$; training on controls, testing on blind: average prediction accuracy V1: $50.9 \pm 1.7\%$, A1: $52.3 \pm 2.2\%$).

We tested whether the between-group classification in V1 and A1 is significantly lower than the aforementioned between-group classification in VTC. The results of such between-region comparisons in classifier performance should be interpreted with care, but to make the conditions at least a bit comparable, we reduced all regions to a standard size. More specifically, we selected the 100 vertices with the best generalized linear model (GLM) fits across subjects for VTC, A1, and V1 and ran the blind-to-sighted (auditory-to-visual) classification analysis on these data points. When statistically quantifying the differences in performance across subjects and condition pairs in VTC and the other two regions of interest (ROIs), we find that the between-group classification in VTC is higher than in V1 ($P = 0.0017$) and A1 ($P = 0.0203$).

Consistency of the Functional Map Across Subjects. Fig. 7 provides a detailed picture of the overall degree of similarity in the cortical layout of the functional responses in VTC among subjects, between subject groups, and between modalities. The figure format also provides a complete view of the interindividual variability. As can be seen, the visual maps show a high internal consistency ($r = 0.49$), whereas the average correlation of the auditory maps in the controls is much lower ($r = 0.05$). Interestingly, from this matrix it becomes clear that the internal consistency of the auditory maps of the blind group is higher than that of the controls

($r = 0.15$ vs. $r = 0.03$, $P < 0.001$). Moreover, the correlation of the auditory maps of the blind individuals with the visual maps of the controls is significantly stronger than the correlation of the auditory maps of the controls with their visual maps ($r = 0.10$ vs. $r = 0.05$, $P < 0.01$).

Discussion

Studying the visual system in blind individuals offers a rare window into the development of functional specificity in the brain. In this paper, we investigated the contention that the categorical selectivity and functional organization of VTC can develop independently of visual experience. We tested two hypotheses: (*i*) that face-, body-, object-, and scene-related natural sounds elicit distinct responses in VTC of congenitally blind individuals, and (*ii*) that the functional architecture of VTC in these blind participants shows significant overlap with the layout of this brain region in sighted
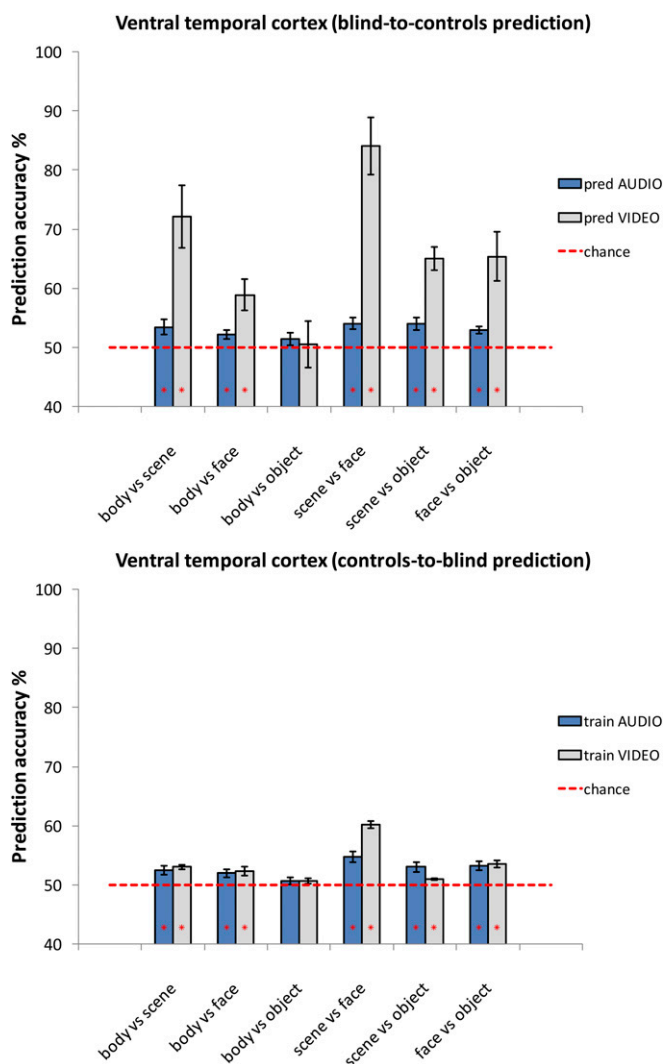
Fig. 6. (*Upper*) Blind-to-controls generalization performance: prediction accuracies observed when the classifier was trained on the auditory trials from the blind participants and tested on either the auditory or the visual trials from the sighted controls. (*Lower*) Controls-to-blind generalization performance: prediction accuracies observed when the classifier was trained on either the visual or the auditory trials from the sighted controls and tested on the auditory trials from the blind subjects. The red dashed line indicates chance level (50%) and is not used for assessment of statistical significance. An asterisk (*) in a bar reflects a significant decoding accuracy compared with the empirically estimated chance level. *$P < 0.05$, FDR-corrected at $q = 0.05$.

van den Hurk et al.

www.manaraa.com

controls. We found that VTC in the blind is indeed sensitive to all four categories, showing distinct responses that were decodable within and to a large extent across subjects. However, it is important to note that the category selectivity in the blind subjects for the auditory stimuli is not as strong as the selectivity in the controls for visual stimuli. Still, the auditory maps in VTC are stronger in the blind group than in controls. Finally, we demonstrated a striking similarity in the cortical response layout of VTC in blind individuals and sighted controls, demonstrating that categorical sensitivity in the extrastriate cortex does not require visual input.

Layout of VTC in Blind Individuals Corresponds to the Layout in Sighted Controls. Through statistical contrasts and the spatial mapping of selectivity, we have demonstrated that the global functional layout of VTC is remarkably similar to the well-known functional architecture described in the literature (34, 35). Directly correlating the functional map of the blind individuals with the visual responses in the sighted controls yields a substantial similarity.

VTC in Blind Persons Functionally Discriminates Among Categories. Using MVPA, we demonstrated that VTC yields robust and distinct responses to the four categories used in this study. Moreover, we were able to predict visual responses in sighted control subjects by training a classifier on the neural responses in the blind individuals, revealing an important functional overlap between the cortical layouts in the two groups. These findings, in combination with the functional topography presented in the congenitally blind subjects, provide reliable evidence that the category-selective ventral cortex does not require visual experience to develop its functional architecture. This notion is strengthened by strong discriminatory response in VTC to the four categories shown by even our anophthalmic participants. This result counters the alternative explanation that a brief period of low-quality visual experience in the early youth of category 2 and 3 blind participants might have functionally shaped their visual brain regions.

A partially intact functional architecture of VTC in individuals without visual experience can have two main interpretations. The first is that the categorical selectivity is innate; this proposal would be at odds with the hypothesis that a retinotopic protomap becomes fine-tuned over time via correlations between retinotopic and cate-

gorical input (9–11). In this scenario, the eccentricity bias of higher-order visual regions is a purely hard-wired principle, strictly following the retinotopic organization of the earlier stages of visual processing. In this case, the functional specificity of VTC is merely an extension of this principle, which would make testable predictions about the computational mechanisms that underlie visual categorical processing (36, 37). Alternatively, the development of categorical selectivity in VTC could be driven in part by multimodal or supramodal principles (38). In this view, a particular categorical sensitivity is hard-wired in the extrastriate cortex from birth. Through early development and exploration of the outer world, this part of the brain is fine-tuned via sensory input. Visual input is the primary modality for shaping category sensitivity, but when this sensory input is lacking from birth, other modalities can take over. This interpretation in terms of intermodality interactions might also extend to our finding that our control subjects show a less pronounced functional topography in the auditory modality than do blind individuals. If the organization of VTC were fully innate and multimodal from birth, one could reasonably expect a solid response to auditory stimulation in sighted individuals as well.

Indeed, our findings across all tests suggest stronger, not weaker, responses in VTC to auditory stimuli in blind than in sighted participants: stronger functional selectivity in the analyses shown in Fig. 1, a tendency for stronger decoding in the analyses of Fig. 2, and stronger map consistency in the matrices of Fig. 7. This strong auditory activation in VTC of blind participants is all the more striking because visual imagery, which is an important possible explanation for auditory activation in visual cortex in sighted individuals (39), cannot explain the auditory activation in congenitally blind participants.

Even though the category-selective maps uncovered in the blind through auditory stimuli tended to be stronger in blind than in sighted individuals, these auditory maps are still much weaker than the category-selective maps found with visual stimulation. Thus, when we conclude that the category-selective functional organization in VTC develops independently from visual experience, we refer solely to the basic map that is similar across modalities and is present in the blind and to the auditory activation of this map that is even stronger in the blind. Still, it is striking that the activity pattern elicited by an object-related sound in visual cortex of a congenitally blind individual nicely
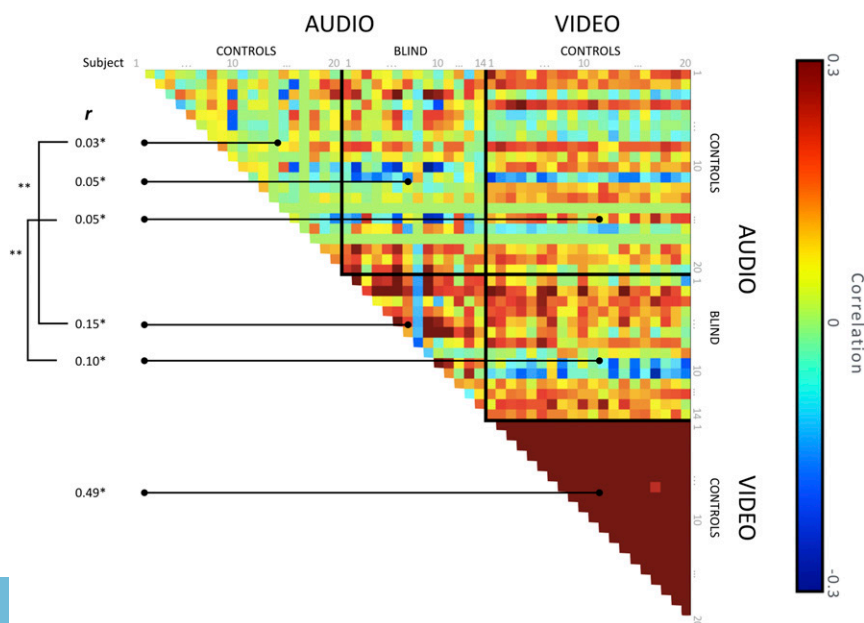

Fig. 7. Map consistency matrix across individual subjects. Each cell in the matrix represents the correlation of the functional maps (four conditions × V vertices) of VTC of two subjects. Before computing the correlation for each subject pair, we subtracted the mean across conditions per vertex. The matrix is subdivided in six submatrices comparing auditory and visual responses between controls and blind subjects. For each submatrix, we computed the average correlation coefficient across subjects (excluding the diagonal) and statistically assessed these means by ranking them against 1,000 means obtained through permutation. The maps of the blind subjects are based on the first four auditory runs, making them directly comparable with those of the control subjects. Note that for two control subjects we did not record the auditory trials; the corresponding rows/columns have the value 0 and are not included in the statistical assessment. *P < 0.0001, **P < 0.01, for the difference between two submatrices, FDR-corrected.

www.manaraa.com

predicts the activity pattern elicited by a visual movie clip of that same object in a sighted participant.

What Does VTC Represent in Blind Individuals? The activation of categorical representations in VTC of blind individuals raises some fundamental questions. First, what neural pathways mediate the categorical representations activated in VTC through auditory stimulation? Earlier work has demonstrated that nonvisual sensory input can be conveyed into the visual cortex in both blind and sighted (albeit to a smaller extent) individuals (38, 40). These nonvisual signals might reach visual cortex through thalamocortical projections or corticocortical connections, transferring information directly from primary sensory areas or indirectly from higher-order (multisensory) areas (e.g., the superior temporal sulcus). Moreover, in blind individuals these existing pathways will be altered through cross-modal plasticity via rewiring and/or unmasking or strengthening of existing connections. It has been shown that the process of synaptic pruning, which typically removes about 40% of the synaptic connections of visual cortex in early life (41), is limited in the case of early sensory deprivation. A decreased level of pruning might, at least in part, underlie the auditory-to-visual cross-modal plasticity in the early blind (42–45): the absence of visual input during early development allows the synaptic connections to strengthen in blind individuals (46). This cross-modal plasticity predicts a stronger response in visual cortex of blind persons when auditory stimuli are processed and perhaps could explain why we observed a stronger selectivity for auditory stimuli in VTC of the congenitally blind participants. However, in ventral visual cortex we see not only a stronger overall auditory response but also a regional distribution of category preference that mimics the preference in the sighted brain. To explain this distribution as the result of pruning, we need to assume that the synapses that run from the auditory stream to visual cortex early in development already have a rudimentary category preference. Recent findings regarding the visual word form area illustrate the importance of early connectivity patterns for the functional development of visual cortex. In a longitudinal study, structural connectivity patterns (but not functional responses) measured in children at a prereading age could predict the cortical location of the visual word form area measured 3 y later when children had learned to read (47).

Another fundamental question regarding categorical representations in VTC concerns the information that is represented: What does VTC of the blind person encode when it is presented with auditory stimuli, and why is the selectivity of VTC responses stronger in congenitally blind individuals? Although our data allow cross-decoding between the blind and sighted participants, this cross-decoding does not give direct insights into the information that is represented in VTC in blind individuals. Several non-mutually exclusive interpretations are possible. For instance, VTC could be encoding the auditory properties of the sound stimuli. Alternatively, the sounds trigger might more abstract category-related representations such as form, semantics, or perhaps even linguistics. We could argue that the cross-decoding with the sighted individuals makes some of these scenarios implausible, because there is no evidence for them in the literature on sighted individuals (e.g., for the representation of auditory dimensions). However, it is possible that VTC in blind individuals encodes one of these dimensions but VTC in sighted controls does not. Thus, the informational content could be very different in the two subject groups. We know only that the content gives rise to the same preferences at the category level and that even such a mere overlap in functional topography between the two groups leads to the significant cross-decoding. Note that such a situation with multiple correlated biases also seems to be a general property of visual category representations in sighted individuals. For example, in sighted individuals the category preference for scenes is correlated with a feature preference for higher spatial frequencies, straight lines, and eccentric stimulation (for more discussion, see ref. 48).

Finally, even though we find significant cross-decoding, our results do not rule out the possibility that part of the category preferences in the blind individuals is idiosyncratic to this population. Indeed, in Fig. 7, the auditory maps of blind subjects seem to be more similar to the auditory maps of other blind participants than to the visual maps of the sighted participants. This finding is surprising, because the visual maps of the sighted are more robust and more consistent among themselves, and might suggest that the maps of the blind participants have peculiarities that are not present in the maps of the sighted.

V1 and A1. The finding that V1 discriminates among the four sound categories only in the blind illustrates the brain's ability to reorganize itself when a brain region is deprived of input (49). Previous work in the blind has demonstrated that when the brain is deprived of visual input, early visual cortex can show involvement in complex cognitive tasks, such as listening to spoken language (50) and spatially informative sounds (51). Earlier work in sighted individuals also has demonstrated that auditory scenes elicit discriminable responses in early visual cortex (52), although our sighted controls did not show this effect. The exact functional mechanisms behind this selectivity for auditory stimuli in V1 are still to be investigated. A possible explanation might be that, because we focused primarily on higher-level categorical boundaries, we did not control for all low-level auditory properties of our stimuli. Features such as spectral properties and modulation over time may differ among categories and could have activated V1 differently in our blind participants.

In A1, we found that selectivity for the auditory conditions was stronger in the sighted controls than in the blind participants. The direction of this difference was unexpected, given that blind participants rely upon auditory perception much more than do the controls and do not have any visual input to compete for attentional resources. This finding is all the more striking because the strength of the selectivity for auditory stimuli tended to go in the opposite direction in ventral visual cortex, with stronger selectivity for auditory stimuli. Still, this seemingly reduced functionality of nondeprived brain regions has been reported before in studies on early blind individuals (53, 54). It seems as if some of the processing in nondeprived cortex has been replaced by the responses in deprived visual areas (55). The finding is also helpful for the interpretation of our findings in visual cortex, because the weaker decoding for the sounds in the auditory cortex of the blind rules out the possibility that the stronger decoding in VTC is related to an enhanced processing of the sounds.

Interestingly, the activation patterns in V1 and A1 did not allow cross-decoding between visual and auditory stimuli. Note that we should be careful not to attach too many conclusions to this outcome, given that it is possible that the between-subject classification fails for more technical reasons, such as the activation patterns being organized differently in different subjects or are organized at a finer scale than in VTC.

Natural Sounds Related to Visual Categories. The use of natural sounds to investigate category-selectivity in ventral visual cortex has proven to stimulate categorical regions in visual cortex of blind individuals successfully, providing an interesting perspective on the domain × modality interaction that was recently proposed (56). Note that our study does not allow us to compare the response to natural sounds directly with the response to spoken words in blind individuals. In addition, we emphasize that our finding of shared category preferences between blind and sighted individuals does not pinpoint the exact informational content that underlies this between-group similarity; the informational representation is an important question for future research.

Experimental Procedures

Participants. Fourteen congenitally blind (five females, mean age 37.1 y) and 20 sighted adults with normal or corrected-to-normal vision (11 females, mean age 34.5 y) took part in this study. Because of technical problems, we could not record auditory runs for two of the sighted subjects, resulting in 18 subjects with both visual and auditory data and two subjects with only visual data. Participants were screened for fMRI compatibility, signed informed consent, and were financially compensated for their participation. The study was approved by the medical ethical committee of Katholieke Universiteit Leuven.

Gradation of Blindness. The blind participants had to meet two inclusion criteria that are typical in the literature of congenital blindness: (*i*) that the onset of blindness was congenital, and (*ii*) that the participant had no (memory of) object or shape perception. We assigned the blind volunteers to three categories of blindness: category 1: anophthalmia, i.e., no eyeballs developed in utero; category 2: no (memory of) light–dark perception; and category 3: light–dark perception in one or both eyes. See Table 1 for more details. Objectively, category 1 blind participants are the only participants whom one can be certain had no visual perception even in early postnatal weeks or months.

Stimuli. We used four visual categories for both visual and auditory stimulation: faces, body parts, artificial objects, and scenes. Visual stimuli consisted of 64 short (~1,800 ms) movie clips with 16 unique clips per category. In the face category, clips portrayed actions such as laughing, chewing, blowing a kiss, and whistling. In the body parts category, clips showed scratching, hand clapping, finger snapping, bare foot steps, and knuckle cracking. In the objects category we created clips of a car, a washing machine, a clock, tools, a mug, a glass, a ball, and a fan. In the scenes category, the clips depicted waves crashing on a beach, a restaurant overview, a train station, traffic, a forest, a lake, and a grass field. Auditory stimuli consisted of 64 short (~1,800 ms) audio clips with 16 unique clips per category. Stimuli were matched in overall sound intensity by normalizing the rms of the sound pressure levels. In the face category, clips reflected facial sounds such as laughing, chewing, blowing a kiss, and whistling. In the body category, sounds illustrated scratching, hand clapping, finger snapping, and foot steps. In the objects category, sounds depicted various artificial man-made objects and tools, such as a car starting, a washing machine, a bouncing ball, fluid poured into a glass, and the ripping of paper. In the scenes category, sounds reflected waves crashing on a beach, a calm lake, a forest with birds, a crowded restaurant, a train station, and a busy road. We attempted to match each visual stimulus with a sound stimulus, that is, the visual movie clip of clapping hands and the sound of clapping hands.

Design. With the sighted participants we scanned four runs each in the auditory and visual modalities. The auditory runs were presented before the visual runs. The blind volunteers were presented with eight auditory runs. Within each run, the stimuli were presented in a block design, with four blocks per condition for each run. In each block, eight stimuli were presented with a 200-ms interstimulus interval. The conditions were presented in an ABCD order, in which "A," "B," "C," and "D" refer to the four different conditions. The blocks of each condition were separated by a 200-m interval, until all four conditions were shown. After each ABCD block, a fixation cross was presented for 12 s. The ABCD presentation order was counterbalanced within and between runs to account for possible order effects. Each run lasted ~7.5 min.

Task. In the scanner, subjects were asked to perform a one-back task in which they had to rate the conceptual dissimilarity of each presented stimulus to the previous stimulus on a scale of 1–4 with a button press. We gave each subject the verbal example of a cat meowing and a dog barking being conceptually similar, and a cat meowing and a car starting being dissimilar. Our rationale was that this task would encourage the participants to process the categorical information in the stimuli attentively. Before scanning, but after the participants were provided with hearing protection, we adjusted the volume to an individually comfortable audio level.

To verify whether subjects perceived the stimuli as intended, we investigated if the subjects' rating of dissimilarity when one category block ended and a new category began was significantly higher than the dissimilarity rating within a category block. For this analysis, we excluded subjects who gave fewer than five responses at the between-block trials (which might happen because our stimulus speed was relatively fast to allow swift responses at a block change in this subjective comparison task). In addition, when multiple button presses were given in a trial, we only used the first response. Task performance shows that participants perceived the stimuli in a categorical manner. For the blind participants ($n = 11$), the mean dissimilarity ($\pm$ SEM) of the first stimulus in a category block to the last stimulus in the previous block was 3.41 ± 0.12 across subjects, significantly higher [t (10) = 6.05, $P < 0.001$] than the mean dissimilarity of stimuli within a block,

2.33 ± 0.16 across subjects. For the control subjects ($n = 15$), the among-category dissimilarity rating was 3.62 ± 0.11 on average, whereas the within-category dissimilarity was rated significantly lower [2.40 ± 0.11, t (14) = 8.8, $P < 0.001$]. This difference demonstrates that subjects picked up the category boundaries.

As an additional and conservative estimate of stimulus recognition and categorization, we designed two simple behavioral tasks. The first experiment was intended to quantify how well the sound stimuli could be categorized correctly. Twenty healthy volunteers (not included in the imaging experiment) were presented twice with all the sound stimuli in a random order. Participants were asked to determine the correct condition with a button press. On average, the face sounds were correctly categorized 89% of the time, the body sounds 76%, scene sounds 96%, and object sounds 89%. Note that these accuracies reflect the stimuli perceived in isolation, rather than in blocks of eight stimuli as during the scanning paradigm. Thus, these accuracies are a conservative estimate of perceived category membership in a blocked presentation. The second experiment was intended to test whether the individual stimuli could be recognized correctly. Twenty healthy volunteers (not included in the imaging experiment or in the first behavioral test) were presented with the individual sounds. On screen, 50 alphabetically ordered options were presented from which the participant had to pick the correct answer. Across subjects, 83% of the stimuli were recognized correctly on average. Divided per category, recognition accuracies were as follows: face sounds 98%, body sounds 81%, scene sounds 69%, and object sounds 86%. Because some of the answer options closely resembled other options in the same category, we also computed the accuracy with which subjects picked an answer from the correct category: face sounds 98%, body sounds 84%, scene sounds 96%, and object sounds 96%.

Data Acquisition Parameters. Functional and anatomical images were acquired on a 3T Philips Ingenia CX scanner (Department of Radiology, University of Leuven) with a 32-channel head coil. Each functional run consisted of 225 T2*-weighted echoplanar images (EPIs) [1.875 × 1.875 mm in-plane voxel size, 32 2.2-mm slices, interslice gap 0 mm, repetition time (TR) = 2,000 ms, echo time (TE) = 30 ms, 112 × 112 matrix]. In addition to the functional images, we collected a high-resolution T1-weighted anatomical scan for each participant (182 slices, resolution 0.98 × 0.98 × 1.2 mm, TR = 9.6 ms, TE = 4.6 ms, 256 × 256 acquisition matrix). Stimuli were presented using custom-written MATLAB R2014a code (MathWorks Inc.) and Psychtoolbox-3 (57) via an NEC projector with an NP21LP lamp that projected the image onto a screen the participant viewed through a mirror. The viewing distance was ~64 cm.

Data Preprocessing. Functional and anatomical data were preprocessed using the BrainVoyager QX2.8 package (Brain Innovation). The remainder of the data processing and analyses were performed using custom-written MATLAB code in combination with the NeuroElf v0.9c toolbox (NeuroElf.net). Functional volumes were first corrected for differences in slice scan time and 3D head motion using three translation and three rotation parameters. Subsequently, linear trends and low-frequency temporal drifts were removed from the data using a high-pass filter, removing temporal frequencies below four cycles per run. After the preprocessing, functional data were coregistered to the high-resolution anatomical volume and were normalized to Talairach space. For each subject, a white-matter surface reconstruction was made for each hemisphere separately.

Trial Estimation. Individual responses to the experimental trials were estimated by fitting a double-gamma hemodynamic response function (HRF) to each voxel's time course, using the resulting beta as trial estimate. The trials were labeled according to their corresponding condition (face, body, scene, object). The functional voxel-based data were subsequently sampled to the reconstructed vertex-based surface (integrated in depth along the vertex normals from −1.0 mm to 3.0 mm using linear interpolation).

Cortex-Based Alignment. A cortex-based alignment (CBA) procedure (58) was applied as implemented in the BrainVoyager QX 2.8 software package. For the left and right hemisphere individually, each subject's curvature information was aligned to the dynamic group ($n = 34$) average. On the resulting average surface reconstruction, VTC was defined manually by observing anatomical boundaries: the occipitotemporal sulcus, posterior transverse collateral sulcus, parahippocampal gyrus, and the anterior tip of the MFS (35). By using the vertex migration information that resulted from the CBA procedure, we could create anatomically matched VTC masks for each subject to use in the subsequent analyses. In addition to VTC, we also used V1 and A1 as ROIs by selecting the calcarine sulcus and Heschl's gyrus, respectively, bilaterally on the average cortex reconstruction.

www.manaraa.com

Table 1. Blind participant demographics

ID no.	Sex	Age, y	Cause of blindness	Residual light perception	Category of blindness
1	Female	19	Anophthalmia	None	1
2	Male	62	Anophthalmia	None	1
3	Male	24	Anophthalmia	None	1
4	Female	39	Retinopathy of prematurity	None	2
5	Male	51	Leber's congenital amaurosis	None	2
6	Female	39	Congenital glaucoma	None	2
7	Male	29	Leber's congenital amaurosis	None	2
8	Male	42	Leber's congenital amaurosis	L/D in right eye	3
9	Male	36	Leber's congenital amaurosis	L/D in both eyes	3
10	Male	28	Leber's congenital amaurosis	L/D in both eyes	3
11	Female	48	Congenital glaucoma	L/D in both eyes	3
12	Male	37	Leber's congenital amaurosis	L/D in both eyes	3
13	Male	34	Etiology unknown	L/D in both eyes	3
14	Female	32	Perinatal infection (exact etiology unknown)	L/D in both eyes	3

Categories of blindness: 1, anophthalmia; 2, no (memory of) light–dark perception; 3, light–dark perception but no shape/border detection. L/D, light–dark perception.

Univariate Analyses. We first performed a univariate group analysis on the surface-based beta maps after CBA for the visual and auditory sessions individually. For each vertex in our ROI, the mean response (beta) to each condition was computed by first averaging the responses within each subject and then averaging across subjects. Based on these averages, we selected the condition with the strongest response for each vertex. We then performed a paired-samples t test across subjects between the condition that yielded the strongest response in that vertex and the mean response of the remaining three conditions, and the t statistic was assigned to the vertex. The resulting maps were corrected for multiple comparisons by maintaining a false-discovery rate (FDR) with a proportion of false discoveries (q) = 0.05.

Topographical Selectivity Maps. In addition to univariate statistical mapping, we plotted the unthresholded selectivity after normalizing for global differences in response. More specifically, we first computed the average response across subjects to each of the four conditions per vertex, after which we computed the average response across all vertices in VTC. We then subtracted this average per condition from each vertex and mapped the condition that had the strongest normalized beta for that vertex. This process resulted in distinct clusters of vertices that spatially distinguish themselves from their surround in terms of selectivity for a particular condition.

Within-Subject Classification. For the classification procedure, we used a linear support-vector machine (l-SVM) (59, 60), based on the LIBSVM algorithm (https://www.csie.ntu.edu.tw/~cjlin/libsvm/). We selected features (vertices) based on the anatomically defined VTC mask that was created after cortex-based alignment. Each trial was normalized (z-scored) to obtain a mean response of 0 and an SD of 1 across all vertices within our ROI. For each pairwise comparison of categories, we trained the classifier with the labeled training data using a leave-one-run-out cross-validation. Thus we ran the classification procedure four or eight times (depending on whether the participant was sighted or blind), using each run once as test data; then the final classification accuracy was computed by averaging over the four or eight folds. We used a randomization procedure to estimate the distribution parameters under the null hypothesis. For each of the cross-validation folds in the original classification, 1,000 permutations were run in which we randomly permuted the condition labels, restricted to a label shuffle within each run (with no shuffling across runs). From the 4 × 1,000 or 8 × 1,000 prediction accuracies, we obtained the 1,000 mean prediction accuracies under the null hypothesis by averaging over all four or eight folds. By ranking the true prediction accuracy against the ranked permutation scores, we subsequently assessed the significance of the classification for a particular subject.

To test the classification performance across all participants per condition pair, we tested the prediction accuracies against the averaged permutation accuracies using a nonparametric Wilcoxon sign rank test (61). The P values were corrected for multiple comparisons with an FDR correction with q = 0.05.

To compare the classification results between the blind and healthy subjects directly, we first averaged the prediction accuracies across category pairs per subject group. Then, we tested the mean difference between the groups against the distribution of means obtained by 10,000 shuffles of group labels across subjects.

Between-Subject Classification. The rationale for the between-subject classification analyses was to quantify the extent to which the response profile of a given subject could be generalized to all other subjects. The procedure was largely similar to the within-subject analyses, with some small adjustments. First, we ran the classification analysis for each subject pair, i.e., the algorithm was trained on subject 1 and subsequently was tested on subjects 2–20. The 19 prediction accuracies were averaged and assigned to subject 1 as the generalization index for that subject. This procedure was repeated for all six pairwise comparisons of the categories. Here we also used a nonparametric Wilcoxon sign rank test (61). The P values were corrected for multiple comparisons with an FDR correction with q = 0.05. It is important to note that MVPA is often used to study neural representations that might not align spatially across subjects (62, 63). The rationale of our between-subject classification fits with the use of MVPA to study larger-scale maps (64, 65) and was appropriate, given that the category preferences that we target have been shown to fall in similar regions in different subjects, as already demonstrated for faces (1, 66), bodies (2), and scenes (5).

Testing the Topographical Consistency Across Subjects. We directly compared the consistency of the functional spatial topography between subjects, disregarding univariate effect size differences between modalities. First, for each subject and for the auditory (AUD) and visual (VIS) runs separately, we created a 4 × V matrix, which contains the mean response amplitude (beta) per condition for each vertex. This value was obtained by averaging all betas across all runs per condition. To make the maps of the blind subjects comparable to those of the sighted controls in terms of the number of runs, we based the maps of the blind participants on only the first four runs. The resulting 4 × V matrix captures the full response topography of VTC of a subject in a given modality. We then computed the correlation between each subject pair, i.e., the Pearson correlation between the auditory (AUD) data of subject 1 and subject 2, and so forth. In addition to within-modality comparisons, we also computed the correlation coefficients between modalities, i.e., comparing AUD subject 1 and the visual data (VID) of subject 2. To prevent unspecific activity differences of vertices from biasing the correlation measures to be above zero even if no selectivity would be present, we subtract the vertex's mean response across conditions from each vertex (7, 8, 67). The resulting correlation coefficients were assigned to the upper diagonal of the matrix. Note that although the correlation matrix visually resembles a representational dissimilarity matrix (68, 69), the information content of our correlation matrix is fundamentally different: the rows and columns correspond not to conditions but to a combination of subjects and stimulus modalities.

To perform statistics on the pairwise correlation coefficients found between subjects, we used a randomization approach. For each pairwise comparison between subjects, we permuted the four runs of one of these subjects (the shuffled subject) by shuffling the four labels. There are 24 permutations of the four category labels, and because we have four runs, there are a total of 24^4 possible shuffles. We randomly sampled 1,000 shuffles from these possible permutations of labels. For each permutation, we computed the 4 × V matrix as described above but with randomly assigned condition labels per run. Then, the correlations were computed between the unaffected subject map and the 1,000 shuffled maps of the paired subject. By repeating this procedure for each subject pair, we created 1,000 correlation matrices under the null hypothesis.

For all statistical assessments of the correlation results, for a particular submatrix of interest (e.g., the between-subject consistency of the functional maps related to the visual stimulation), we averaged the correlation coefficients from the appropriate cells and ranked this average against the distribution of averaged correlation coefficients from the permutations. When comparing two submatrices, we computed the average correlation coefficients of the two respective submatrices and subtracted these values from each other. This process was repeated for the 1,000 permutation correlations for both submatrices; then the true difference score was ranked against the distribution of permutations.

1. Kanwisher N, Yovel G (2006) The fusiform face area: A cortical region specialized for the perception of faces. *Philos Trans R Soc Lond B Biol Sci* 361:2109–2128.
2. Downing PE, Jiang Y, Shuman M, Kanwisher N (2001) A cortical area selective for visual processing of the human body. *Science* 293:2470–2473.
3. Schwarzlose RF, Baker CI, Kanwisher N (2005) Separate face and body selectivity on the fusiform gyrus. *J Neurosci* 25:11055–11059.
4. Malach R, et al. (1995) Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc Natl Acad Sci USA* 92:8135–8139.
5. Epstein R, Kanwisher N (1998) A cortical representation of the local visual environment. *Nature* 392:598–601.
6. Grill-Spector K, Kourtzi Z, Kanwisher N (2001) The lateral occipital complex and its role in object recognition. *Vision Res* 41:1409–1422.
7. Haxby JV, et al. (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293:2425–2430.
8. Op de Beeck HP, Haushofer J, Kanwisher NG (2008) Interpreting fMRI data: Maps, modules and dimensions. *Nat Rev Neurosci* 9:123–135.
9. Hasson U, Harel M, Levy I, Malach R (2003) Large-scale mirror-symmetry organization of human occipito-temporal object areas. *Neuron* 37:1027–1041.
10. Levy I, Hasson U, Avidan G, Hendler T, Malach R (2001) Center-periphery organization of human object areas. *Nat Neurosci* 4:533–539.
11. Hasson U, Levy I, Behrmann M, Hendler T, Malach R (2002) Eccentricity bias as an organizing principle for human high-order object areas. *Neuron* 34:479–490.
12. Deen B, et al. (2017) Organization of high-level visual cortex in human infants. *Nat Commun* 8:13995.
13. Kouider S, et al. (2013) A neural marker of perceptual consciousness in infants. *Science* 340:376–380.
14. He C, et al. (2013) Selectivity for large nonmanipulable objects in scene-selective visual cortex does not require visual experience. *Neuroimage* 79:1–9.
15. Mahon BZ, Anzellotti S, Schwarzbach J, Zampini M, Caramazza A (2009) Category-specific organization in the human brain does not require visual experience. *Neuron* 63:397–405.
16. Reich L, Szwed M, Cohen L, Amedi A (2011) A ventral visual stream reading center independent of visual experience. *Curr Biol* 21:363–368.
17. Striem-Amit E, Cohen L, Dehaene S, Amedi A (2012) Reading with sounds: Sensory substitution selectively activates the visual word form area in the blind. *Neuron* 76:640–652.
18. Pietrini P, et al. (2004) Beyond sensory images: Object-based representation in the human ventral pathway. *Proc Natl Acad Sci USA* 101:5658–5663.
19. Kitada R, et al. (2013) Early visual experience and the recognition of basic facial expressions: Involvement of the middle temporal and inferior frontal gyri during haptic identification by the early blind. *Front Hum Neurosci* 7:7.
20. Kitada R, et al. (2014) The brain network underlying the recognition of hand gestures in the blind: The supramodal role of the extrastriate body area. *J Neurosci* 34:10096–10108.
21. Wong RO (1999) Retinal waves and visual system development. *Annu Rev Neurosci* 22:29–47.
22. Ackman JB, Burbridge TJ, Crair MC (2012) Retinal waves coordinate patterned activity throughout the developing visual system. *Nature* 490:219–225.
23. Verma AS, Fitzpatrick DR (2007) Anophthalmia and microphthalmia. *Orphanet J Rare Dis* 2:47.
24. Coullon GSL, Jiang F, Fine I, Watkins KE, Bridge H (2015) Subcortical functional reorganization due to early blindness. *J Neurophysiol* 113:2889–2899.
25. Watkins KE, et al. (2012) Language networks in anophthalmia: Maintained hierarchy of processing in 'visual' cortex. *Brain* 135:1566–1577.
26. Watkins KE, et al. (2013) Early auditory processing in area V5/MT+ of the congenitally blind brain. *J Neurosci* 33:18242–18246.
27. Weiner KS, et al. (2014) The mid-fusiform sulcus: A landmark identifying both cytoarchitectonic and functional divisions of human ventral temporal cortex. *Neuroimage* 84:453–465.
28. Man K, Kaplan JT, Damasio A, Meyer K (2012) Sight and sound converge to form modality-invariant representations in temporoparietal cortex. *J Neurosci* 32:16629–16636.
29. Oosterhof NN, Tipper SP, Downing PE (2012) Visuo-motor imagery of specific manual actions: A multi-variate pattern analysis fMRI study. *Neuroimage* 63:262–271.
30. Akama H, Murphy B, Na L, Shimizu Y, Poesio M (2012) Decoding semantics across fMRI sessions with different stimulus modalities: A practical MVPA study. *Front Neuroinform* 6:24.
31. Quadflieg S, et al. (2011) Puddles, parties, and professors: Linking word categorization to neural patterns of visuospatial coding. *J Cogn Neurosci* 23:2636–2649.
32. Kaplan JT, Man K, Greening SG (2015) Multivariate cross-classification: Applying machine learning techniques to characterize abstraction in neural representations. *Front Hum Neurosci* 9:151.
33. Etzel JA, Gazzola V, Keysers C (2008) Testing simulation theory with cross-modal multivariate classification of fMRI data. *PLoS One* 3:e3690.
34. Grill-Spector K (2003) The neural basis of object perception. *Curr Opin Neurobiol* 13:159–166.
35. Grill-Spector K, Weiner KS (2014) The functional architecture of the ventral temporal cortex and its role in categorization. *Nat Rev Neurosci* 15:536–548.
36. Henriksson L, Mur M, Kriegeskorte N (2015) Faciotopy-A face-feature map with face-like topology in the human occipital face area. *Cortex* 72:156–167, 10.1016/j.cortex.2015.06.030.
37. van den Hurk J, Pegado F, Martens F, Op de Beeck HP (2015) The search for the face of the visual homunculus. *Trends Cogn Sci* 19:638–641.
38. Ricciardi E, Bonino D, Pellegrini S, Pietrini P (2014) Mind the blind brain to understand the sighted one! Is there a supramodal cortical functional architecture? *Neurosci Biobehav Rev* 41:64–77.
39. Reddy L, Tsuchiya N, Serre T (2010) Reading the mind's eye: Decoding category information during mental imagery. *Neuroimage* 50:818–825.
40. Qin W, Yu C (2013) Neural pathways conveying novisual information to the visual cortex. *Neural Plast* 2013:864920.
41. Innocenti GM, Price DJ (2005) Exuberance in the development of cortical networks. *Nat Rev Neurosci* 6:955–965.
42. Lewis LB, Saenz M, Fine I (2010) Mechanisms of cross-modal plasticity in early-blind subjects. *J Neurophysiol* 104:2995–3008.
43. Bavelier D, Neville HJ (2002) Cross-modal plasticity: Where and how? *Nat Rev Neurosci* 3:443–452.
44. Pascual-Leone A, Amedi A, Fregni F, Merabet LB (2005) The plastic human brain cortex. *Annu Rev Neurosci* 28:377–401.
45. Merabet LB, Pascual-Leone A (2010) Neural reorganization following sensory loss: The opportunity of change. *Nat Rev Neurosci* 11:44–52.
46. Collignon O, et al. (2011) Functional specialization for auditory-spatial processing in the occipital cortex of congenitally blind humans. *Proc Natl Acad Sci USA* 108:4435–4440.
47. Saygin ZM, et al. (2016) Connectivity precedes function in the development of the visual word form area. *Nat Neurosci* 19:1250–1255.
48. Bracci S, Op de Beeck H (2016) Dissociations and associations between shape and category representations in the two visual pathways. *J Neurosci* 36:432–444.
49. Florence SL, Kaas JH (1995) Large-scale reorganization at multiple levels of the somatosensory pathway follows therapeutic amputation of the hand in monkeys. *J Neurosci* 15:8083–8095.
50. Bedny M, Richardson H, Saxe R (2015) "Visual" cortex responds to spoken language in blind children. *J Neurosci* 35:11674–11681.
51. Striem-Amit E, Amedi A (2014) Visual cortex extrastriate body-selective area activation in congenitally blind people "seeing" by using sounds. *Curr Biol* 24:687–692.
52. Vetter P, Smith FW, Muckli L (2014) Decoding sound and imagery content in early visual cortex. *Curr Biol* 24:1256–1262.
53. Jiang F, Stecker GC, Fine I (2014) Auditory motion processing after early blindness. *J Vis* 14:4.
54. Dormal G, Rezk M, Yakobov E, Lepore F, Collignon O (2016) Auditory motion in the sighted and blind: Early visual deprivation triggers a large-scale imbalance between auditory and "visual" brain regions. *Neuroimage* 134:630–644.
55. Bock AS, Fine I (2014) Anatomical and functional plasticity in early blind individuals and the mixture of experts architecture. *Front Hum Neurosci* 8:971.
56. Bi Y, Wang X, Caramazza A (2016) Object domain and modality in the ventral visual pathway. *Trends Cogn Sci* 20:282–290.
57. Brainard DH (1997) The psychophysics toolbox. *Spat Vis* 10(4):433–436.
58. Goebel R, Esposito F, Formisano E (2006) Analysis of functional image analysis contest (FIAC) data with brainvoyager QX: From single-subject to cortically aligned group general linear model analysis and self-organizing group independent component analysis. *Hum Brain Mapp* 27:392–401.
59. Formisano E, De Martino F, Valente G (2008) Multivariate analysis of fMRI time series: Classification and regression of brain responses using machine learning. *Magn Reson Imaging* 26:921–934.
60. Cortes C, Vapnik V (1995) Support-vector networks. *Mach Learn* 20:273–297.
61. Wilcoxon F (1945) Individual comparisons by ranking methods. *Biom Bull* 1:80–83.
62. Kamitani Y, Tong F (2005) Decoding the visual and subjective contents of the human brain. *Nat Neurosci* 8:679–685.
63. Haynes J-D, Rees G (2006) Decoding mental states from brain activity in humans. *Nat Rev Neurosci* 7:523–534.
64. Konkle T, Oliva A (2012) A real-world size organization of object responses in occipitotemporal cortex. *Neuron* 74:1114–1124.
65. Connolly AC, et al. (2012) The representation of biological classes in the human brain. *J Neurosci* 32:2608–2618.
66. Kanwisher N, McDermott J, Chun MM (1997) The fusiform face area: A module in human extrastriate cortex specialized for face perception. *J Neurosci* 17:4302–4311.
67. Sayres R, Grill-Spector K (2008) Relating retinotopic and object-selective responses in human lateral occipital cortex. *J Neurophysiol* 100:249–267.
68. Kriegeskorte N (2008) Representational similarity analysis – Connecting the branches of systems neuroscience. *Front Syst Neurosci* 2:4.
69. Kriegeskorte N, Kievit RA (2013) Representational geometry: Integrating cognition, computation, and the brain. *Trends Cogn Sci* 17:401–412.