

Impact of a transient neonatal visual deprivation on the development of the ventral occipito-temporal cortex in humans

Received: 23 January 2025

Accepted: 15 October 2025

Published online: 17 November 2025

 Check for updates

Stefania Mattioni ^{1,2} ✉, Mohamed Rezk², Xiaoqing Gao³, Junghyun Nam⁴, Zhong-Xu Liu ⁵, Remi Gau ², Valérie Goffaux², Andrea I. Costantino ⁶, Hans Op de Beeck ⁶, Terri Lewis⁷, Daphne Maurer⁷ & Olivier Collignon ^{2,8} ✉

How does sensory experience shape the development of the visual brain? To answer this eluding question, we examine brain responses to visual categories in a rare group of cataract-reversal individuals who experienced a short transient period of early blindness. Encoding of low-level visual properties is impaired in the early visual cortex (EVC) of cataract-reversal participants, whereas categorical responses in downstream ventral occipito-temporal cortex (VOTC) are preserved. In controls, degrading visual input to mimic the visual deficits of cataracts produces cascading disruptions extending from EVC to VOTC, unlike in the cataract group. A deep neural network trained on altered visual input reproduces this dissociation, supporting the brain findings. These results demonstrate that while EVC is permanently affected by early deprivation, categorical coding in VOTC shows resilience, highlighting different sensitive periods for specific brain regions and computations.

The study of the consequence of an early and transient period of visual deprivation has long served as a compelling experimental model to causally test the need of sensory experiences for the functional development of brain regions¹.

Seminal studies have unveiled the existence of sensitive and critical developmental periods during which early life experience significantly shapes the functional tuning of the brain². Studies in mice^{3,4}, cats^{1,5–7} and monkeys⁸ have demonstrated profound and enduring effects of a brief and transient period of monocular or binocular deprivation in early life on responses of the early visual cortex (EVC) (for comprehensive reviews, see ref. 9). Similarly, humans who experience transient postnatal blindness due to dense bilateral cataracts show permanently reduced visual acuity^{10–13} and alteration in the response and retinotopic organization of the EVC^{14,15}.

The impact of a transient neonatal visual deprivation beyond EVC has been less explored; yet it is commonly assumed that these regions are altered due to their connection to, and dependence from, lower-level properties extracted from EVC^{16,17}. It has, for instance, been suggested that the transient absence of early visual experience could disrupt the typical development of category selectivity in the ventral occipito-temporal cortex (VOTC), potentially even to a greater extent than it impacts EVC^{15–18}. Most of the studies that have explored categorical representations in VOTC among cataract-reversal individuals have focused on the processing of faces. Some of these studies have reported impairments in tasks involving the localization of internal facial features^{19–21}, face memory tasks¹⁹, ratings of facial attractiveness²², automatic gaze following¹³ accompanied by abnormal²³ or absent²⁴ electrophysiological responses (N170) to face stimuli and altered connectivity within the face-processing network²⁵.

¹Department of Experimental Psychology & Center for Cognitive Neuroscience (CCN), UGent, Gent, Belgium. ²Institute for Research in Psychology (IPSY) & Neuroscience (IoNS), UCLouvain, Louvain-la-Neuve and Brussels, Belgium. ³Center for Psychological Sciences, Zhejiang University, Zhejiang, China. ⁴The Rotman Research Institute in Baycrest Health Sciences, University of Toronto, Toronto, ON, Canada. ⁵Department of Behavioral Sciences, University of Michigan-Dearborn, Dearborn, MI, USA. ⁶Department of Brain & Cognition, Leuven Brain Institute, KU Leuven, Leuven, Belgium. ⁷Department of Psychology, Neuroscience & Behaviour, McMaster University, Hamilton, ON, Canada. ⁸HES-SO Valais-Wallis, The Sense Innovation and Research Center, Lausanne & Sion, Switzerland. ✉e-mail: stefania.mattioni@ugent.be; olivier.collignon@uclouvain.be

Similarly, a study involving monkeys raised without selective exposure to faces has reported the absence of face domain in the brain after sight recovery²⁶.

These results are however not without controversies. Other studies showed that cataract-reversal individuals can successfully perform several tasks including face detection²³, face/nonface categorization²⁷, discrimination among individual faces based on the shape of the eyes, the mouth or the external contour²⁸ or the extraction of social information from faces, such as facial and emotional expression^{29,30}. In the same line, monkeys raised without exposure to any faces exhibited a preference for human and monkey faces in photographs, along with the ability to discriminate between human faces as well as monkey faces³¹.

No previous studies have simultaneously investigated the concurrent development of early and higher-level visual regions across different visual categories in the context of cataract-reversals. Moreover, so far, most studies comparing brain responses between cataract-reversal and controls have relied on univariate techniques. However, univariate and contrast approaches applied to a restricted set of stimuli are limited in their capacity to reveal the precise nature of the underlying brain computations that may differ between cataracts and controls, and how these potential changes in computation vary according to brain regions³².

Here, we examined functional differences along the ventral occipito-temporal stream between cataract reversal individuals and controls using decoding and representational similarity analysis (RSA)^{31,33} applied to functional Magnetic Resonance Imaging (fMRI) signals and artificial neural network models. Using a diverse set of images belonging to multiple visual categories including faces, bodies, houses, tools and words (Fig. 1A), we concomitantly probed alteration in the brain coding of either low-level (e.g., spatial frequencies) or high-level (e.g., categorical) visual features of afferent visual input. We also included a control fMRI experiment where we presented modified visual stimuli to additional control subjects to simulate the visual

deficits associated with cataract reversal (i.e., reduced visual acuity and nystagmus), allowing us to account for partially impaired vision at the time of testing. Additionally, we directly compared the brain's response along the visual cortical hierarchy with the representation of different deep neural network (DNN) layers, thought to emulate how information is sequentially extracted in VOTC³⁴.

We show that while early deprivation leaves lasting deficits in early visual cortex, categorical coding in ventral occipito-temporal regions remains resilient—revealing differential susceptibility across the visual hierarchy and challenging the view that the development of high-level areas critically depends on early visual experience. Complementary DNN simulations further support this interpretation, showing how structured categorical representations can emerge despite degraded early input.

Results

Visual acuity

The control group has a mean level of logMar of 0.01 (SD 0.07, min: -0.05, max:0.17), the cataract-reversal group (see Table 1) has a mean level of logMar of 0.21 (SD 0.18, min -0.05, max: 0.61). As expected^{10,12,13}, an independent samples T-test revealed a significant difference (Fig. 2A) of the level of visual acuity between the two groups ($t(28) = -4.1; p = 0.0003$).

Behavioral response in the fMRI

During the fMRI data acquisition, participants were asked to identify instances where two images did not repeat. We excluded one subject in the control group, one subject in the cataract-reversal group and one subject in the control experiment (CONblurry1 and CONblurry2) based on poor performance during fMRI data acquisition (under 2.5 standard deviations from the group mean).

After the exclusions, the mean values of accuracy were 94% in the control group, 91% in the cataract-reversal group, 94% in the CON-blurry1 experiment, and 95% in the CON-blurry2 experiment (Fig. 2B). Independent samples t-tests did not show any significant difference between the accuracy values across groups (pFDR between 0.63 and 0.99).

Amount of gaze displacement

Ten out of the 15 cataract-reversal subjects reported experiencing nystagmus (Table 1), a comorbid condition often associated with early visual deprivation and characterized by involuntary and repetitive eye movements. The small tracking window of typical in-scanner eye trackers in combination with the fact that participants wore individually fitted in-scanner corrective glasses (the same glasses that they wore during the visual acuity test) and the presence of nystagmus made the use of eye-tracker impracticable. We instead employed deepMRye³⁵ to examine the eye movement patterns of our subjects and used this data to account for this potential confounding factor in our study (Fig. 2C).

The average amount of gaze displacement inside the scanner, computed through deepMRye, was significantly higher in the cataract-reversal group compared to all the other groups (Fig. 2C; CAT vs CON ($t(28) = 2.61; p = 0.01; pFDR = 0.02$); CAT vs CON-blurry1 ($t(25) = -3.1; p = 0.004, pFDR = 0.02$); CAT vs CON-blurry2 ($t(25) = -2.78; p = 0.01, pFDR = 0.02$). Instead, the amount of gaze displacement did not differ between the control group and the CON-blurry1 ($t(27) = 0.15; p = 0.88, pFDR = 0.94$), nor between the control group and the CON-blurry2 ($t(27) = 0.08; p = 0.94, pFDR = 0.94$).

This difference between the cataract-reversal group and the others was expected, due the presence of nystagmus in most of the cataract-reversal individuals, and validates the indices provided by deepMRye. As described in the method section, we included the variance for the X and Y gaze positions as regressors of no-interest in

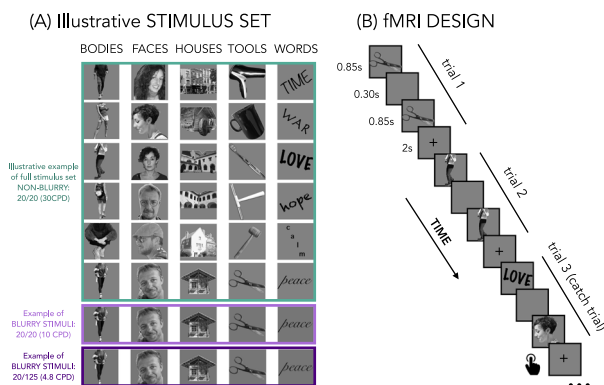


Fig. 1 | Stimulus set and experimental design. **A** illustrates representative examples of the stimulus categories used in the experiment. The green box encompasses the illustrative set of stimuli used in the main experiment, with each column representing a category and each row depicting different exemplars within the category, totaling 30 images. For the bodies, faces, houses, and tools categories, we present six example images per category that resemble the original experimental stimuli, which cannot be shared due to copyright restrictions. The six word exemplars are the original stimuli created by the authors for this experiment. Additionally, the light and dark purple boxes illustrate one exemplar per category in two control experiments where the images were spatially filtered at two different levels of blurriness: 10 Cycles per degree (CPD), representing the average visual acuity level of the reversal-ataract group (logMar=0.21), and 4.8 CPD, which denotes the lowest level of acuity recorded among the cataract-reversal individuals (logMar=0.61). All the stimuli, without any compression, can be found in the OSF project (<https://doi.org/10.17605/OSF.IO/BECDR>) of this study at this link. **B** depicts the experimental design during fMRI.

Table 1 | Characteristics of reversal cataract participants

Subjects	Age(y)	Sex	Length of deprivation (days)	Acuity (LogMar)	Nystagmus
CAT1	39	M	124	0.17	YES
CAT2	34	F	47	0.14	YES
CAT3	22	F	84	0.28	YES
CAT4	27	F	132	0.20	N.A.
CAT5	38	F	44	0.19	N.A.
CAT6 ^a	27	F	107	0.31	YES
CAT7	23	F	6	0.25	YES
CAT8	32	M	94	0.61	YES
CAT9	26	M	77	0.52	YES
CAT10	37	M	143	0.09	YES (mild)
CAT11	18	F	10	0.14	NO
CAT12	25	F	65	-0.05	YES (mild)
CAT13	26	M	21	0.26	YES
CAT14	27	M	118	0.09	NO
CAT15	23	M	3	0.04	NO

^aSub CAT6 has been excluded based on the low performance in the fMRI task.

our GLM (see next paragraph), to control for the impact of different magnitude of eye movement on the fMRI data activity.

Univariate analysis

We conducted this analysis to examine the activated regions (within and between groups) for all stimuli compared to the baseline (SI Fig. 1) and for each category (Fig. 3, SI Fig. 2) in contrast to all other categories. It is crucial to note that our experimental design was optimized for multivariate analyses, and the rapid event-related presentation typically does not yield the best signal for univariate analysis. Here, we present the results for the contrasts of each category versus the others (Fig. 3, SI Fig. 2). Please refer to the supplemental material (SI Fig. 1) for the more general analysis, focusing on the contrast of all visual stimuli versus baseline within each group and between groups

In Fig. 3, the results from the univariate contrast of each category versus the others for each group individually are depicted (panels A, B, C, D and E). These maps are presented at a statistically significant p -value of 0.001 uncorrected. Note that, since the [words > others] contrast did not yield any significant results, we report here the [words > houses] contrast (panel E).

For each category when compared to the others, we generally observed the activation of the known category selectivity portion of the brain: the extrastriate body area (Downing et al., 2001) for [bodies > others] emerged in all groups (Fig. 3, panel A); the contrast [face > others] elicited activation in the face network including the fusiform face area and the occipital face area (Kanwisher et al., 1997) in all groups (Fig. 3 panel B); the contrast [houses > others] activated the parahippocampal cortex (Epstein and Kanwisher, 1998) in all groups (Fig. 3 panel C); the contrast [tools > others] elicited the activity of the left lateralized tool region (Bracci et al., 2016) in controls and cataract-reversal, not in the ConBlurry 1 and 2 (Fig. 3 panel D); finally the contrast [words > houses] activated a region close to the left VWFA (Dehaene et al., 2002) in Controls, Cataract and Controls Blurry1, however in Controls and Cataract also the right counterpart showed some activity (Fig. 3, panel E).

The comparison between controls and the cataract-reversal group revealed less activation in small portions of V1 in the cataract-reversal group for the bodies (SI Fig. 2 panel A1) and tools (SI Fig. 2 panel D1) categories, with no difference observed in VOTC between the two groups in any contrast. Comparisons between the two degraded vision control conditions (ConBlurry1 and ConBlurry2) and

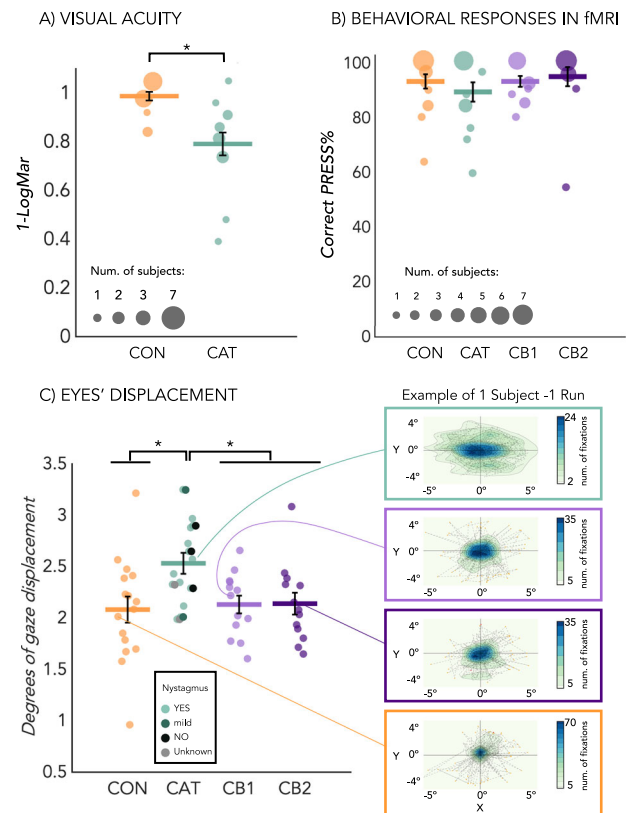


Fig. 2 | Visual acuity, behavioral response, and eye displacement across experimental conditions. **A** Visual acuity, computed as 1 minus the Logarithm of the Minimum Angle of Resolution (logMAR), is depicted for Cataract Reversals (CAT) in green and Control (CON) subjects in orange. Data are shown as mean \pm SEM across participants ($n = 16$ controls, $n = 14$ cataract-reversal individuals). Each n refers to an independent biological subject. Dot size reflects subject count, with smaller dots representing one subject and larger dots indicating up to seven subjects. Horizontal lines denote group means, and black bars signify standard errors. Asterisks mark significant differences between groups assessed with independent-samples two-sided t -tests (CAT vs CON: $t(28) = -4.1$, $p = 0.0003$). **B** Behavioral responses during fMRI are shown as mean \pm SEM across participants ($n = 16$ controls in orange, $n = 14$ cataract-reversal individuals in green, $n = 14$ Controls Blurry1 in light purple, $n = 14$ Controls Blurry2 in dark purple). Each n refers to an independent biological subject. Dot sizes correspond to subject count. **C** Eye displacement during the fMRI experiment, analyzed using deepMReye, are shown as mean \pm SEM across participants ($n = 16$ controls, $n = 14$ cataract-reversal individuals, $n = 14$ Controls Blurry1, $n = 14$ Controls Blurry2). Each n refers to an independent biological subject. Each dot represents one subject, with horizontal bars indicating group averages and vertical black lines denoting standard errors. Additionally, on the right side, an example of the deepMReye output for an average subject in each group during one run of the experiment is provided. Movement intensity is represented by color spread, with darker shades indicating more frequent fixation. Asterisks denote significant between-group differences, tested with independent-samples two-sided t -tests, FDR-corrected across comparisons (CAT vs CON: $t(28) = 2.61$, $pFDR = 0.02$; CAT vs ConBlurry1: $t(25) = -3.1$, $pFDR = 0.02$; CAT vs ConBlurry2: $t(25) = -2.78$, $pFDR = 0.02$; non-significant contrasts: $pFDR > 0.9$). Source data are provided as a Source Data file.

both the Control and Cataract groups revealed scattered differences in activation. The most prominent effects emerged for the house category, with both ConBlurry1 and ConBlurry2 showing reduced activity in the visual place network compared to both Controls and Cataracts (SI Fig. 2, panels C3–C6). Additional differences were observed for the ConBlurry2 group, which showed decreased activation in posterior occipital regions for bodies compared to both Controls (SI Fig. 2, panel A5) and Cataracts (SI Fig. 2, panel A6), and for faces when compared to Controls only (SI Fig. 2, panel B5). These

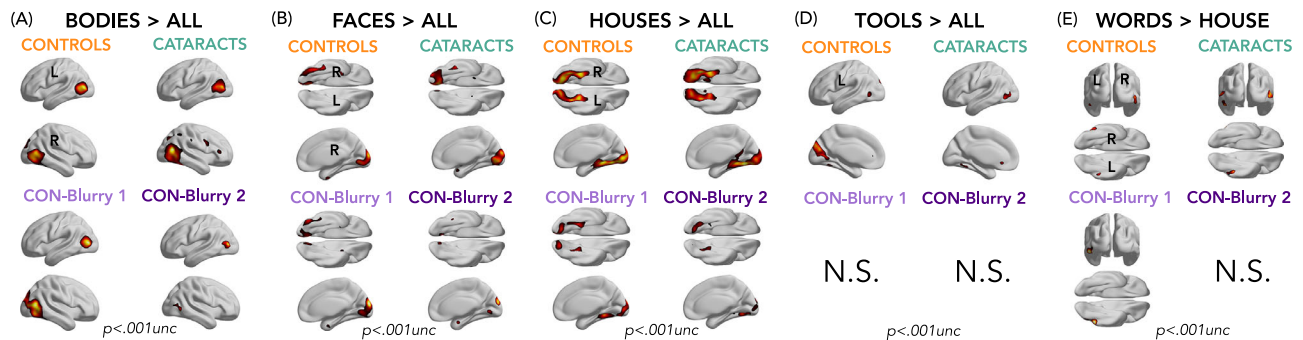


Fig. 3 | Results from univariate analyses across categories. Each column displays results for one category: bodies (A), faces (B), houses (C), tools (D), and words (E). Each plot shows within-group activation maps (voxel-wise threshold $p < 0.001$, uncorrected). For each category, we contrasted that category against all others, except for words (E), where we show words > houses because words > all did not yield suprathreshold voxels at this threshold across groups. BrainNet Viewer was

used for the visualization of brain maps²⁰. Analyses were performed in SPM12. First-level contrast images were additionally smoothed (6 mm FWHM) and entered into group-level one-sample t -tests (within groups); all tests were two-sided. Maps are shown at voxel-wise $p < 0.001$, uncorrected (no multiple-comparisons correction applied for this univariate figure). Sample sizes (biological replicates): $n = 16$ Controls, $n = 14$ Cataract-reversal, $n = 14$ Controls Blurry1, $n = 14$ Controls Blurry2.

effects were variable across comparisons and did not reveal a systematic pattern

These results, along with those concerning the general contrast [all stimuli>baseline] (see supplemental Fig. 1), do not reveal any significant differences in the activity of VOTC in cataract-reversal individuals compared to controls. This analysis is also supported by the comparison of beta values extracted from independent region of interests (see supplemental section “Univariate analysis – ROIs approach” and SI Fig. 3) and the use of Bayesian statistics (see supplemental section material section “Bayesian statistics to support similarity in brain data between controls and cataract-reversal participants”). Additionally, the findings from comparisons with control-blurry 1 and control-blurry 2 conditions are not always straightforward and are challenging to interpret. Therefore, we conducted a series of multivariate pattern analyses to further explore potential differences in categorical representation in VOTC between cataract-reversal individuals and controls, as well as to better elucidate the representation in the control-blurry 1 and control-blurry 2 conditions.

Representational similarity analysis (RSA): low vs high-levels representational models

Representational similarity analysis enabled us to simultaneously probe the low-level and categorical representations throughout the ventral occipito-temporal cortex in our participants. We correlated the brain representation of our stimuli with two representational model based either on the low-level visual features of the images (i.e., HmaxC1 model) or on their categorical membership (see Fig. 4H).

The resulting maps from the whole brain searchlight RSA analyses for both models and in each group separately are represented in the first row of Fig. 4A–D.

As expected, the HmaxC1 model elicits the higher correlation with the early visual cortex, while the categorical model is significantly correlated with VOTC activity in all groups (Fig. 4A–D).

The comparison between controls and cataract-reversals showed a significantly lower correlation in cataract-reversals with the HmaxC1 model only in V1, while there was no group differences in the correlation with the categorical model in VOTC (Fig. 4E).

In both blurry1 and blurry2 control groups, there was a reduced correlation with the HmaxC1 model in a big portion of V1 compared to the control group and a lower correlation with the categorical model in VOTC (see Fig. 4F, G).

Finally, when we directly compared the cataract-reversal group with the controls-blurry1 (Fig. 4I) and with the controls-blurry2 (Fig. 4J) we did not observe any difference in the correlation with the HmaxC1

model in V1. However, we did find a significantly lower correlation with the categorical model in VOTC for both controls-blurry1 and controls-blurry2 when compared with the cataract-reversal group.

Decoding analyses

We expanded the RSA analyses by incorporating two additional decoding analyses. One of these analyses focused on decoding categorical representation in VOTC. This analysis was implemented to strengthen our observation of no group differences using RSA with a categorical model of stimuli encoding (see Fig. 5). However, whereas RSA aims to identify maximal relations, decoding seeks to pinpoint maximal distances between patterns and, in this instance, categories. The results from the decoding analysis closely matched those obtained from the RSA analysis, providing strong support for our observation of no differences between the cataract and control group in how categories are encoded in VOTC. Due to the conceptual overlap of these findings, we decided to present the 5-way decoding between categories analysis in the supplemental material (see SI Fig. 5).

The second addition is the 6-way item decoding, applied within each category, which complements the RSA analysis with the HmaxC1 model. As we conducted five separate analyses, decoding the items within each category and including only the six items from the same category at a time, we anticipate that the algorithm’s decision-making process will rely mostly on low and middle visual properties (since the categorical membership is identical between the 6 decoded items). Our observations confirm this expectation, as we consistently found that in all groups V1 is the primary brain region encoding our stimuli across the 5 categories tested (separately) (Fig. 5, top row). You can also refer to SI Fig. 6 for an overlay of the 5 categories, demonstrating that the results are consistently similar across the different categories. These maps resulted from the whole brain searchlight 6-way decoding analyses for each group and for each category separately. No other brain regions were found to be involved in this task, even when considering lower p values.

Interestingly, when we contrast the results from different groups with independent sample T -tests, clear differences emerged. Contrasts producing significant results are reported in Fig. 5 (panels A 2-4; B 2-4; C 2-4; D 2-4; E 2-4).

The comparison between controls and cataract-reversals revealed significantly lower decoding activity in V1 among cataract-reversal individuals compared to controls. This consistent pattern of results was evident across all categories: bodies (Fig. 5, Panel A2), faces (Fig. 5, Panel B2), houses (Fig. 5, Panel C2), tools (Fig. 5, Panel D2), and words (Fig. 5, Panel E2). These findings align with the results obtained from

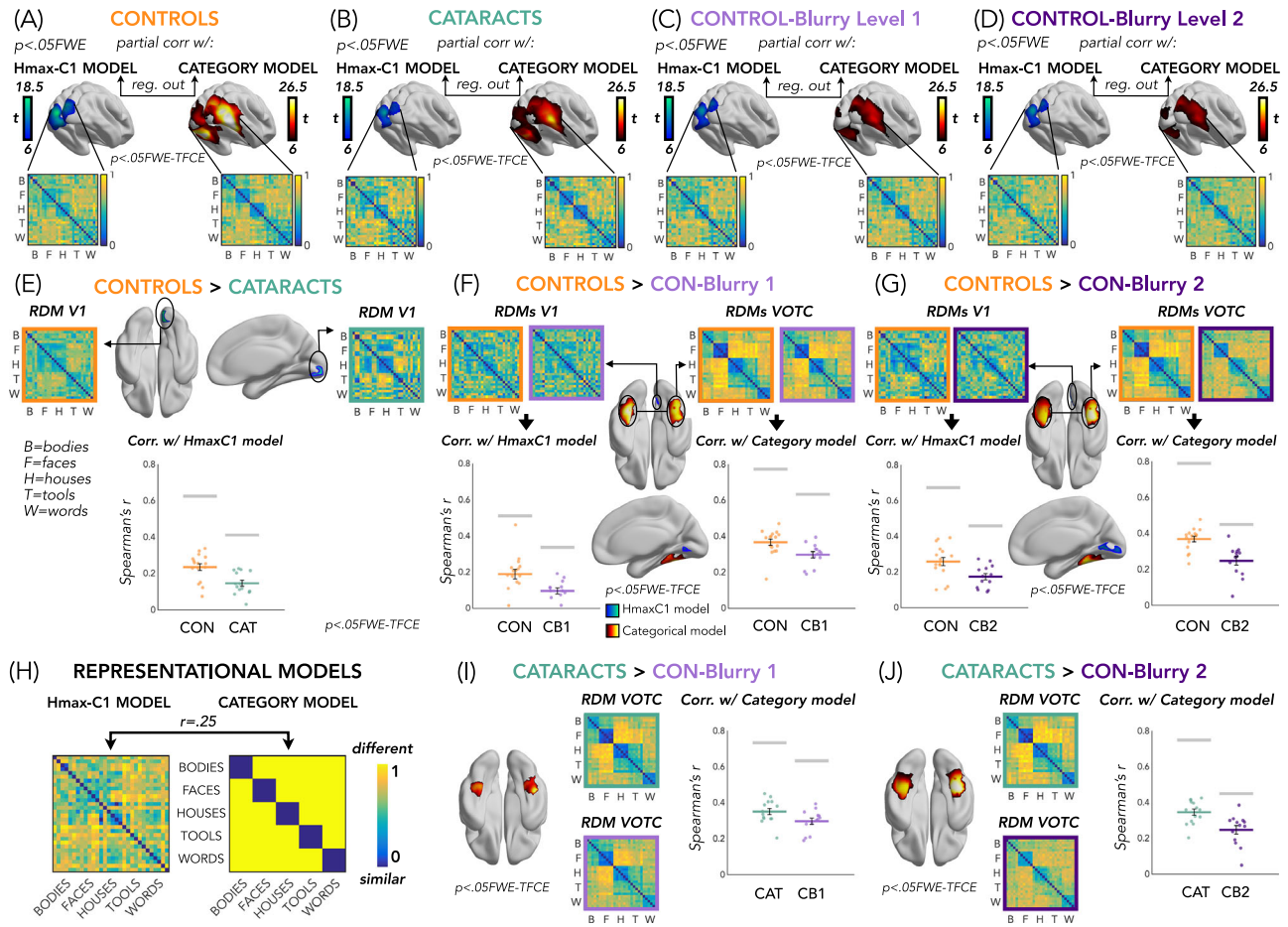


Fig. 4 | Results from Searchlight RSA Analysis with Hmax-C1 and Categorical Model. The top row presents results from the searchlight analysis within each group and condition: controls (A), Cataract-reversals (B), Controls-blurry exp 1 (C), and Controls-blurry exp 2 (D). Within each panel, Spearman's r of brain dissimilarity with the Hmax-C1 model is depicted on the left, and Spearman's r with the categorical model is depicted on the right. Below the brain maps, representational dissimilarity matrices (RDMs) extracted from significant clusters are displayed. The second row presents between-group contrasts: Controls > Cataracts (E), Controls > ConBlurry1 (F), and Controls > ConBlurry2 (G). H shows the two representational models: Hmax-C1 (left) and categorical (right). I, J show Cataract > ConBlurry1 and Cataract > ConBlurry2, respectively. For each subject and searchlight sphere (100 voxels), representational dissimilarity matrices (RDMs) were computed from stimulus-specific activation patterns and compared with the two models using Spearman's partial correlation, regressing out shared variance

between models. Correlation values were Fisher-transformed and entered into one-sample t -tests (within groups, two-sided) and two-sample t -tests (between groups, two-sided). Group contrasts were corrected for multiple comparisons using threshold-free cluster enhancement (TFCE) with family-wise error (FWE) correction applied within independent masks: V1 for the Hmax-C1 model, VOTC for the categorical model. Maps are displayed at $p < 0.05$ FWE corrected. For regions showing significant group differences, RDMs from both groups and their correlations with the two models are shown alongside dot plots. Dot plots are presented for visualization only and are not statistically tested to avoid circularity. Dot plots display individual subject data (independent biological replicates; $n = 16$ Controls in orange, $n = 14$ Cataract in green, $n = 14$ ConBlurry1 in light purple, $n = 14$ ConBlurry2 in dark purple) with mean \pm SEM. Source data for dot plots are provided as a Source Data file. Contrasts without suprathreshold voxels are not shown. BrainNet Viewer was used for the visualization of brain maps¹²⁰.

the split half analysis and RSA with the HmaxC1 model, providing robust evidence that individuals who underwent cataract reversal exhibit permanent impaired representation of low-level visual properties in V1. Indeed, they demonstrate that their primary visual cortex faces challenges in utilizing these visual features to differentiate items within the same category.

Following a similar procedure to the previous analyses, we conducted comparisons between controls and the two control conditions: controls-blurry1 (Fig. 5, third row) and controls-blurry2 (Fig. 5, fourth row). In both control conditions (blurry1 and blurry2), V1 exhibited reduced decoding accuracy compared to the control group. These results were consistent across all categories and closely resembled the findings observed in the contrast between controls and cataract-reversals. Direct comparisons between the cataract-reversal group and the control-blurry1 & control-blurry2 groups did not reveal any significant differences.

Between- and within-category decoding analyses in different layers and conditions of the DNN

Deep neural networks are increasingly used as models of human brain processing and have proven valuable for testing hypotheses about the computational principles that constrain how the brain operates.

Our fMRI results revealed a double dissociation: reduced between-category decoding in VOTC for both ConBlurry1 and ConBlurry2 (but not for cataract-reversal individuals), and reduced within-category decoding in V1 for both ConBlurry groups and cataract-reversal individuals. Therefore, we performed this final DNN analysis to assess whether a similar dissociation would emerge in the hierarchical representations of a deep neural network trained on similarly degraded visual input. This allowed us to test whether the observed brain patterns could be reproduced in a model relying solely on visual experience, thereby isolating the role of visual input quality in shaping categorical representations.

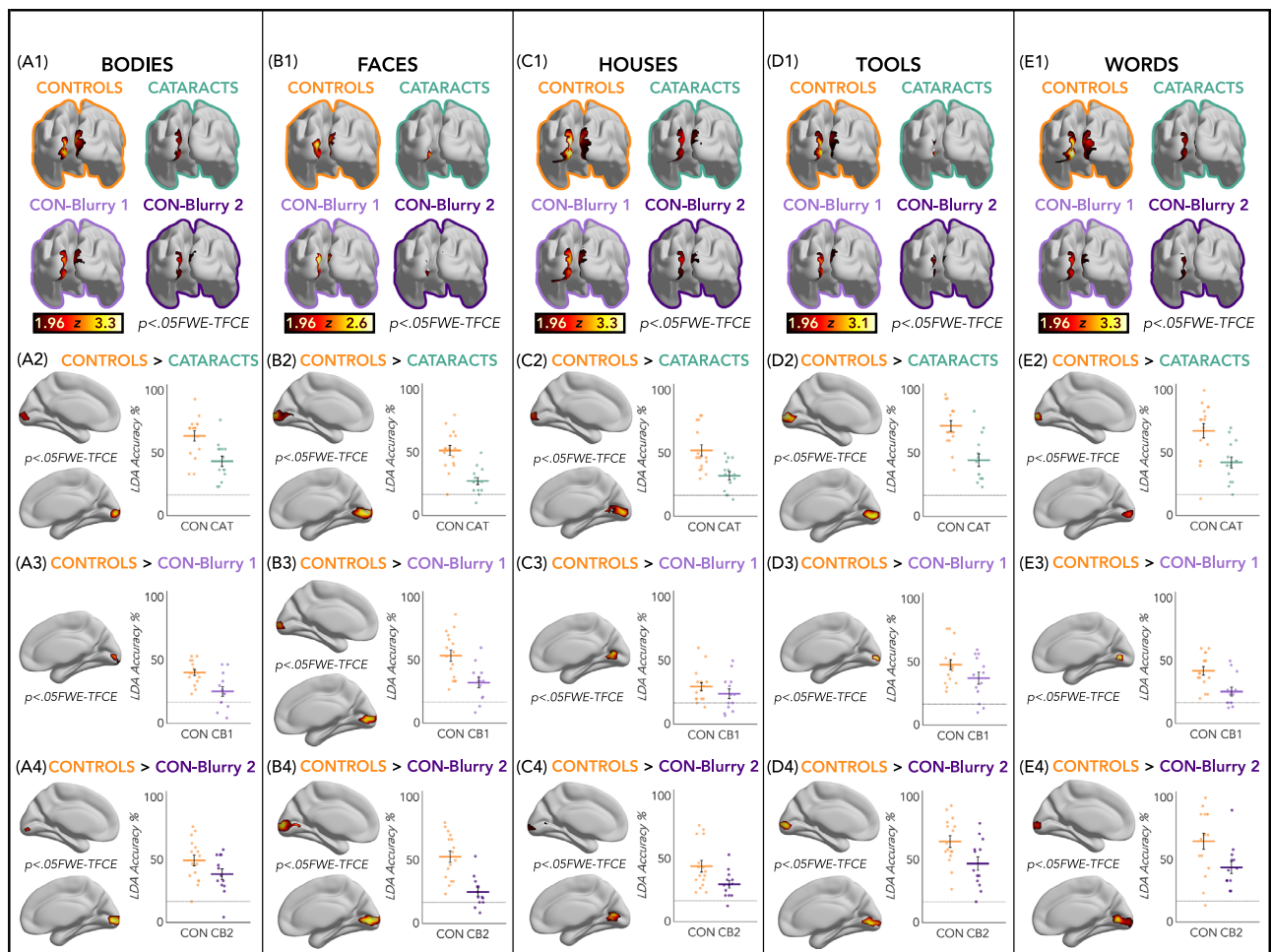


Fig. 5 | Within-category decoding analyses. Results from five different decoding analyses, each executed on a subset of the data, focusing on one category at a time: bodies (A), faces (B), houses (C), tools (D), and words (E). The top row (A1, B1, C1, D1 & E1) shows brain maps depicting regions with significant decoding accuracy for distinguishing the six within-category items within each group. The second row (A2, B2, C2, D2, E2) displays contrasts between Controls and Cataracts. The third row (A3, B3, C3, D3, E3) shows contrasts between Controls and Controls-Blurry1, and the fourth row (A4, B4, C4, D4, E4) shows contrasts between Controls and Controls-Blurry 2. Contrasts not depicted did not yield significant group differences in occipital or occipito-temporal regions. Decoding analyses were implemented using a searchlight approach (100-voxel spheres) with linear discriminant analysis (LDA). For each category, a six-way multiclass decoding was run to discriminate between the six exemplars, using a leave-one-run-out cross-validation

across runs. At the group level, voxel-wise one-sample *t*-tests (two-sided) were performed against chance for each group separately, and two-sample *t*-tests (two-sided) were used for between-group contrasts. Family-wise error (FWE) correction with threshold-free cluster enhancement (TFCE) was applied within an independent V1 mask, consistent with our a priori expectation that within-category discrimination relies primarily on low-level visual features. All maps are displayed at $p < 0.05$ FWE corrected. Dot plots below significant group differences display individual subject accuracies (independent biological replicates; $n = 16$ Controls in orange, $n = 14$ Cataract in green, $n = 14$ ConBlurry1 in light purple, $n = 14$ ConBlurry2 in dark purple), shown as mean \pm SEM. Dot plots are presented for visualization purposes only and were not statistically tested to avoid circularity. Source data for dot plots are provided as a Source Data file. BrainNet Viewer was used for the visualization of brain maps¹²⁰.

All decoding results across different DNN layers are presented in Fig. 6.

To further examine the effects of altered visual experience on categorical representation, we conducted 5-way between-category decoding analyses across the full hierarchy of DNN layers, from early (V1-like) to late (VOTC-like) representations. As shown in Fig. 6B, decoding accuracy increased progressively from early to higher layers across all conditions. However, substantial differences emerged between the three visual input regimes.

In the control-like model, decoding performance rose steeply from V1-like layer (LF1) to the final classifier layers, reaching near-ceiling accuracy by the intermediate Layer Features.7 (LF7). The cataract-like model started with low performance in the early layers but showed a marked improvement in mid-to-high layers, eventually matching control-like decoding levels by LF9 and maintaining comparable accuracy through the second-to-last layer (LC2) and even

outperforming the controls-like performance in the final VOTC-like layer (LC5). In contrast, the ConBlurry-like model also showed an initial rise in performance but plateaued at lower decoding accuracies, never reaching control-like levels at any layer, despite outperforming the cataract-like model in the Layer Features.4 (LF4).

Statistical comparisons confirmed these observations. Significant differences (all $p < 0.001$, FDR corrected) were observed between the control-like and ConBlurry-like models at all layers despite the first V1-like layer, with the ConBlurry-like model consistently showing lower decoding accuracy. The cataract-like model, in contrast, began with significantly reduced decoding at early layers (LF1, LF4 and LF7) but caught up with control-like performance by LF9, with no significant difference between cataract-like and control-like models at LF9 and LC2, and even significantly higher accuracy in Cataracts-like compared to Control-like condition in the final LC5 layer. Interestingly, although ConBlurry-like performance was slightly higher than cataract-like in

recognition system. Using a multivariate framework incorporating human brain imaging and artificial neural networks, we examined the impact of visual deprivation early in life on the development of the visual object recognition system. Our findings revealed a preserved categorical representation in the ventral visual pathway (VOTC) despite a brief and transient period of visual deprivation early in life and permanent alterations in the response properties of the early visual cortex (EVC) for low-level visual features.

The EVC of neonates demonstrates an advanced level of connectivity comparable to the one observed in adults^{36–38}, exhibits ocular and orientation columnar organization^{39,40} displays selectivity for visual features similar to adult levels⁴¹ and advanced retinotopic organization²⁶. Moreover, the eccentricity-related resting-state connectivity profile of EVC appears roughly preserved in congenitally and permanently blind adults^{42,43}. Do these findings suggest that EVC development occurs independently of sensory experience? Our results demonstrate otherwise by revealing that cataract-reversal individuals have altered representation of low-level visual features, as measured by reduced RSA correlation with a computational EVC-like model (Fig. 4) and by reduced item-decoding analyses within each category (Fig. 5; see also the supplemental analysis: *Between-groups correlation of V1 and VOTC brain dissimilarity matrices* and SI Fig. 8). These results align with former animal studies indicating lasting functional deficits in EVC properties in kittens, monkeys, and mice exposed to early and transient visual deprivation^{2,40,44}. Studies involving human participants after a transient period of early visual deprivation also suggested lasting deficit in low- and mid-level visual abilities^{45–47} and alteration in the response properties of EVC¹⁴. Altogether, our results show how early visual experience plays a crucial role in maintaining and refining the functional development of early visual circuits, despite a prototopographic organization at birth^{14,48}.

In contrast to the notion that EVC displays a relatively high level of functional and structural maturity at birth^{37,39,41}, the extrastriate visual system has long been thought to display rather immature neonatal functional and anatomical organization^{49–52}. Combined with the idea that downstream higher-level regions acquire their functional tuning based on structuring inputs they receive from earlier regions in the cortical hierarchy, this led to the assumption that visual deprivation induces a cascaded pattern of impairments, with its magnitude increasing as a function of the synaptic distance from the earlier visual input⁵³. For instance, recent studies on human cataract-reversal individuals reported more altered visual processing in extrastriate compared to striate brain regions^{15,18,54}, suggesting that the functional impact of visual deprivation becomes more prominent downstream in the visual processing hierarchy⁵⁵. A preliminary report even suggested that monkeys raised in darkness for a year show no impairment in EVC but impairment in VOTC, where activity for visual stimuli is present but not selective for specific categories⁵⁶. Our findings challenge these assumptions. Despite clear differences in EVC response properties, we did not observe differences between the cataract-reversal group and control group in VOTC response. In this region, the reliability of activity patterns for each item was equivalent between the two groups (split-half correlation measurements; see SI Fig. 4), and we observed a comparable categorical representation expressed by univariate contrasts (Fig. 3, SI Fig. 2, SI Fig. 3), RSA correlation with a categorical model (Fig. 4), and multiclass decoding analyses between categories (Fig. SI 5). Such convergence in results using multiple analytical streams clearly challenges the view that visual areas most mature at birth are the most resilient to experience-dependent plasticity^{26,57}. Our study instead suggests that early postnatal visual experience is not a prerequisite to develop categorical selectivity in VOTC. These results also interrogate the notion that categorical domain formation is primarily influenced by early visual experience¹⁶.

An important difference between our and previous studies is the length of visual deprivation. We examined the impact of a brief period

of postnatal visual deprivation, lasting approximately 100 days on average. By contrast, previous studies examined significantly longer periods of deprivation in both humans^{14,15} and primates^{53,56}. Crucially, the time between the eyes opening and the test in previous studies was significantly shorter than the several years that elapsed between surgery and testing in our study (Table 1). In our study, we tested cataract-reversal individuals many years after surgery. This extensive visual experience might have contributed to the categorical representation in VOTC. Indeed, previous behavioral studies have documented a gradual recovery of many visual abilities following congenital cataract surgery. For example, visual acuity in children treated early continues to improve until about age two, even if it doesn't reach typical levels^{10,46}. Longitudinal studies of individuals treated after extended blindness also show that, despite some low-level impairments (e.g., reduced acuity), these individuals demonstrate proficiency in many mid- and high-level visual tasks, including shape matching, visual memory, image segmentation, and face discrimination, localization & identification, and gender classification^{12,58}. This behavioral recovery suggests parallel recovery at the neural level, particularly in brain regions involved in these visual functions, which would include VOTC. Longitudinal studies examining brain changes in various regions of the visual cortex following cataract removal would be essential to understand the mechanisms supporting this recovery trajectory at the brain level.

Importantly, previous studies involving cataract reversal participants with similar period of transient deprivation reported behavioral impairments in fine-grained categorical processing^{19–21}. These impairments may stem from initial visual analysis deficits in EVC or potential impairments in VOTC representation at more detailed levels. We do not rule out the possibility of VOTC impairments in more fine-grained representations of visual categories, particularly in tasks that require feature binding^{47,59}, such as holistic face processing. For the face domain specifically, previous research has indeed suggested that the neural processes involved in discerning fine details of a specific face may vary from those implicated in recognizing generalized facial configurations^{60,61}. This warrants investigation in future studies.

Our results support the existence of varying sensitive periods along the visual processing hierarchy, leading to distinct developmental impacts of sensory deprivation^{62–64}. This aligns with the idea that different visual skills—and the brain regions supporting them—may vary in their susceptibility to early visual deprivation¹² supporting the existence of multiple sensitive periods across different brain regions^{62–64}. However, which brain regions are more or less affected by early visual deprivation remains controversial. In striking contrast with the idea that impairments increase from EVC to more downstream regions⁶⁴, we instead found that early blindness compromises the representation of low-level visual information in EVC, while visual categorical representation in VOTC appears unaffected.

To explore whether EVC impairment was due to early visual deprivation or current visual impairment at the time of testing, we included several control analyses. First, we extracted and regressed out information on each subject's eye movements during the experiment (Fig. 2C) to minimize data variance due to different eye movement profiles across groups. Correlation analyses between brain data and visual acuity values for the cataract-reversal individuals was not significant, suggesting that the observed EVC results were not linked to visual acuity at the moment of testing. To directly assess whether the neural profile observed in the cataract-reversal group could be explained by degraded visual input at the time of testing, we included two control conditions (CON-Blurry1 and CON-Blurry2) in which the original stimuli were manipulated to simulate reduced acuity and nystagmus in typical participants. These conditions mimicked the visual experience of cataract-reversal individuals while preserving typical visual development. Notably, both blurry conditions led to degraded representations in both EVC and VOTC, indicating that acute

visual degradation affects both early and high-level visual areas. In contrast, cataract-reversal individuals showed a selective impairment in EVC but preserved categorical representations in VOTC. This dissociation strongly suggests that their neural profile cannot be explained solely by current visual input quality, but rather reflects long-term developmental adaptations following early visual deprivation.

These results are also supported by recent findings demonstrating that cataract-reversal individuals exhibit long-term plasticity in late-visual white matter pathways, beyond what is expected from typical maturation⁶⁵. Importantly, these structural changes were predictive of behavioral recovery. Altogether, this suggests that visual experience, albeit degraded, can drive functional and structural adaptation in high-level visual regions.

To better understand the mechanisms underlying the preserved categorical organization in the ventral occipito-temporal cortex (VOTC) of cataract-reversal individuals, we turned to deep neural networks (DNNs) as a computational model of hierarchical visual processing. DNN are not perfect models of their biological counterparts (for a nuanced review, see ref. 66). Nevertheless, they are among the most successful models available for predicting human visual behavior and neural responses across the visual hierarchy^{67–70} and can thereby serve as valuable approximations of visual processing mechanisms⁷¹. Most crucial for the results in this paper, these systems offer a systematic methodology for directly examining the consequences of controlled manipulations of sensory experience that experiments with humans, for both ethical and practical reasons, do not provide. More precisely, DNNs allow us to simulate different types of visual experience and deprivation with precise control, offering a powerful way to test hypotheses that are difficult to address directly in clinical populations. Therefore, the DNN analyses aim to shed light on how categorical representations in VOTC can emerge or be preserved despite an early period of visual deprivation. One possibility is that, following sight restoration, individuals with congenital cataracts gradually learn to associate degraded or impoverished visual inputs with semantic or categorical structure, leveraging experience-driven learning mechanisms. DNNs provide a useful framework for testing this hypothesis: by simulating altered visual histories and observing whether high-level categorical representations can still emerge, we can assess whether degraded input alone is sufficient to build structured representations akin to those found in VOTC.

We first investigated whether hierarchical processing in DNNs could replicate key properties of the human visual system. As expected, and consistent with previous findings³⁴, early DNN layers showed higher representational similarity with brain activity in EVC, while deeper layers correlated more with anterior regions of the ventral stream. In line with our results, the EVC-like layer representations were more similar to EVC activity in the control group than in cataract-reversal or visually degraded (ConBlurry) conditions. Conversely, in the VOTC-like layer, representations in the cataract-reversal model were comparable to those of controls, while the visually degraded model showed weaker similarity (SI Fig. 9B). These findings suggest that DNNs can capture not only the hierarchical structure of visual representations but also the functional impact of different visual experiences, setting the stage for further modeling of deprivation scenarios.

Building upon this similarity, we implemented three network variants reflecting the visual experiences of our human groups in the fMRI experiment (Fig. 6A): a Control-like network (trained and tested on clear images), a Cataract-like network (trained and tested on blurred images), and a ConBlurry-like network (trained on clear images and tested on blurred images). These conditions allowed us to evaluate the distinct effects of early visual deprivation followed by recovery versus acute input degradation at test.

In line with the fMRI data, both Cataract-like and ConBlurry-like models showed reduced between-category decoding accuracy in early layers, mirroring impairments in EVC. However, a key difference emerged in deeper layers: the Cataract-like model progressively improved and reached Control-like performance by layer LF9, suggesting that prolonged experience with intact input following early deprivation can support the development of categorical representations. In contrast, the ConBlurry-like model remained impaired at all levels, indicating that even temporary input degradation at test can disrupt categorical structure, despite normal prior experience (Fig. 6B).

This distinction was also evident in the within-category decoding analyses (Fig. 6C), which were restricted to early-to-middle layers (LF1 to LF9) based on our biological data implicating early visual cortex in this process. While the Cataract-like model improved with depth and matched Control-like performance by LF9, the ConBlurry-like model consistently showed lower accuracy in every layer, across all categories. These results suggest that despite early impairments in low-level visual processing, extended visual experience can restore the capacity for fine-grained visual discrimination, whereas acute degradation, even in the absence of developmental deprivation, limits representational precision.

Together, these findings support the idea that long-term visual experience plays a critical role in shaping and recovering high-level visual representations, and that DNNs can effectively model how different types of input quality influence representational outcomes. By distinguishing between the effects of degraded input experienced throughout life versus at test only, our DNN analysis helps clarify why categorical selectivity in VOTC may be preserved in cataract-reversal individuals but not in individuals experiencing acute visual degradation. More generally, these results also reveal that DNNs represent an interesting approach to evaluate the consequence of an early and transient period of deprivation and therefore pave the way to manipulate various features of deprivation that are notoriously difficult to control in humans (e.g., exact period and severity of deprivation) to generate new hypotheses on the functional consequences of sensory deprivation⁷².

While experience-driven learning from degraded input following sight restoration offers one plausible explanation, there is also a complementary, though not mutually exclusive, possibility. The maintenance of categorical selectivity in VOTC despite early visual deprivation and EVC functional impairment could potentially be related to its connectivity with downstream regions that could exert a top-down regulation for the development of the categorical coding of VOTC^{73–79}. Categorical domains within the VOTC might have an inherent predisposition to process these specific categories such as faces, words, landscapes through specialized connections to temporoparieto-frontal networks related to social cognition, linguistic processing, and spatial interaction with the environment, respectively^{80–82}. Indeed, besides enabling object recognition, VOTC is likely a foundational brain region in determining the relevance of objects to behavior, for instance to engage in social interaction, object manipulation and in navigating the environment⁸³. Such pattern of large-scale connectivity might explain in part the maintained categorical coding of VOTC despite deprivation and EVC alteration⁸².

Early visual deprivation triggers crossmodal activity for touch or sounds in EVC^{84–86}, even years after visual restoration^{87–89}, and sound stimulation can suppress early visual responses in cataract reversal individuals⁹⁰. Crossmodal responses during a sensitive period could occupy synaptic space in the EVC, potentially interfering with visual recovery. In VOTC, crossmodal recruitment by sound in blind individuals appears to follow the category-selective profile of the region and is also partially observable in sighted individuals^{75,91–93}, suggesting that it may promote functional maintenance^{94,95}, rather than interfere with sight recovery. In other words, the selective alteration of visual

function in specific regions might also relate to region-specific impact of crossmodal reorganization triggered by early blindness⁸⁴.

By showing how a brief period of early-life visual deprivation permanently affects information encoding in EVC while leaving the categorical coding in VOTC intact, our study highlights how distinct regions in the human object visual recognition system are differently affected by a transient period of postnatal blindness. Our results challenge the conventional belief that high-level brain regions predominantly depend on early visual experiences and intact downstream visual input for their development, and instead show how the development of the categorical coding in VOTC shows resilience to deprivation and altered upstream input. The inclusion of CON-Blurry1 and CON-Blurry2 control conditions was essential to isolate the developmental effects of early visual deprivation from the immediate impact of poor visual input. These conditions mimicked the reduced acuity and nystagmus of cataract-reversal individuals in participants with typical development. The distinct profile observed in the cataract-reversal group – impaired EVC and spared VOTC – stands in contrast to the two control (blurry) conditions, who exhibited widespread degradation in EVC and VOTC, thereby reinforcing the interpretation that the VOTC in cataract-reversal individuals has undergone compensatory developmental mechanisms to preserve high-level visual representations.

Our study offers a crucial refinement to prevailing models of visual development, highlighting that sensitive periods are not fixed boundaries but interact with the type and timing of input, and that later visual experience can still support the emergence of robust categorical coding in the human brain.

Our work provides neurobiological and modeling evidence supporting the conclusion that the notion of critical periods needs to be defined differently for different visual skills, rather than as a unitary construct applicable to all aspects of visual development.

Methods

Participants

Fifteen cataract-reversal individuals (7 male, 8 female; mean age \pm SD = 28.26 \pm 6.32 years, range = 18–39 years; average length of deprivation \pm SD = 69 \pm 48 days; see Table 1) and seventeen matched controls (11 male, 6 female; mean age \pm SD = 29.12 \pm 5.71 years, range = 21–39 years) participated in the experiment. An additional group of fourteen matched controls (7 male, 7 female; mean age \pm SD = 27.86 \pm 4.91 years, range = 21–39 years) took part in a control experiment in which the visual properties of the stimuli were altered. All participants reported their sex/gender as part of routine demographic information. Sex/gender was not used as a factor in the study design because we had no specific hypotheses regarding sex- or gender-related effects, and our sample size was insufficient to perform meaningful sex- or gender-based analyses. All participants were naïve to the purpose of the experiment, provided written informed consent prior to participation, and were monetarily compensated. The study was approved by the research ethics committees of the University of Toronto and McMaster University. Prior to testing, participants completed a training session to familiarize themselves with the tasks. Based on poor performance during fMRI acquisition, we excluded one participant from the control group, one from the cataract-reversal group, and one from the control-experiment group (see Behavioral Results).

Assessment of visual acuity

The visual acuity was acquired for the cataract-reversal (see Table 1) and the control subjects undergoing the main experiment using the Freiburg Visual Acuity Test^{96,97}. Participants were tested at a distance of 210 cm from the stimulation monitor (CRT monitor of screen width of 53.5 cm, and screen resolution of 800 \times 600 pixels). The test run comprised 30 Landolt-C trials that were presented in a random orientation. The size of the stimuli was adjusted on each trial based on

the previous responses to estimate the most probable visual acuity threshold by using a maximum-likelihood staircase procedure⁹⁸. The participants performed a forced choice task where they indicated the orientation of the C-stimulus gap (up, down, right, left). The Logarithm of the Minimum Angle of Resolution (logMar) was used as a measure of visual acuity. During this visual acuity test, the subjects wore the same corrective glasses they used in the scanner during the fMRI experiment. Therefore, the visual acuity measured represents a corrected level, reflecting the maximal acuity the subjects could achieve at the time of the test. The visual acuity of the two groups were compared using a two-sample *t*-test. Furthermore, the visual acuity values were used to conduct correlations with brain data.

Stimuli and procedure for fMRI experiment

Visual stimuli were projected on a screen (frame rate: 60 Hz; screen resolution 1920 \times 1080 pixels) behind the scanner. Participants viewed the screen (distance from head = 45 cm) through a mirror mounted on the MRI head coil. Participants performed a task during the experiment by responding with two MR-compatible response buttons.

The stimulus set included 6 images in each of 5 different categories: bodies, faces, houses, tools and words (30 images; see Fig. 1A for illustrative examples of the stimulus set). The images were black and white pictures (500 \times 500 pixels) collected from internet. They were placed in the center of the screen on a gray (129 RGB) background.

Before entering the scanner, each participant was familiarized with the stimuli to ensure perfect recognition. In the fMRI event-related experiment each trial consisted of the same stimulus repeated twice. Rarely (10% of the occurrences), a trial was made up of two different consecutive stimuli (catch trials). Only in this case, participants were asked to press a key with the right index finger if the second stimulus belonged to the living category and with their right middle finger if the second stimulus belonged to the non-living category. This procedure ensured that the participants attended and processed the stimuli. Each pair of stimuli lasted 2 s (850 msec per stimulus interleaved with 300 msec of a blank screen) and the inter-trial interval (i.e., fixation cross) was 2 s long for a total of 4 s for each trial (see Fig. 1B). Within the fMRI session participants underwent 5 runs, with the exception of one control subject that, due to technical issues, underwent only 4 runs. Each run contained 2 repetitions of each of the 30 stimuli, 6 catch trials and two 20s-long fixation periods without stimuli (one in the middle and another at the end of the run). The total duration of each run was 304 s. The presentation of trials was pseudo-randomized: two stimuli from the same category (i.e., bodies, faces, houses, tools and words) were never presented in subsequent trials. The stimuli delivery was controlled using Matlab R2016b (<https://www.mathworks.com>) Psychophysics toolbox (<http://psychtoolbox.org>).

In the control experiment that we ran in a separate group of participants, the visual properties of the images were altered to mimic the lower acuity and nystagmus of cataract-reversal subjects. Aside from image alteration, the procedure was identical to the one described for the original experiment. Nystagmus is a condition characterized by involuntary, rhythmic eye movements that can occur in people who have been treated for bilateral congenital dense cataract (10 of our cataract-reversals participants had a nystagmus, see Table 1). To simulate the nystagmus of cataract-reversal participants we applied to the visual stimulation a pendular movement in the horizontal plane. The like-nystagmus movement was applied to the image for the total time of the presentation (i.e., 850 msec) with a frequency of 3.5 Hz. The maximal displacement of each image was of 150 pixels, corresponding to 3.23 degrees of visual angle (corresponding to the larger gaze's displacement among the cataract subjects, see Fig. 2C). Therefore, each image was shifting of 3.23 degrees in the horizontal direction for 3 times during the 850msec of presentation. To mimic the lower acuity of the cataract-reversals group we blurred the original stimulus set,

applying a lowpass filter. We selected 2 different cutoff levels: 10 cycles per degree (CPD) corresponding to a visual acuity of 20/60 (average acuity level of our cataract-reversal group corresponding to a logMar of 0.21) and 4.8 CPD corresponding to an acuity of 20/125 (lower acuity level of the cataract-reversal group corresponding to a LogMar of 0.61). Therefore, we created two versions of the control experiment. In both versions the nystagmus was applied in the same way, while the level of blurring was either 10 CPD (CON-Blurry 1) or 4.8 CPD (CON-Blurry 2). These control conditions were included to disentangle the effects of degraded visual input at testing from the long-term developmental consequences of early visual deprivation. All these stimuli, without any compression, can be found in the OSF project (<https://doi.org/10.17605/OSF.IO/BECDR>) at this [link](#). A separate group of 14 subjects took part in both versions of the control experiment. Each subject performed 5 runs of the control experiment-blurry 1 and 5 runs off the control experiment- blurry 2.

MRI data acquisition

Structural and functional data were acquired at the Sherman Health Science Research Center (York University, Toronto, Canada) using a Siemens MAGNETOM 3 T PrismaFit MRI scanner with a standard 64 channel head coil. Structural high-resolution T1-weighted MPRAGE scans were acquired using parallel imaging (GRAPPA factor = 2). The acquisition parameters for the T1 images included a repetition time (TR) of 2300 ms, echo time (TE) = 2.62 ms, voxel size = 1 mm isotropic voxels, number of slices=192, and a flip angle of 9°. Functional scans were acquired using T2*-weighted echoplanar BOLD imaging. The acquisition parameters included simultaneous interleaved multi-slice acquisition using parallel imaging with multiband acceleration = 2, phase encoding acceleration = 3, TR = 1170 ms, TE = 30 ms, voxel-size = 2 mm isotropic, number of slices = 51, flip angle = 66°, and echo-spacing 0.68 ms. The first four initial scans of each run were discarded to allow for equilibrium magnetization.

fMRI preprocessing

fMRI data was preprocessed in statistical parametric mapping software (SPM12 – Wellcome Department of Imaging Neuroscience, University College London, UK) implemented in Matlab R2016b (Mathworks, Inc.). Preprocessing steps included slice time correction, EPI alignment to the mean functional image with a 2nd degree B-spline interpolation, co-registration of the functional volumes to the structural image and normalization to the Montreal Neurological Institute (MNI) template. For the univariate analysis only, we also performed a spatial smoothing with a Gaussian kernel of FWHM of six millimeters on the volume time series.

Eye movement extraction from fMRI data

Ten out of our 16 cataract-reversal participants presented nystagmus. To control for the possible different pattern of eye movement in our groups, we used DeepMReye, an open source framework for eye-tracking that does not require a camera. It is based on a convolutional neural network (CNN) that reconstructs the viewing behavior of individuals directly from the MR signal of their eyeballs³⁵. At each image acquisition, DeepMReye captures the multi-voxel pattern of the eyes and use a CNN to predict the gaze location based on that pattern. To train the CNN, the researchers required an independent measure of gaze location, which was obtained from previous studies using camera-based eye tracking or fixation targets (see below for further details). Notably, although independent gaze information was necessary for the CNN training, it is not required when applying the trained CNN to new data⁹⁹.

To obtain the “eye-tracking” results from our fMRI data, we performed the preprocessing using bidsMReye (version 0.3.0+24.gbe1f5da.dirty), a BIDS app relying on deepMReye (@deepmreya) to decode eye motion from fMRI time series data.

The data of each BOLD run underwent co-registration conducted using Advanced Normalization Tools (ANTs, RRID:SCR_004757) within Python (ANTsPy). First, each participant’s mean EPI was non-linearly co-registered to an average template. Second, all voxels within a bounding box that included the eyes were co-registered to a preselected bounding box in our group template to further improve the fit.

Each voxel within those bounding box underwent two normalization steps. First, the across-run median signal intensity was subtracted from each voxel and sample and was divided by the median absolute deviation over time (temporal normalization). Second, for each sample, the mean across all voxels within the eye masks was subtracted and divided by the standard deviation across voxels (spatial normalization). Importantly, this method can decode gaze position at a temporal resolution higher than the one of the imaging protocol (sub-TR resolution)³⁵.

Voxels time series were used as inputs for generalization decoding using a pre-trained model 1 to 6 from deepMReye from OSF. This model was trained on the following datasets: guided fixations (@alexander_open_2017), smooth pursuit (@nau_real-motion_2018, @politi_rapid_2022, @nau_hexadirectional_2018), free viewing (@julian_human_2018).

For each run the following values were computed: (1) the variance for the X gaze position; (2) the variance for the Y gaze position; (3) the framewise gaze displacement; (4) the number of outliers for the X gaze position; (5) the number of outliers for the Y gaze position; (6) the number of outliers for the gaze displacement. Outliers were robustly estimated using an implementation of @carling_resistant_2000.

We used the average amount of gaze displacement across runs, after excluding the outliers, to test with independent samples T-tests whether there was a difference in the amount of eye movement between our groups (Fig. 2C).

We included the variance for the X and Y gaze positions as regressors of no-interest in our GLM (see next paragraph), to control for the impact of eye movement on the fMRI data activity.

General linear model

The pre-processed images for each participant were first analyzed using a general linear model (GLM).

For the univariate analyses, for each subject, a design matrix was formed using a predictor for each stimulus category (bodies, faces, houses, tools, words) in each run. These regressors of non-interest were also added: 1 regressor of no-interest for the catch trials, 6 head-motion regressors of no-interest, 2 eye movement regressors of no-interest (the variance for the X gaze position and the variance for the Y gaze position) and 1 constant.

For the multivariate analyses, for each of the 5 runs we included 40 regressors: 30 regressors of interest (each stimulus), 1 regressor of no-interest for the catch trials, 6 head-motion regressors of no-interest, 2 eye movement regressors of no-interest (the variance for the X gaze position and the variance for the Y gaze position) and 1 constant. From the GLM analysis we obtained a β -image for each stimulus (i.e., 30 pictures) in each run, for a total of 150 (30 × 5) beta maps.

Statistical procedure and brain masks for TFCE correction

For each one-sample *t*-test within-group and two-sample *t*-tests performed to compare the effects between groups, results were corrected using the non-parametric threshold free cluster enhancement (TFCE) method combined with a FWE correction¹⁰⁰.

All univariate and searchlight analyses in the paper were run whole brain. But the threshold free cluster enhancement (TFCE) correction for the two-sample *t*-tests was applied within specific masks, based on a-priori hypotheses. We were a priori interested in how our visual stimuli were processed for their low-level visual features and for their high-level categorical visual properties. We therefore created a mask including the primary visual cortex (V1) and the ventral occipito-

temporal cortex (VOTC) by combining these regions from the JuBrain Anatomy Toolbox bilaterally (a.k.a. SPM Anatomy Toolbox - v2.2b¹⁰¹; the *Human occipital cytoarchitectonic area 1*, *hOc1*¹⁰², *Area FG1* situated in the posterior and medial regions of the fusiform gyrus, *Area FG2* located in the posterior fusiform gyrus and extending to the occipito-temporal sulcus, *Area FG3*, anterior to FG1, placed in the medial fusiform gyrus and expanding till the collateral sulcus¹⁰³ and *Area FG4*, anterior to FG2, including the lateral portion of the fusiform gyrus (FG) and extending to the occipito-temporal sulcus¹⁰³. For a similar parcellation of the human ventral visual stream see also ref. 104. These masks can be found at this [link](https://doi.org/10.17605/OSF.IO/BECDR), on the OSF project (<https://doi.org/10.17605/OSF.IO/BECDR>) linked to this work.

We chose not to average multivariate values within those anatomical ROIs (V1 and VOTC), but instead used searchlight analysis to preserve the spatial richness of the data. Nonetheless, since our focus, especially for groups' comparisons, was on early visual cortex (V1) and occipito-temporal cortex (VOTC), we applied these masks for multiple comparisons correction, in order to constrain inference to theoretically relevant areas.

The correction was applied either across the full mask, including bilateral V1 and bilateral VOTC (for univariate analysis and split-half analysis), solely in the VOTC mask (for RSA with categorical model and 5-way categorical decoding), or exclusively in the V1 mask (for RSA with Hmax-C1 model and within-category decoding), based on a priori hypotheses that are explained in each respective section below.

Univariate analysis

Using SPM12 we computed within and between groups contrast maps comparing (1) all visual stimuli vs baseline; (2) bodies vs all other categories; (3) faces vs all other categories; (4) houses vs all other categories; (5) tools vs all other categories; (6) words vs all other categories. Since the last contrast did not show any activation across groups, even at a low correction threshold, we also ran and reported the contrast *words > houses*. These contrast images were further spatially smoothed with a Gaussian kernel of 6 mm FWHM before conducting group-level analyses. They were input into a series of corresponding one-sample t-tests within each of the four groups and two-sample t-tests between each possible group comparison (SI Fig. 1 for visual stimuli vs baseline and Fig. 3 & SI Fig. 2 for all other contrasts). Maps are shown at voxel-wise $p < 0.001$, uncorrected.

Since several regions within the VOTC are known to show category-specific selectivity we conducted a supplemental analysis to examine whether activation in these areas was reduced in cataract-reversal patients compared to controls. Specifically, we extracted beta values from 12 predefined ROIs: left and right EBA, OFA, FFA, PPA, and TOS, as well as the left LOC and left VWFA. We then performed both frequentist and Bayesian statistical analyses on these values to assess potential group differences. Further methodological details and the full results of this analysis are reported in the Supplementary Material (For frequentist statistics see supplemental section 'Univariate analysis – ROIs approach' & SI Fig. 3. For Bayesian statistics, supplemental section: 'Bayesian statistics to support similarity in brain data between controls and cataract-reversal participants', SI fig 7 and Supplemental Table 1).

Split-half analysis

We ran a split-half analysis to investigate how reliable/stable were the patterns of activity produced by the visual stimuli in both cataract and control subjects and in the two versions of the control experiment. Further methodological details and results are provided in the Supplementary Information in the section 'Split-half analysis and SI Fig. 4).

Representational similarity analysis (RSA): high-level vs low level representational models

We further assessed whether the representational content encoded in the occipito-temporal cortex differed across groups using RSA^{31,105}.

This analysis was performed using the CoSMoMvPA¹⁰⁶ toolbox, implemented in Matlab R2016b (Mathworks). RSA is based on the concept of representational dissimilarity matrix (RDM): a square matrix where the columns and rows correspond to the number of the conditions (i.e., n 30 image stimuli, therefore a matrix of 30×30 in this experiment) and it is symmetrical about a diagonal of zeros. Each cell contains the dissimilarity index between two stimuli¹⁰⁷. This abstraction from the activity patterns themselves represents the main strength of RSA, allowing a direct comparison of the information carried by the representations in different groups and between brain and models^{31,108}. Crucially for the present study, RSA allowed us to use the same set of visual stimuli to isolate both the categorical/high-level and the visual/low-level representations all along the ventral occipito-temporal cortex of our participants. For each 100-voxels searchlight sphere and in each subject, we extracted the RDM, computing the dissimilarity between the spatial patterns of activity for each pair of conditions/images. To do so, we extracted the stimulus-specific BOLD estimates from the contrast images (i.e., SPM T-maps) for all the 30 image stimuli separately. Then, we used Pearson's correlation to compute the distance (i.e., $1-r$) between each pair of patterns.

Since the RDMs are symmetrical matrices, for all the RSA analyses we only used the upper triangular RDM, excluding the diagonal to avoid inflating correlation values¹⁰⁹.

Then, we used Spearman's partial correlation to compare the selective brain representation of each searchlight sphere with two representational models: a categorical and a low-level visual model.

The categorical RDM (Fig. 4H right side) assumes that image stimuli from the same category gather together into 5 distinct clusters representing the 5 main categories (i.e., (1) bodies, (2) faces, (3) houses, (4) tools, (5) words).

The Hmax model is a computational model of object recognition in the cortex that has been designed to reflect the hierarchical organization of the visual cortex¹¹⁰ in a series of layers from V1 to inferotemporal (IT) cortex^{111,112}. To build our low-level visual model we used the output from the V1- complex cells layer (also called C1 layer¹¹². The inputs for the model are the gray-value luminance images presented in the experiment. Each image is first analyzed (i.e., filtered) by an array of simple cells (S1) units at 4 different orientations and 16 scales. At the next C1 layer, the image is subsampled through a local Max pooling operation over a neighborhood of S1 units in both space and scale, but with the same preferred orientation (Serre et al.¹¹²). C1 layer stage corresponds to V1 cortical complex cells, which show some tolerance to spatial shift and size (Serre et al.¹¹²). The outputs of all complex cells were concatenated into a vector as the V1 representational pattern of each image^{31,113}. We, finally, built the (30×30) RDM computing 1-Pearson's correlation of each pair of vectors (Fig. 4H left side). The Hmax-C1 RDM was significantly correlated with the categorical RDM ($r = 0.25$, $p < 0.001$). The use of partial correlation between the brain RDMs and the models, allowed us to regress out this shared correlation between the two models and to look at the unique correlation between the brain RDM and each model independently.

For each model, the output correlation values with the brain RDM were Fisher transformed and assigned to the center voxel of each searchlight sphere. We, therefore, obtained two separate correlation brain maps, one for each model. To estimate a group-level statistic we performed a voxel-wise t-test against baseline, for each group separately: the control group, the cataract-reversal group, the control-blurry1 group and the control-blurry2 group.

We also performed two-sample t-tests to check for differences between groups. For these groups' contrasts the TFCE correction was applied within the V1 mask for the correlations with the HmaxC1-model and within the VOTC mask for the correlations with the categorical model (for details about the masks, see the section above: *Statistical procedure and brain masks for TFCE correction*). This choice was guided by our a priori expectation—also supported by our within-group

whole-brain results (Fig. 4 panels A-B-C-D) – regarding where these two models would be most represented in the brain (i.e., HmaxC1 in V1 and the categorical model in VOTC).

Decoding analysis between and within categories

We complemented the RSA analyses with two different decoding analyses, one targeting the categorical representation in VOTC (5-way categories decoding) and one targeting the low-level visual representation in V1 (6-way items decoding for each category). These analyses were performed using the CoSMoMVPA¹⁰⁶ toolbox, implemented in MATLAB R2016b (Mathworks).

The 5-way categories decoding analysis is similar, on the theoretical level, to the RSA analysis testing the relation between brain representation and the categorical model. In both analyses, we expect that in VOTC, items from the same category have more similar patterns of activity compared to items from different categories. To foreshadow our results, since we observed no differences across control and cataract-reversals groups in the relation between brain representation in VOTC and the categorical model, we decided to carry out the between-category decoding analysis to corroborate our observation of no differences across these groups in how VOTC encodes categories of visual stimuli. However, given the conceptual redundancy of these analyses we placed the 5-way decoding between categories in the supplemental material (SI Fig. 5).

The 6-way item decoding within (Fig. 5) each category is an analysis that will, instead, add complementary information to the RSA analysis with the HmaxC1 model.

In our study, the selection of the categorical model was based on our own design (we created a stimulus set with 5 well defined categories). For the low-level visual model, we selected the HmaxC1 model that is well known to fit well the representation implemented in V1^{111,112}. However, there are many visual features that we could have tested separately (e.g., rectilinearity/curvilinearity; simple silhouette model; gist model; etc.), with the downside of inflating the multiple comparisons problem. But the 6-way item decoding does not require an a-priori selection of the low-level feature or model to be tested. In this case, in fact, the decoding algorithm will select by itself the features to use to classify the items. In addition, since this analysis is implemented within each category (we run 5 independent analyses within 5 different subsets of data, one for each category), we circumvent the confound of partial collinearity between categorical membership and low-level features (e.g., all faces share similar visual properties).

In all the classification analyses (between- and within- categories) we ran multiclass decoding analyses. In the decoding analysis between categories (5-way category decoding) we tested the discriminability of patterns for the five categories using a linear discriminant analysis (LDA). Since categories are known to be represented in VOTC, we expected that this task would be mostly performed in VOTC. In the decoding analyses within each category (6-way items decoding for each category) we tested the discriminability of brain voxel activation patterns for the six items within each category, using LDA. We run 5 separate analyses, one for each category. Since the low-level visual features of the items within each category were quite variable, we expected that V1 would mostly engage in such image specific decoding (e.g., the classification of the 6 images of bodies would be mostly based on the different low level visual properties of the 6 bodies)¹¹⁴. We, therefore, obtained five separate accuracy maps, one for each category.

In all decoding analyses, we performed a leave-one-run-out cross-validation procedure using beta-estimates from 4 runs in the training set, and the beta-estimates from the remaining independent run to test the classifier, with iterations across all possible training and test sets. This procedure was implemented in each 100-voxels sphere using a searchlight approach^{105,115}. Classification accuracy for each sphere was assigned to the central voxel of the sphere to produce accuracy maps.

The resulting accuracy maps were then smoothed with an 8-mm Gaussian kernel.

To estimate a group-level statistic we performed a voxel-wise t-test against chance level, for each group separately: the control group, the cataract-reversal group, the CON-blurry1 group and the CON-blurry2 group. We also performed two-sample t-tests to check for differences between groups. For the groups' contrasts in the between-categories decoding analysis, the TFCE correction was applied within the VOTC mask since we expect this kind of classification to be based on categorical membership. For the groups' contrasts in the within-category decoding analyses instead, the TFCE correction was applied within the V1 mask because, as mentioned above, we expected that V1 would mostly engage in such image specific decoding. For details about the masks, see the section above: *Statistical procedure and brain masks for TFCE correction*.

Deep neural network analyses

Earlier studies comparing representations between Deep Neural Networks (DNNs) and the human brain revealed a spatio-temporal hierarchical alignment: as DNN layer number increased, DNN representations exhibited stronger correlation with cortical representations emerging later in time, and in progressively downstream brain regions along both the dorsal and ventral visual pathways³⁴. We leveraged this alignment to gain further insights into the mechanisms underlying visual information processing after transient blindness early in life¹¹⁶.

More specifically, the DNN analyses aim to shed light on how categorical representations in the VOTC of cataract-reversal individuals are established despite early visual deprivation. One possibility is that, following sight restoration, these individuals learn to associate degraded or impoverished visual inputs with semantic or categorical information through different learning mechanisms. DNNs offer a useful framework for testing this second possibility: by simulating altered visual experience and observing whether high-level categorical representations can still emerge, we can assess whether exposure to degraded visual input alone might support the development of structured representations similar to those observed in VOTC.

To explore the alignment between deep neural network (DNN) outcomes and human functional magnetic resonance imaging (fMRI) data, we used AlexNet¹¹⁷.

For more details on the model architecture and data preparation please refer to Supplemental Material section titled: “DNN supplemental details and analyses”.

Training and testing conditions. Three distinct training and testing paradigms were designed to model the different visual processing conditions experienced by our human participants (see the section about participants and control experiment for a detailed description of these conditions in the fMRI experiment):

Intact-to-Intact ('Control' Network): In this baseline condition, the network was both trained and tested on standard, unaltered images. This setup served as the control to assess the network's performance under normal visual conditions.

Blurry-to-Blurry ('Cataract' Network): For this condition, both training and testing phases involved images that had been previously blurred to simulate cataract-like visual impairments. The blurring was applied as detailed in the data preparation subsection. This approach was designed to mimic the visual degradation typical in cataract-affected vision, allowing us to study the impact of such impairments on the network's performance and internal representations.

Note that the blurring only mimics vision after cataract removal, and not the near-blindness levels prior to the surgery. Based on the literature described in the introduction, we know that this pre-surgery deprivation will be particularly detrimental for the development of responses in early visual cortex due to the existence of an early

sensitive period. To mimic this aspect of the development in the cataract patients in at least a crude manner, the weights of the first convolutional layer were frozen (i.e., not trained) during the training phase of the Cataract-like network.

Intact-to-Blurry (‘Control-Blurry’ Network): Here, the network was trained on standard, unaltered images but tested on blurred ones. This paradigm simulates the experience of the control-blurry conditions, where individuals with normal vision were presented with blurred images. The application of the Butterworth filter for testing mirrors the procedure used in the Blurry-to-Blurry condition.

Features extraction. All the DNN activations in this study were extracted from the layers ‘features.1’, ‘features.4’, ‘features.7’, ‘features.9’, ‘classifier.2’ and ‘classifier.5’ during the testing phase on our 5-categories dataset. These layers were selected to span a progression from early to late stages of visual processing within the network. Specifically, features.1 is the ReLU activation following the first convolutional layer and is used to capture early, low-level visual representations; this layer is referred to as the V1-like layer. At the other end of the hierarchy, classifier.5—the ReLU activation following the penultimate fully connected layer—was chosen to reflect high-level categorical representations and is referred to as the VOTC-like layer. The intermediate layers (features.4, features.7, features.9, and classifier.2) provide a graded continuum of representational complexity, enabling a more detailed characterization of how visual information evolves across the network.

Analyses. Our goal was to apply the same analytical approaches used for the fMRI data, namely RSA and decoding, to the DNN data. In the main text, we report the results of the decoding analyses (between- vs. within-category), which we believe are better suited to address our hypotheses in this context. The RSA analysis, on the other hand, relied on the Hmax model, and since early layers of both HMAX and AlexNet are computational models of early visual cortex, they are naturally correlated. As a result, this comparison may offer limited additional insight and is therefore included only in the Supplemental Material (see also SI Fig. 9B). Additionally, we performed a direct correlation between the RDMs from brain ROIs and RDMs extracted from DNN layers (in the intact-to-intact DNN condition). This analysis is reported in the Supplemental Material (see also SI Fig. 9C).

Between- and within-category decoding analyses in different layers and conditions of the DNN. For each DNN condition, we replicated the decoding analyses previously conducted on the brain data. We followed the exact same procedure described for the fMRI data (see ‘Decoding Analysis’ in the Methods section), with the only difference being that we used activations extracted from various DNN layers instead of brain activity patterns.

To focus on the most relevant comparisons, we restricted our decoding analyses based on findings from the biological data and prior knowledge. Specifically, for the between-category decoding, we analyzed activations across the full range of DNN layers, from the V1-like to the VOTC-like layers, based on the assumption that both low-level and high-level stimulus properties can contribute to this type of categorization. In contrast, for the within-category decoding, we focused on the V1-like and intermediate layers, as this analysis revealed effects primarily in early visual regions in the brain.

We implemented non-parametric permutations to conduct statistical analyses. For each layer/condition, we determined the statistical difference from zero using a permutation test with 10,000 iterations. This involved constructing a null distribution for the classification values by computing them after randomly shuffling the conditions’ labels.

Additionally, we assessed the statistical difference between each group-like condition using a permutation test. We generated a null

distribution for the difference in classification values of the two conditions by computing the difference after randomly shuffling the conditions’ labels, repeating this step 10,000 times.

To determine the statistical significance of our results, we compared the observed result to the null distribution. This comparison involved calculating the proportion of observations in the null distribution that had a value higher than the one obtained in the real test. To address multiple comparisons, all p-values were corrected using false discovery rate (FDR)¹¹⁸.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Raw fMRI data are not publicly available as full anonymity of the participants cannot be guaranteed, even after defacing the MRI images and due to the lack of explicit consent from our participants. Pre-processed data, brain masks and statistical outputs can be found on OSF: <https://doi.org/10.17605/OSF.IO/BECDR>. Source data for any other graph are provided with this paper. Source data are provided with this paper.

Code availability

Codes used to process the data can be found on GitHub: <https://github.com/SteMat9787/Neurocat3>¹¹⁹.

References

- Hubel, D. H. & Wiesel, T. N. Receptive fields of cells in striate cortex of very young, visually inexperienced kittens. *J. Neurophysiol.* **26**, 994–1002 (1963).
- Knudsen, E. I. Sensitive periods in the development of the brain and behavior. *J. Cogn. Neurosci.* **16**, 1412–1425 (2004).
- Gordon, J. A. & Stryker, M. P. Experience-dependent plasticity of binocular responses in the primary visual cortex of the mouse. *J. Neurosci.* **16**, 3274–3286 (1996).
- Prusky, G. T. & Douglas, R. M. Developmental plasticity of mouse visual acuity. *Eur. J. Neurosci.* **17**, 167–173 (2003).
- Wiesel, T. N. & Hubel, D. H. Effects of visual deprivation on morphology and physiology of cells in the cat’s lateral geniculate body. *J. Neurophysiol.* **26**, 978–993 (1963).
- Wiesel, T. N. & Hubel, D. H. Single-cell responses in striate cortex of kittens deprived of vision in one eye. *J. Neurophysiol.* **26**, 1003–1017 (1963).
- Wiesel, T. N. & Hubel, D. H. Comparison of the effects of unilateral and bilateral eye closure on cortical unit responses in kittens. *J. Neurophysiol.* **28**, 1029–1040 (1965).
- Le Vay, S., Wiesel, T. N. & Hubel, D. H. The development of ocular dominance columns in normal and visually deprived monkeys. *J. Comp. Neurol.* **191**, 1–51 (1980).
- Mitchell, D. E. & Maurer, D. Critical periods in vision revisited. *Annu. Rev. Vis. Sci.* **8**, 291–321 (2022).
- Lewis, T. L., Maurer, D. & Brent, H. P. Development of grating acuity in children treated for unilateral or bilateral congenital cataract. *Invest. Ophthalmol. Vis. Sci.* **36**, 2080–2095 (1995).
- Lewis, T. L. & Maurer, D. Effects of early pattern deprivation on visual development. *Optom. Vis. Sci.* **86**, 640–646 (2009).
- Sinha, P. & Held, R. Sight restoration. *F1000 Med. Rep.* **4**, 17 (2012).
- Zohary, E. et al. Gaze following requires early visual experience. *Proc. Natl. Acad. Sci. USA* **119**, e2117184119 (2022).
- Heitmann, C. et al. Early visual experience refines the retinotopic organization within and across visual cortical regions. *Curr. Biol.* **33**, 4950–4959.e4 (2023).
- Sourav, S., Bottari, D., Kekunnaya, R. & Röder, B. Evidence of a retinotopic organization of early visual cortex but impaired

- extrastriate processing in sight recovery individuals. *J. Vis.* **18**, 22 (2018).
16. Arcaro, M. J. & Livingstone, M. S. On the relationship between maps and domains in inferotemporal cortex. *Nat. Rev. Neurosci.* **22**, 573–583 (2021).
17. Arcaro, M. & Livingstone, M. A Whole-brain topographic ontology. *Annu. Rev. Neurosci.* **47**, annurev-neuro-082823-073701 (2024).
18. Sourav, S., Bottari, D., Shareef, I., Kekunnaya, R. & Röder, B. An electrophysiological biomarker for the classification of cataract-reversal patients: a case-control study. *EClinicalMedicine* **27**, 100559 (2020).
19. De Heering, A. & Maurer, D. Face memory deficits in patients deprived of early visual input by bilateral congenital cataracts. *Dev. Psychobiol.* **56**, 96–108 (2014).
20. Le Grand, R., Mondloch, C. J., Maurer, D. & Brent, H. P. Early visual experience and face processing. *Nature* **410**, 890–890 (2001).
21. Robbins, R. A., Nishimura, M., Mondloch, C. J., Lewis, T. L. & Maurer, D. Deficits in sensitivity to spacing after early visual deprivation in humans: a comparison of human faces, monkey faces, and houses. *Dev. Psychobiol.* **52**, 775–781 (2010).
22. Gupta, P. et al. Vulnerability of facial attractiveness perception to early and multi-year visual deprivation. *Dev. Sci.* **26**, e13258 (2023).
23. Mondloch, C. J. et al. The effect of early visual deprivation on the development of face detection. *Dev. Sci.* **16**, 728–742 (2013).
24. Röder, B., Ley, P., Shenoy, B. H., Kekunnaya, R. & Bottari, D. Sensitive periods for the functional specialization of the neural system for human face processing. *Proc. Natl. Acad. Sci. USA* **110**, 16760–16765 (2013).
25. Grady, C. L., Mondloch, C. J., Lewis, T. L. & Maurer, D. Early visual deprivation from congenital cataracts disrupts activity and functional connectivity in the face network. *Neuropsychologia* **57**, 122–139 (2014).
26. Arcaro, M. J. & Livingstone, M. S. A hierarchical, retinotopic proto-organization of the primate visual system at birth. *eLife* **6**, e26196 (2017).
27. Gandhi, T. K., Singh, A. K., Swami, P., Ganesh, S. & Sinha, P. Emergence of categorical face perception after extended early-onset blindness. *Proc. Natl. Acad. Sci. USA* **114**, 6139–6143 (2017).
28. Mondloch, C. J., Robbins, R. & Maurer, D. Discrimination of facial features by adults, 10-year-olds, and cataract-reversal patients. *Perception* **39**, 184–194 (2010).
29. Gao, X., Maurer, D. & Nishimura, M. Altered representation of facial expressions after early visual deprivation. *Front. Psychol.* **4**, 878 (2013).
30. Gilad-Gutnick, S. et al. Motion’s privilege in recognizing facial expressions following treatment for blindness. *Curr. Biol.* **34**, 4047–4055.e3 (2024).
31. Kriegeskorte, N. Representational similarity analysis – connecting the branches of systems neuroscience. *Front. Syst. Neurosci.* <https://doi.org/10.3389/neuro.06.004.2008> (2008).
32. Haxby, J. V., Connolly, A. C. & Guntupalli, J. S. Decoding neural representational spaces using multivariate pattern analysis. *Annu. Rev. Neurosci.* **37**, 435–456 (2014).
33. Mattioni, S. Exploring brain representations through the lens of similarity structures. *Nat. Rev. Psychol.* **3**, 506–506 (2024).
34. Cichy, R. M., Khosla, A., Pantazis, D., Torralba, A. & Oliva, A. Comparison of deep neural networks to spatio-temporal cortical dynamics of human visual object recognition reveals hierarchical correspondence. *Sci. Rep.* **6**, 27755 (2016).
35. Frey, M., Nau, M. & Doeller, C. F. Magnetic resonance-based eye tracking using deep neural networks. *Nat. Neurosci.* **24**, 1772–1779 (2021).
36. Burkhalter, A. Development of forward and feedback connections between areas V1 and V2 of human visual cortex. *Cereb. Cortex* **3**, 476–487 (1993).
37. Coogan, T. A. & Van Essen, D. C. Development of connections within and between areas V1 and V2 of macaque monkeys. *J. Comp. Neurol.* **372**, 327–342 (1996).
38. Horton, J. C. & Hocking, D. R. Timing of the critical period for plasticity of ocular dominance columns in macaque striate cortex. *J. Neurosci.* **17**, 3684–3709 (1997).
39. Blasdel, G., Obermayer, K. & Kiorpes, L. Organization of ocular dominance and orientation columns in the striate cortex of neonatal macaque monkeys. *Vis. Neurosci.* **12**, 589–603 (1995).
40. Crair, M. C., Gillespie, D. C. & Stryker, M. P. The role of visual experience in the development of columns in cat visual cortex. *Science* **279**, 566–570 (1998).
41. Wiesel, T. N. & Hubel, D. H. Ordered arrangement of orientation columns in monkeys lacking visual experience. *J. Comp. Neurol.* **158**, 307–318 (1974).
42. Bock, A. S. et al. Resting-state retinotopic organization in the absence of retinal input and visual experience. *J. Neurosci.* **35**, 12366–12382 (2015).
43. Striem-Amit, E. et al. Functional connectivity of visual cortex in the blind follows retinotopic organization principles. *Brain* **138**, 1679–1695 (2015).
44. Wiesel, T. N. Postnatal development of the visual cortex and the influence of environment. *Nature* **299**, 583–591 (1982).
45. Maurer, D., Ellemberg, D. & Lewis, T. L. Repeated measurements of contrast sensitivity reveal limits to visual plasticity after early binocular deprivation in humans. *Neuropsychologia* **44**, 2104–2112 (2006).
46. Maurer, D. & Lewis, T. L. Visual acuity: the role of visual input in inducing postnatal change. *Clin. Neurosci. Res.* **1**, 239–247 (2001).
47. McKyton, A., Ben-Zion, I., Doron, R. & Zohary, E. The limits of shape recognition following late emergence from blindness. *Curr. Biol.* **25**, 2373–2378 (2015).
48. Huberman, A. D., Feller, M. B. & Chapman, B. Mechanisms underlying development of visual maps and receptive fields. *Annu. Rev. Neurosci.* **31**, 479–509 (2008).
49. Bourne, J. A. & Rosa, M. G. P. Hierarchical development of the primate visual cortex, as revealed by neurofilament immunoreactivity: early maturation of the middle temporal area (MT). *Cereb. Cortex* **16**, 405–414 (2006).
50. Kiorpes, L. & Movshon, J. A. Neural limitations on visual development in primates. in *The Visual Neurosciences*, 2-vol. set (eds Chalupa, L. M. & Werner, J. S.) 159–173 (The MIT Press, 2003). <https://doi.org/10.7551/mitpress/7131.003.0016>.
51. Spriet, C., Abassi, E., Hochmann, J.-R. & Papeo, L. Visual object categorization in infancy. *Proc. Natl. Acad. Sci. USA* **119**, e2105866119 (2022).
52. Zhang, J. et al. Mapping postnatal mouse brain development with diffusion tensor microimaging. *NeuroImage* **26**, 1042–1051 (2005).
53. Hyvärinen, J., Carlson, S. & Hyvärinen, L. Early visual deprivation alters modality of neuronal responses in area 19 of monkey cortex. *Neurosci. Lett.* **26**, 239–243 (1981).
54. Pitchaimuthu, K. et al. Steady state evoked potentials indicate changes in nonlinear neural mechanisms of vision in sight recovery individuals. *Cortex* **144**, 15–28 (2021).
55. Röder, B. & Kekunnaya, R. Effects of Early Visual Deprivation. in *Oxford Research Encyclopedia of Psychology* (Oxford University Press, 2022). <https://doi.org/10.1093/acrefore/9780190236557.013.839>.
56. Arcaro, M., Schade, P. & Livingstone, M. Preserved cortical organization in the absence of early visual input. *J. Vis.* **18**, 27 (2018).
57. Röder, B. & Kekunnaya, R. Visual experience dependent plasticity in humans. *Curr. Opin. Neurobiol.* **67**, 155–162 (2021).
58. Ostrovsky, Y., Andelman, A. & Sinha, P. Vision following extended congenital blindness. *Psychol. Sci.* **17**, 1009–1014 (2006).

59. Putzar, L., Hötting, K., Rösler, F. & Röder, B. The development of visual feature binding processes after visual deprivation in early infancy. *Vis. Res* **47**, 2616–2626 (2007).
60. Kobylkov, D. & Vallortigara, G. Face detection mechanisms: nature vs. nurture. *Front. Neurosci.* **18**, 1404174 (2024).
61. Tsao, D. Y. & Livingstone, M. S. Mechanisms of face perception. *Annu. Rev. Neurosci.* **31**, 411–437 (2008).
62. Maurer, D. Critical periods re-examined: evidence from children treated for dense cataracts. *Cogn. Dev.* **42**, 27–36 (2017).
63. Maurer, D. & Lewis, T. L. Sensitive Periods in Visual Development. in *The Oxford Handbook of Developmental Psychology* Vol.1 (ed Zelazo, P. D.) 201–234 (Oxford University Press, 2013). <https://doi.org/10.1093/oxfordhb/9780199958450.013.0008>.
64. Röder, B., Kekunnaya, R. & Guerreiro, M. J. S. Neural mechanisms of visual sensitive periods in humans. *Neurosci. Biobehav. Rev.* **120**, 86–99 (2021).
65. Pedersini, C. A. et al. White matter plasticity following cataract surgery in congenitally blind patients. *Proc. Natl. Acad. Sci. USA* **120**, e2207025120 (2023).
66. Wichmann, F. A. & Geirhos, R. Are deep neural networks adequate behavioral models of human visual perception? *Annu. Rev. Vis. Sci.* **9**, 501–524 (2023).
67. Cadena, S. A. et al. Deep convolutional models improve predictions of macaque V1 responses to natural images. *PLOS Comput. Biol.* **15**, e1006897 (2019).
68. Lindsay, G. W. Convolutional neural networks as a model of the visual system: past, present, and future. *J. Cogn. Neurosci.* **33**, 2017–2031 (2021).
69. Schrimpf, M. et al. Brain-Score: Which Artificial Neural Network for Object Recognition is most Brain-Like? Preprint at <https://doi.org/10.1101/407007> (2018).
70. Storrs, K. R., Kietzmann, T. C., Walther, A., Mehrer, J. & Kriegeskorte, N. Diverse deep neural networks all predict human inferior temporal cortex well, after training and fitting. *J. Cogn. Neurosci.* 1–21 https://doi.org/10.1162/jocn_a_01755 (2021).
71. Vogelsang, L., Vogelsang, M., Pipa, G., Diamond, S. & Sinha, P. Butterfly effects in perceptual development: A review of the ‘adaptive initial degradation’ hypothesis. *Dev. Rev.* **71**, 101117 (2024).
72. Vogelsang, L. et al. Potential downside of high initial visual acuity. *Proc. Natl. Acad. Sci. USA* **115**, 11333–11338 (2018).
73. Howell, A. L., Osher, D. E., Li, J. & Saygin, Z. M. The intrinsic neonatal hippocampal network: rsfMRI findings. *J. Neurophysiol.* **124**, 1458–1468 (2020).
74. Li, J., Osher, D. E., Hansen, H. A. & Saygin, Z. M. Innate connectivity patterns drive the development of the visual word form area. *Sci. Rep.* **10**, 18039 (2020).
75. Mattioni, S. et al. Categorical representation from sound and sight in the ventral occipito-temporal cortex of sighted and blind. *eLife* **9**, e50732 (2020).
76. Mattioni, S., Rezk, M., Battal, C., Vadlamudi, J. & Collignon, O. Impact of blindness onset on the representation of sound categories in occipital and temporal cortices. *eLife* **11**, e79370 (2022).
77. Osher, D. E. et al. Structural connectivity fingerprints predict cortical selectivity for multiple visual categories across cortex. *Cereb. Cortex* **26**, 1668–1683 (2016).
78. Saygin, Z. M. et al. Anatomical connectivity patterns predict face selectivity in the fusiform gyrus. *Nat. Neurosci.* **15**, 321–327 (2012).
79. Saygin, Z. M. et al. Connectivity precedes function in the development of the visual word form area. *Nat. Neurosci.* **19**, 1250–1255 (2016).
80. Hannagan, T., Amedi, A., Cohen, L., Dehaene-Lambertz, G. & Dehaene, S. Origins of the specialization for letters and numbers in ventral occipitotemporal cortex. *Trends Cogn. Sci.* **19**, 374–382 (2015).
81. Mahon, B. Z. & Caramazza, A. What drives the organization of object knowledge in the brain? *Trends Cogn. Sci.* **15**, 97–103 (2011).
82. Powell, L. J., Kosakowski, H. L. & Saxe, R. Social origins of cortical face areas. *Trends Cogn. Sci.* **22**, 752–763 (2018).
83. Conway, B. R. The organization and operation of inferior temporal cortex. *Annu. Rev. Vis. Sci.* **4**, 381–402 (2018).
84. Bavelier, D. & Neville, H. J. Cross-modal plasticity: where and how? *Nat. Rev. Neurosci.* **3**, 443–452 (2002).
85. Frasnelli, J., Collignon, O., Voss, P. & Lepore, F. Crossmodal plasticity in sensory loss. in *Progress in Brain Research* Vol. 191 233–249 (Elsevier, 2011).
86. Van Ackeren, M. J., Barbero, F. M., Mattioni, S., Bottini, R. & Collignon, O. Neuronal populations in the occipital cortex of the blind synchronize to the temporal dynamics of speech. *eLife* **7**, e31640 (2018).
87. Collignon, O. et al. Long-lasting crossmodal cortical reorganization triggered by brief postnatal visual deprivation. *Curr. Biol.* **25**, 2379–2383 (2015).
88. Dormal, G. et al. Tracking the evolution of crossmodal plasticity and visual functions before and after sight restoration. *J. Neurophysiol.* **113**, 1727–1742 (2015).
89. Guerreiro, M. J. S., Putzar, L. & Röder, B. Persisting cross-modal changes in sight-recovery individuals modulate visual perception. *Curr. Biol.* **26**, 3096–3100 (2016).
90. Sourav, S. et al. Sound suppresses earliest visual cortical processing after sight recovery in congenitally blind humans. *Commun. Biol.* **7**, 118 (2024).
91. Amedi, A., Malach, R., Hendler, T., Peled, S. & Zohary, E. Visuo-haptic object-related activation in the ventral visual pathway. *Nat. Neurosci.* **4**, 324–330 (2001).
92. Pietrini, P. et al. Beyond sensory images: object-based representation in the human ventral pathway. *Proc. Natl. Acad. Sci. USA* **101**, 5658–5663 (2004).
93. Van Den Hurk, J., Van Baelen, M. & Op De Beeck, H. P. Development of visual category selectivity in ventral visual cortex does not require visual experience. *Proc. Natl. Acad. Sci. USA* **114**, E4501–E4510 (2017).
94. Heimler, B., Weisz, N. & Collignon, O. Revisiting the adaptive and maladaptive effects of crossmodal plasticity. *Neuroscience* **283**, 44–63 (2014).
95. Makin, T. R. & Krakauer, J. W. Against cortical reorganisation. *eLife* **12**, e84716 (2023).
96. Bach, M. The freiburg visual acuity test - automatic measurement of visual acuity. *Optom. Vis. Sci.* <https://doi.org/10.1097/00006324-199601000-00008> (1996).
97. Bach, M. The freiburg visual acuity test-variability unchanged by post-hoc re-analysis. *Graefes Arch. Clin. Exp. Ophthalmol.* <https://doi.org/10.1007/s00417-006-0474-4> (2006).
98. Lieberman, H. R. & Pentland, A. P. Microcomputer-based estimation of psychophysical thresholds: the best PEST. *Behav. Res. Methods Instrum.* **14**, 21–25 (1982).
99. Krajbich, I. No camera needed with MR-based eye tracking. *Nat. Neurosci.* **24**, 1641–1642 (2021).
100. Smith, S. & Nichols, T. Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *NeuroImage* **44**, 83–98 (2009).
101. Eickhoff, S. B. et al. A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage* **25**, 1325–1335 (2005).
102. Amunts, K., Malikovic, A., Mohlberg, H., Schormann, T. & Zilles, K. Brodmann’s AREAS 17 AND 18 BROUGHT INTO STEREOTAXIC SPACE—WHERE AND HOW VARIABLE? *NeuroImage* **11**, 66–84 (2000).

103. Lorenz, S. et al. Two new cytoarchitectonic areas on the human mid-fusiform gyrus. *Cereb. Cortex* bhv225 <https://doi.org/10.1093/cercor/bhv225> (2015).
104. Rosenke, M. et al. A cross-validated cytoarchitectonic atlas of the human ventral visual stream. *NeuroImage* **170**, 257–270 (2018).
105. Kriegeskorte, N., Goebel, R. & Bandettini, P. Information-based functional brain mapping. *Proc. Natl. Acad. Sci. USA* **103**, 3863–3868 (2006).
106. Oosterhof, N. N., Connolly, A. C. & Haxby, J. V. CoSMoMVA: multi-modal multivariate pattern analysis of neuroimaging data in Matlab/GNU Octave. *Front. Neuroinform.* **10**, 27 (2016).
107. Kriegeskorte, N. & Kievit, R. A. Representational geometry: integrating cognition, computation, and the brain. *Trends Cogn. Sci.* **17**, 401–412 (2013).
108. Kriegeskorte, N. & Mur, M. Inverse MDS: inferring dissimilarity structure from multiple item arrangements. *Front. Psychol.* **3**, 245 (2012).
109. Ritchie, J. B., Bracci, S. & Op De Beeck, H. Avoiding illusory effects in representational similarity analysis: What (not) to do with the diagonal. *NeuroImage* **148**, 197–200 (2017).
110. Hubel, D. H. & Wiesel, T. N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.* **160**, 106–154 (1962).
111. Riesenhuber, M. & Poggio, T. Hierarchical models of object recognition in cortex. *Nat. Neurosci.* **2**, 1019–1025 (1999).
112. Serre, T., Wolf, L., Bileschi, S., Riesenhuber, M. & Poggio, T. Robust object recognition with cortex-like mechanisms. *IEEE Trans. Pattern Anal. Mach. Intell.* **29**, 411–426 (2007).
113. Khaligh-Razavi, S.-M. & Kriegeskorte, N. Deep supervised, but not unsupervised, models may explain IT cortical representation. *PLoS Comput. Biol.* **10**, e1003915 (2014).
114. Badwal, M. W., Bergmann, J., Roth, J., Doeller, C. F. & Hebart, M. N. The scope and limits of fine-grained image and category information in the ventral visual pathway. *J. Neurosci.* e0936242024 <https://doi.org/10.1523/JNEUROSCI.0936-24.2024> (2024).
115. Tong, F. & Pratte, M. S. Decoding patterns of human brain activity. *Annu. Rev. Psychol.* **63**, 483–509 (2012).
116. Celeghin, A. et al. Convolutional neural networks for vision neuroscience: significance, developments, and outstanding issues. *Front. Comput. Neurosci.* **17**, 1153572 (2023).
117. Krizhevsky, A., Sutskever, I. & Hinton, G. E. ImageNet classification with deep convolutional neural networks. *Commun. ACM* **60**, 84–90 (2017).
118. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B Stat. Methodol.* **57**, 289–300 (1995).
119. Stefania Mattioni. SteMat9787/Neurocat3: v1.0.0. *Zenodo* <https://doi.org/10.5281/ZENODO.17241924> (2025).
120. Xia, M., Wang, J. & He, Y. BrainNet viewer: a network visualization tool for human brain connectomics. *PLoS One* **8**, e68910 (2013).
- UCL) and the Consortium des Équipements de Calcul Intensif en Fédération Wallonie Bruxelles (CÉCI) funded by the Fond de la Recherche Scientifique de Belgique (F.R.S.-FNRS) under convention 2.5020.11 and by the Walloon Region. OC is a senior research associate at the Fond National de la Recherche Scientifique de Belgique (FRS-FNRS). The project was funded in parts by an ERC starting grant MADVIS (Project: 337573) awarded to OC; the Belgian Excellence of Science (EOS) program (Project No. 30991544) awarded to OC, HoB, VG; a Flagship ERA-NET grant SoundSight (FRS-FNRS PINT-MULTI R.8008.19) awarded to OC, the Fond Jacques Moulart attributed to OC, Fonds voor Wetenschappelijk Onderzoek (FWO) Flanders project GOD3322N and Methusalem project METH/24/003 (HO & AC).

Author contributions

Designed research: S.M., M.R., X.G., D.M., T.L., V.G., A.C., H.o.B., O.C.; Performed research: S.M., M.R., X.G., J.N., Z.L.; Analyzed data: S.M., A.C., R.G. Wrote the paper: S.M., O.C. Writing - Review & Editing: M.R., X.G., T.L., D.M., V.G., A.C., H.o.B. Funding acquisition: O.C.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-65468-7>.

Correspondence and requests for materials should be addressed to Stefania Mattioni or Olivier Collignon.

Peer review information *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025

Acknowledgements

We would like to express our gratitude to Sally Stafford and Joy Williams who have helped with recruiting participants and the data acquisition, respectively. Computational resources have been provided by the supercomputing facilities of the Université catholique de Louvain (CISM/