

Development of visual object recognition

Vladislav Ayzenberg^{1,2}✉ & Marlene Behrmann^{2,3}✉

Abstract

Object recognition is the process by which humans organize the visual world into meaningful perceptual units. In this Review, we examine the developmental origins and maturation of object recognition by synthesizing research from developmental psychology, cognitive neuroscience and computational modelling. We describe the extent to which infants demonstrate early traces of adult visual competencies within their first year. The rapid development of these competencies is supported by infant-specific biological and experiential constraints, including blurry vision and ‘self-curation’ of object viewpoints that best support learning. We also discuss how the neural mechanisms that support object-recognition abilities in infancy seem to differ from those in adulthood, with less engagement of the ventral visual pathway. We conclude that children’s specific developmental niche shapes early object-recognition abilities and their neural underpinnings.

Sections

Introduction

Perceiving object shape

Viewpoint invariance

Object categorization

Innate constraints on object recognition

Summary and future directions

¹Department of Psychology, University of Pennsylvania, Philadelphia, PA, USA. ²Department of Psychology, Carnegie Mellon University, Pittsburgh, PA, USA. ³Department of Ophthalmology, University of Pittsburgh, Pittsburgh, PA, USA. ✉e-mail: vayzenb@sas.upenn.edu; mbehrmann@pitt.edu

Introduction

Visual object recognition is a fundamental cognitive ability that organizes the visual world into meaningful perceptual units. Whether identifying food, threats or a friend in a crowd, object recognition enables observers to map visual information to conceptual content, which guides how they understand and interact with the world around them.

Although it is challenging to formulate a single unifying definition of an object, broadly speaking, objects are bounded visual entities that have a cohesive form that enables them to be segmented from other objects or the background (for reviews, see refs. 1,2). In other words, objects are entities whose properties persist together across time and space. Colloquially, objects include natural and man-made artefacts such as trees and books, as well as animate forms such as dogs, human bodies or faces. Objects can be contrasted with other visual entities such as textures and patterns (such as zebra stripes) or large-scale visual scenes and places (such as a view of a street or mountain), which can be composed of many objects.

The visual projection of a single object onto the retina can vary dramatically across changes in viewpoint, lighting or partial occlusion. Additionally, the appearance of objects within the same object category can vary in innumerable ways, such as in colour, texture, size and the appearance of individual object features. For instance, a dog can be seen from the side or the front, peering through tall grass, or covered in mud (Fig. 1a). Visual examples of the category ‘dog’ can have coats of different colours and textures, vary in size, and have different shaped snouts, tails and ears, and can be represented by simplistic line drawings or even shapes in the clouds (Fig. 1b).

Despite the substantial variability with which objects appear, human object recognition is extremely robust and accurate by adulthood. Within less than a tenth of a second, adults can identify objects across never-before-seen viewpoints and can generalize an object’s category to a novel exemplar^{3–5}. Adults can also learn to recognize novel objects extremely quickly^{6,7}, often after seeing only a single exemplar^{8–10}. One important process by which humans achieve such robust recognition is by constructing a holistic, three-dimensional (3D) shape-based representation that describes the overall, or ‘global’, form of the object, rather than using specific visual features that are visible from a particular object view or are specific to one category exemplar¹¹.

In adults, neural object representations are formed primarily via a hierarchy of processing stages along the ventral visual pathway, which projects from the occipital cortex to the anterior pole of the temporal lobe (Box 1). As information is propagated along this hierarchy, regions derive and compute increasingly more complex visual properties. Simple visual features, such as the specific contours of a dog’s nose, are computed in early visual cortex; intermediate shape-like features, such as the triangular form of a dog’s ear, are computed in mid-level areas such as V4; and large holistic features that map directly to conceptual categories, such as tools, animals and faces, are computed in high-level areas of the occipitotemporal cortex^{12,13}. Accumulating evidence also suggests that the recognition processes of this ventral pathway are complemented by input from regions of the dorsal visual pathway such as the posterior parietal cortex^{14,15}. Specifically, the dorsal pathway might compute properties crucial for forming holistic 3D shape representations, such as depth information^{16,17} and an object’s global structure^{18,19}, which describe the spatial arrangement of object features rather than their appearance^{18,19}. These processes are further supplemented by lateral and recurrent feedback connections within both pathways, which can further support recognition in ambiguous

contexts^{20,21}, by, for example, ‘filling in’ missing information in cases of partial occlusion or poor visibility.

Furthermore, researchers have begun to uncover the computational principles that support object recognition in adults via computational models known as deep neural networks (Box 2). When trained to recognize objects, deep neural networks can achieve human-level performance on a range of recognition tasks²², and their internal structure approximates the anatomical organization and functional profile of the adult ventral visual pathway^{23,24}.

Although substantial progress has been made in understanding the cognitive, neural and computational mechanisms that support mature object recognition, our understanding of how object recognition abilities develop is still limited. This discrepancy is well reflected in the limitations of deep neural networks, which – unlike humans – often require millions of training examples to be able to derive robust object representations^{25,26}. This limitation suggests that deep neural networks lack human-like developmental mechanisms for object recognition. Thus, understanding the origins and development of human object recognition enables a deeper understanding of the mechanisms that support object recognition in adulthood and opens up the possibility of training more human-like machines.

In this Review, we detail the developmental origins and processes that lead to mature object recognition in humans. To better understand the innate mechanisms and crucial life experiences that form the foundation of object-recognition abilities, we focus primarily on object recognition in pre-linguistic infants in the first year of life. We first review early shape-perception abilities in infants, and examine the infant-specific constraints that support the development of robust shape representations. Next, we examine how infants come to use depth and motion information to perceive objects in 3D, and the extent to which this process supports their ability to recognize objects across viewpoints. Finally, we outline how infants use shape information to learn object categories and we characterize the underlying processes that support categorization. Throughout, we illustrate how the rapid development of object recognition abilities is bootstrapped by specific biological and environmental factors that are unique to the developmental niche of infants.

Perceiving object shape

Object recognition involves matching perceived object properties to a pre-existing representation in memory. Perhaps the most diagnostic property of an object’s individual identity and category is its shape^{27,28}. Human adults and older children preferentially identify and categorize objects using shape information, rather than other cues such as colour or texture^{29–31}. Indeed, adults show little decrement in object-recognition performance when non-shape cues are removed, such as when looking at line drawings that lack colour and texture information^{32,33}.

The most common behavioural approach to studying shape perception in infancy is the habituation (or dishabituation) paradigm³⁴ (Fig. 2a,b). In the habituation phase, infants are shown one stimulus repeatedly until their looking time to the stimulus – a proxy for engagement – decreases to a predetermined criterion, which triggers the dishabituation phase. During the dishabituation phase, infants are shown either a familiar stimulus (similar to the habituation stimulus) or a novel stimulus (different from the habituation stimulus). If infants can discriminate between stimuli, their looking time increases (they dishabituate) for the novel stimulus relative to the habituation phase and the familiar stimulus. A conceptually similar approach is sometimes

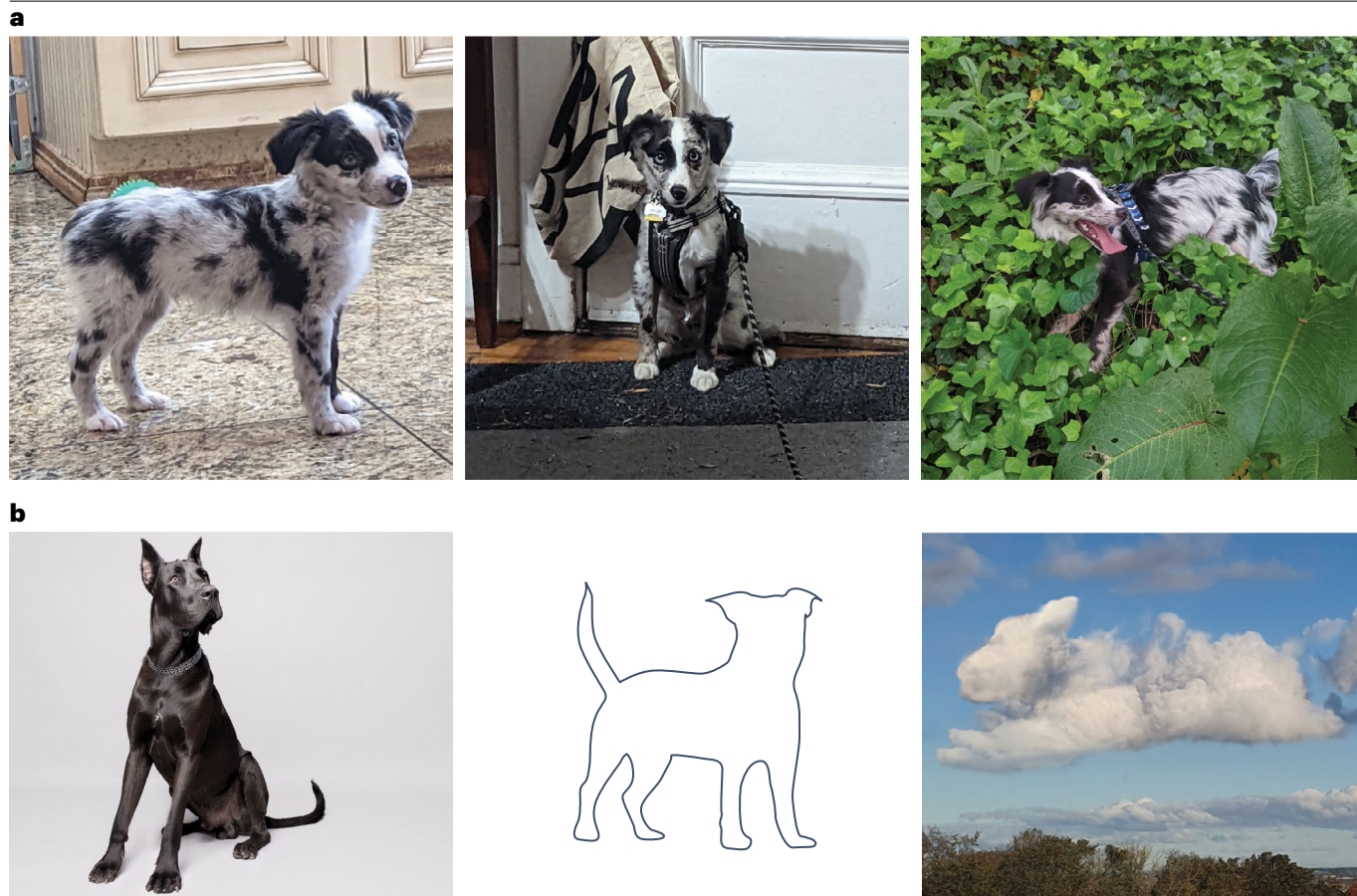


Fig. 1 | Different retinal projections of a dog. **a**, The same dog under different viewing conditions. **b**, Different examples of dogs. Despite variations in low-level image properties and different local visual features, all these images are readily recognized as showing a dog. Credit (Great Dane): ozgurdonmaz/getty.

referred to as the violation-of-expectation paradigm. Using habituation, researchers have shown that newborns can discriminate between simple two-dimensional (2D) shapes, as reflected by the finding that they look longer at a new shape than one identical to habituation³⁵. Newborns also display shape constancy, perceiving a shape as the same when it is seen from a novel 3D orientation, as evidenced by their looking longer at a new shape than at the habituation shape seen from a new orientation³⁶. By 3 months of age, infants can readily categorize familiar stimuli (such as dogs and cats) using only the shape silhouette of the stimulus. Specifically, they dishabituate to an object from a novel category, but not one from a familiar category³⁷ (Fig. 2b).

In this section, we discuss the mechanisms of shape perception in infants and young children. We consider the contributions of feedback connections and the role of visual acuity in the developmental trajectory of this ability.

Neural representations of shape

In adults, shape discrimination and shape constancy are most often ascribed to the lateral occipital cortex, a region in the posterior occipitotemporal cortex that is causally involved in object recognition^{38,39}. Using functional near-infrared spectroscopy (fNIRS) and an adaptation procedure with 6-month-old infants, one study examined whether infant lateral occipital cortex is also sensitive to shape information⁴⁰

(Fig. 2c) (for a review of the strengths and limitations of fNIRS, see refs. 41,42). Infants were shown blocks of trials in which either the shape of the object varied, but the colour and texture remained constant, or the colour and texture of the object varied, but the shape remained constant (Fig. 3a). Lateral occipital cortex in infants showed adaptation only when the shape information remained constant, even though colour and texture varied.

Other fNIRS research has shown that infants' neural representations are more sensitive to shape than colour^{43–45} (Fig. 3b). In these experiments, infants were shown an object entering one side of an occluding screen and then an object exiting from the other side of the screen. In some trials, the exiting object had a different shape (the shape condition) and, in others, it had a different colour (the colour condition) compared to the object that entered the screen. The lateral occipital cortex of infants showed a larger response to the shape change than the colour change condition. By approximately 10 years of age, the lateral occipital cortex of children, as measured using fMRI, shows adult-like functional selectivity to objects^{46–49}, and the responses are correlated with children's ability to recognize objects⁵⁰.

There is also evidence that the dorsal pathway exhibits shape sensitivity from early in development: 3-month-old infants tested with the occlusion paradigm described above exhibited stronger responses to changes in shape than to changes in colour within the

Box 1

The visual system at birth

The relative maturity of the visual system and cortical areas strongly constrains how object recognition can be accomplished at different ages. Evidence from humans and non-human primates suggests that the large-scale anatomical organization of the visual system is largely in place within a few days of birth^{77,206,241,242}. The areas of early visual cortex, occipitotemporal cortex and posterior parietal cortex are already organized along adult-like functional and anatomical boundaries, including functional differentiation into dorsal and ventral pathways and a hierarchical organization within each pathway^{77,242}. Moreover, these regions show a topographic organization with retinotopically based responses to visual input and eccentricity-biased connectivity between areas of early visual and high-level occipital-temporal cortices^{206,242,243}. However, in contrast to adult connectivity, these connections in newborns are predominantly feedforward, with relatively few feedback or lateral connections even at 4 months of age^{78,80}.

The early visual areas (particularly V1–V3) are among the earliest regions of the brain to develop and to display signatures of mature functioning at birth^{200–203}. For instance, single-unit recordings from non-human primates show that V1 in newborns already exhibits selectivity for oriented contours and properties important for 3D shape and depth perception such as binocular disparity^{123,244,245}. Higher-level areas of the ventral pathway, such as the occipitotemporal cortex, are slower to develop and show inconsistent selectivity to individual objects or categories, such as faces, several months after birth^{191,192,194,246,247}. The dorsal pathway is also visually responsive shortly after birth^{242,243}, and continues to develop functionally into late childhood and early adolescence^{48,248}.

There is accumulating evidence from both human and non-human primates that the cytoarchitecture of the dorsal pathway might mature earlier than the ventral pathway^{150,151}. For instance, the posterior parietal cortex shows a more mature laminar organization than the occipitotemporal cortex in newborn monkeys¹⁴⁹, and preliminary data from human infants suggests faster microstructural development of the parietal cortex compared to the inferior temporal cortex^{203,249}. The dorsal pathway is also found to be less capable of functional reorganization than the ventral pathway following disruptions in childhood^{205,250,251} — consistent with earlier maturity and a smaller window of plasticity. However, many classic studies exploring the development of the dorsal pathway based their conclusions partly on the maturity of motion perception^{150,151}, which

might be processed along a separate lateral pathway outside the dorsal pathway^{252,253}.

Another way of measuring the maturity of different regions is by examining the degree to which the time-varying stimulus response of a region is similar between children and adults. Using this approach, studies using naturalistic videos and whole-brain analysis have further confirmed the existence of a posterior-to-anterior pattern of maturation in the ventral and dorsal pathways. The time-varying response of infant early visual cortex to movies is correlated with that of adults²³⁷ and its correlation with adult early visual cortex remains stable across age^{238,239}. By contrast, the correlation between child and adult occipitotemporal cortex and posterior parietal cortex becomes progressively stronger with age, suggesting protracted development of higher-level visual areas relative to lower-level ones^{238,239}. This gradient can also be seen within each pathway, such that object-selective responses are most adult-like in posterior portions of occipitotemporal cortex and posterior parietal cortex and least adult-like in the anterior portions⁴⁸.

Researchers have also suggested that visual behaviours in newborns are supported by subcortical structures^{214,254,255}. The anatomical organization of the subcortex and its projections to the cortex are already in place at birth^{110,256,257}. The superior colliculus and pulvinar are thought to play a key part in object recognition in children and adults^{258–260}. Like the early visual cortex, the adult superior colliculus and the pulvinar are visually responsive and have a topographic organization corresponding to the retina^{257,261}, as well as bilateral connections with the ventral and dorsal visual pathways^{262–265}. Although neuroimaging data linking infant recognition abilities to the subcortex are sparse, some studies find that lesions to subcortical regions in newborns have a greater impact on visual abilities than lesions to occipital cortex^{255,266}. Researchers have even suggested that visual recognition is almost exclusively controlled by subcortical regions in newborns and that cortical regions eventually take on these recognition processes as infants mature^{214,267}.

Existing research suggests that the overall anatomical organization of the brain is largely adult-like at birth but that subcortical and early visual areas mature earlier than higher-level ventral and dorsal regions. Future fine-grained microstructural analysis of the dorsal and ventral regions can confirm whether the dorsal pathway matures earlier than the ventral pathway.

parietal cortex⁵¹. Furthermore, the response of posterior parietal cortex in older children (over 6 years of age) varied parametrically with the degree to which a shape is distorted and was correlated with the child's object-recognition abilities^{48,50}.

Local and global shape properties

The use of local and global properties to represent shape changes over the course of development. Studies with older infants (over 4.5 months of age) have shown that they are sensitive to local properties such as the relative lengths and angles that comprise 2D shapes^{52–55},

as well as to pictorial cues that convey 3D shape (such as line junctions)^{56,57}. However, object recognition in adults is primarily supported by global shape properties, rather than by local shape properties like angles and lines^{58,59}. An extensive literature has revealed that global shape percepts in adults are supported by gestalt principles of perceptual grouping (for a review, see ref. 60).

Many studies have shown that infants begin to construct shape percepts using the gestalt principles of similarity, proximity and goodness of form by at least 4 months of age^{2,61–63}. For instance, in a series of classic studies, researchers examined whether infants implemented the gestalt

principle of good continuation to perceive the two halves of an occluded object as one unified object, rather than as two separate elements (Fig. 2a). Infants were habituated to a display of a diagonally oriented rod whose centre was occluded. The upper and lower halves of the rod were either aligned, supporting good continuation of contours, or not aligned. At test, infants were either shown a display with a complete rod or a display with two segmented rod halves. Two-month-old infants (but not newborns) who were habituated to the aligned rod dishabituated to the segmented rod at test, suggesting that they perceptually grouped the pieces of the occluded rod into a single shape^{64–67}. Importantly, infants who were habituated to the misaligned rod dishabituated to the complete rod, further suggesting that good continuation is needed for the perception of unity and that these results are not explained by a pre-existing preference for the segmented rod. Other research has found that newborns, like adults, show a global-precedence effect to preferentially categorize objects on the basis of global form, such as by the arrangement of local elements, rather than on the basis of the shape of the local elements themselves^{63,68–70}.

In adults, an extensive neuroimaging literature has suggested that perceptual grouping processes are generally accomplished via recurrent feedback connections between higher-level neural regions like lateral occipital cortex and early visual cortex^{71–73}. Computational

work has further shown that incorporating recurrent connections in deep neural networks is crucial for models to recognize objects when images are partially occluded^{20,74}, and the internal representations of the models become better aligned to those of the adult ventral pathway^{75,76}. However, feedback connections might be relatively sparse in the first few months of human life^{77–80}. For instance, post-mortem tracing studies of the brains of human infants showed a normal distribution of feedforward connections, but extremely few feedback connections even at 4 months of age^{78,80}. Moreover, anatomical and functional connectivity studies with pre-term and term (26–44 weeks of gestational age) infants as well as older children (7–17 years of age) have shown that they have far fewer long-range connections compared to adults, and that their connectivity profile is, instead, dominated by local connections^{81,82}.

Nevertheless, by 7 months of age, infants show some evidence of feedback processing, with fNIRS research showing enhancement of early visual cortex signals in preparation of an anticipated stimulus⁸³, and electroencephalography (EEG) research shows a late ‘second wave’ of activity in occipitotemporal cortex when processing a masked stimulus⁸⁴. These feedback processes are nevertheless immature, such that even 15-month-old infants show a much slower second wave of putative feedback processing compared to adults⁸⁴. Consistent with the

Box 2

Deep neural networks

Deep neural networks (DNNs) are hierarchically organized computational models consisting of multiple layers. Each DNN layer consists of many units, or ‘artificial neurons’, that have weighted connections with units in different layers. Most conventional object classification DNNs are trained in a supervised fashion wherein they iteratively attempt to predict the category labels for a set of object images. The accuracy, or ‘error’, of these predictions on each iteration is used to update the weights between units to improve subsequent predictions. This process is repeated many times until accuracy plateaus. Researchers also increasingly use unsupervised, or ‘self-supervised’, training methods, such as methods that generate contrasting categories directly from the training data or ones that update their weights by predicting the next frame of a video¹⁸².

These approaches have led to the high accuracy of DNNs on object-recognition tasks, as well as a strong correspondence to the hierarchical organization and functional profile of the ventral pathway. One study found that DNNs that showed the best object-recognition performance also exhibited the best statistical fit to the multivariate neural responses of mid-level and high-level regions in the adult ventral pathway¹⁹⁹, suggesting that optimizing an artificial system for object recognition leads to representations similar to those in the ventral pathway. Indeed, examination of the internal representations of DNN units reveals qualitatively similar types of feature selectivity, with simple visual features, such as the oriented grating represented in early model layers (akin to early visual cortex), more complex shape-like features in mid-level layers (akin to V4), and holistic object features such as faces in later layers (akin to occipitotemporal cortex)²⁶⁸. Importantly, researchers often find that improving the biological plausibility of a DNN’s architecture, such as

by adding feedback connections²⁰ or local-connectivity constraints²³, improves object-recognition performance and the model’s fit to the ventral pathway. However, the relation between a DNN’s biological plausibility and performance has plateaued in the past few years²⁶⁹, suggesting that newer models are achieving better performance using strategies different from humans.

DNNs can be an excellent artificial ‘model organism’ for the study of development. Like animal models, DNNs can be manipulated, or ‘control-reared’, to have different biological mechanisms or experiences and then used to address questions that are rarely accessible using current child-testing methods. Because DNNs can be constructed to execute the same tasks as infants and young children, researchers can iteratively test what kinds of mechanism are needed for DNNs to exhibit the same behavioural signatures as children^{234,236}.

However, DNNs currently have many limitations. One key limitation is that the training needed for DNNs to accomplish object recognition is often greater than, and qualitatively different from, what humans need^{25,26}. Moreover, despite many training approaches and architecture manipulations, DNNs continue to exhibit difficulties in representing holistic global shape information and instead primarily represent objects via local visual features and textures^{30,270}. One possible explanation for these limitations is that DNNs have not yet been able to approximate the perceptual processing of the dorsal pathway¹⁸ or a human-like learning objective²²⁸. In the future, DNNs might overcome these limitations in part by incorporating developmentally plausible biological and experiential constraints or learning objectives.

Review article

slow development of feedback connections, studies have shown that although older children (3–5 years of age) exhibit perceptual grouping abilities during object recognition⁸⁵, these abilities might not fully mature until adolescence^{86–89}. Although feedback connections exhibit a prolonged developmental trajectory, they might be sufficiently intact to support rudimentary perceptual grouping abilities in infancy.

Visual acuity and holistic recognition

In addition to feedback connections, perceptual grouping in infancy might be supported by infants' lower visual acuity. At birth, infant visual acuity is approximately 20 to 30 times poorer than for adults, with a Snellen equivalent of about 20/600^{90–93}. Acuity increases rapidly over the first few months of life, reaching adult levels for high-contrast

displays by approximately 6 months of age⁹⁴. Rather than being a limitation, the low level of acuity in newborns might be adaptive for their age-specific ecological niche^{95,96}. Even with 20/600 vision, newborns can clearly see a 12-cm stimulus from 60 cm away. This acuity is well within the range for an infant viewing a parent's face at arm's length or the objects of a mobile dangling above a crib. Initial low visual acuity might also be crucial for the development of mature object-recognition abilities. Specifically, impoverished input might reduce competition from distracting background stimuli, enabling infants to learn objects better when they are within view⁹⁷. Indeed, in animals, such as quail, rats and monkeys, exposure to visual stimuli at developmentally inappropriate time points (such as by surgically opening a rat pup's eyes or by shining light patterns through a quail chick's egg) impairs sensory

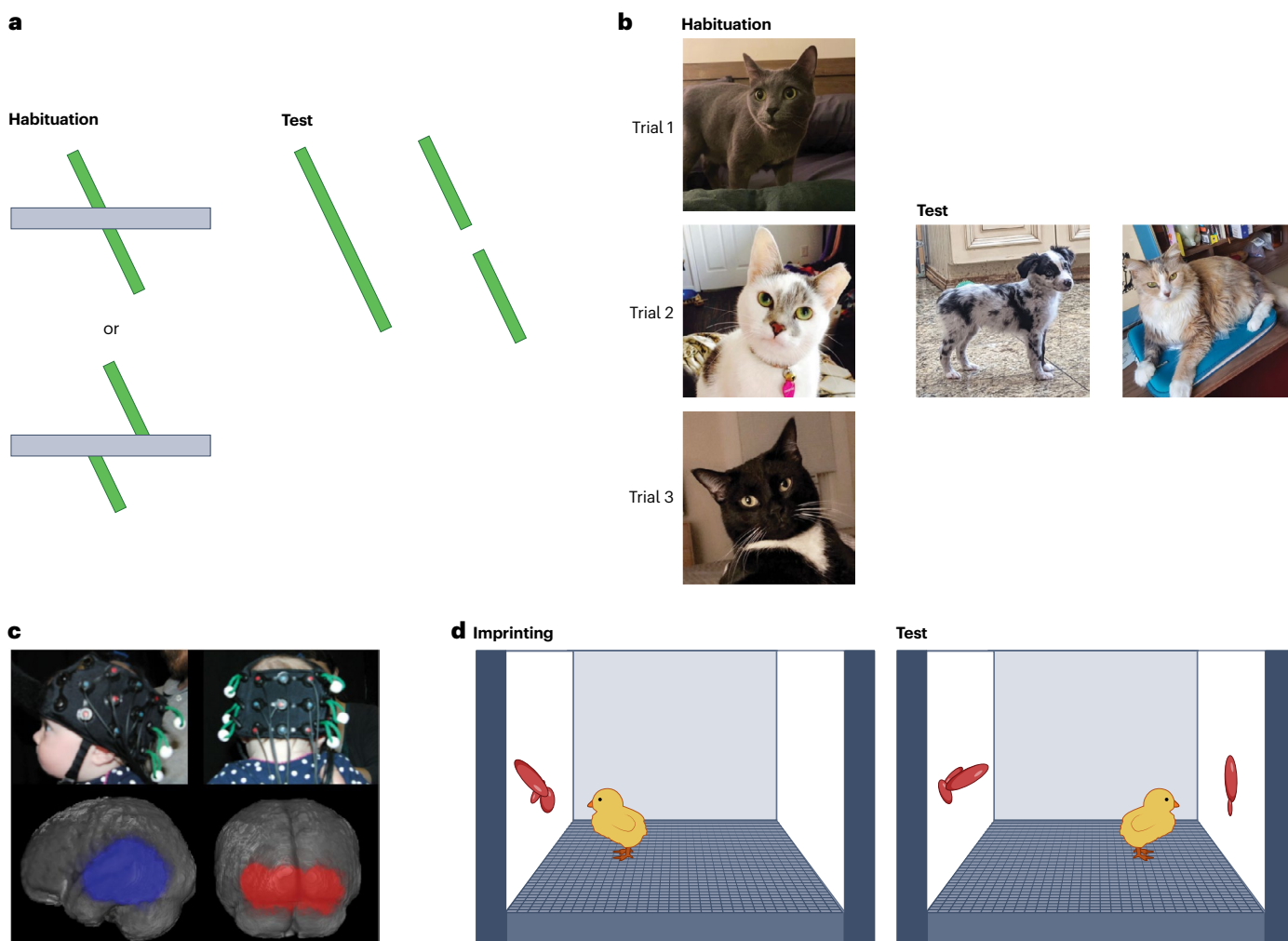


Fig. 2 | Procedures for studying the development of visual perception and object recognition. **a**, In a habituation paradigm, infants are habituated to the repeated presentation of a stimulus and then tested with a familiar stimulus and a novel stimulus. For example, if infants are habituated to an occluded rod where good continuation of contours is preserved, they look longer at the segmented than the complete rod in the test⁸⁴. **b**, A variation of the habituation paradigm, wherein infants are habituated to multiple category exemplars (different cats) and then tested with either a novel exemplar from the same category (a new cat) or an exemplar from a different category (a dog). Paradigm as in ref. 159.

c, An infant wearing an functional near-infrared spectroscopy (fNIRS) cap with electrodes arranged to measure neural signals from temporal cortex (left; blue) or early visual cortex (right; red). Panel **c** is reprinted with permission from ref. 83, PNAS. **d**, In a control-rearing study examining viewpoint-invariance, chicks are first imprinted on one rotating object, and then tested with either the imprinted object from a new orientation or a novel object. Chicks continue to follow the imprinted object, even when it is presented from a novel orientation. Panel **d** is courtesy of Building a Mind Lab.

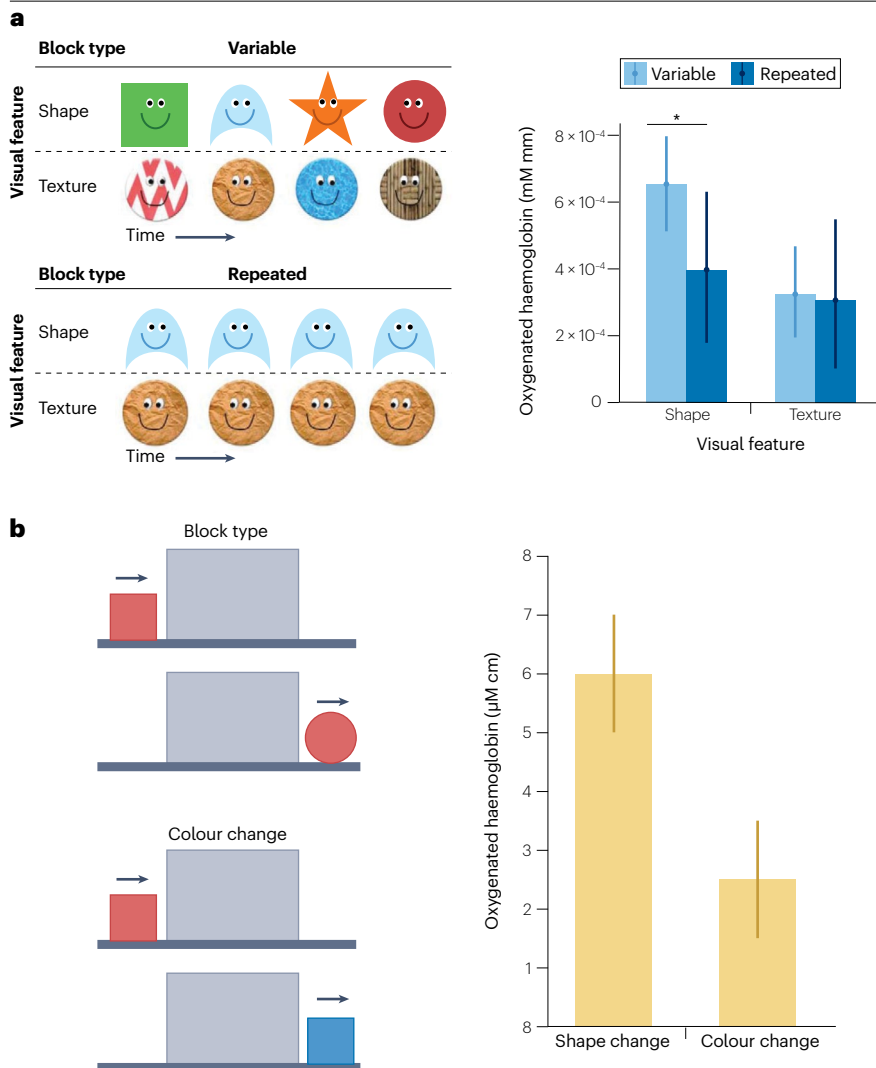


Fig. 3 | Shape sensitivity in the infant brain. Example procedures for measuring shape sensitivity in infants. **a**, In an adaptation procedure, infants are shown blocks of trials in which one stimulus property (shape) either varies or stays the same (left). Infant lateral occipital cortex shows reduced neural activity when object shape stays constant, even when colour and texture information varies (right). Asterisk indicates significant difference between conditions. Panel **a** is adapted from ref. 40, CC BY 4.0 (<https://creativecommons.org/licenses/by/4.0/>). **b**, In an occlusion procedure, infants were shown displays in which an object's shape changed after it moved behind an occluding screen (top), or its colour changed (bottom). Infant lateral occipital cortex and posterior parietal cortex showed greater neural response to the shape-change condition than to the colour-change condition. Data from ref. 240.

development^{98–100} and can impede the animal's ability to discriminate between objects¹⁰¹.

Importantly, low visual acuity might help infants to develop holistic object representations, rather than those based on the appearance of individual object features. For instance, children who have had their sight restored later in childhood (by surgical removal of dense cataracts after 1 year of age) transition from near-total visual deprivation to high visual acuity in a relatively short period of time, and consequently do not experience a period of blurry vision. Although these children gain normal visual function in many domains, they are specifically impaired at integrating visual features in a holistic manner, a process crucial for face recognition and perception of global form^{102–104}.

Consistent with these findings, newborns preferentially categorize objects using global, rather than local, shape cues by relying on low-spatial-frequency, or 'blurry', information⁶⁸. In these experiments, infants were habituated to stimuli with both global and local information, such as a cross shape comprising several smaller cross shapes, and then tested with objects where either the global information changed (a diamond comprising smaller cross shapes) or the local information

changes (a cross comprising smaller diamond shapes). Infants consistently dishabituated to a change in the global, but not the local cue, and – importantly – continued to do so when the displays were filtered to have only low-spatial-frequency information (blurred) but not when they were filtered to have only high-spatial-frequency information (sharpened). Moreover, simulations with deep neural networks have revealed that a training regimen consisting of blurry images at the start of training and high-resolution images at the end of training results in better recognition performance for holistically defined categories, such as faces, than training that consists of high-resolution images alone^{105–107} (Box 2).

The importance of low-spatial-frequency input in the development of object recognition is consistent with the neural development of magnocellular and parvocellular cells in the lateral geniculate nucleus and the broader dorsal and ventral pathways. Magnocellular cells are most sensitive to contrast and low-spatial frequency information, enabling them to support perception of global form¹⁰⁸. These neurons primarily transmit information to the dorsal rather than the ventral pathway, via direct projections from the lateral geniculate nucleus and indirect

projections via early visual cortex¹⁰⁹. By contrast, parvocellular cells are most sensitive to colour and high-spatial frequencies, enabling them to support perception of fine-grained details and local object features. Parvocellular cells project primarily from the lateral geniculate nucleus to the ventral pathway by way of early visual cortex. Mirroring the visual sensitivities of infants to global form and low-spatial-frequency information, magnocellular cells develop anatomically and functionally earlier than parvocellular cells in both humans and monkeys^{80,110–112}. The earlier development of the magnocellular cells relative to parvocellular cells might also then drive the dorsal pathway to mature earlier than the ventral pathway (Box 1).

Infants' early perceptual grouping abilities might be primarily supported by the dorsal pathway and its magnocellular input. Research with human adults has shown that posterior parietal cortex in the dorsal pathway might be involved in representing the spatial arrangement rather than the appearance of the local elements, with the representation of arrangement leading to the formation of global shape percepts^{18,19,113,114}. Also in adults, disruptions to posterior parietal cortex (from damage or transient magnetic stimulation) impair perception of global, but not local, shape properties^{115–117}. By contrast, the ventral pathway, with its strong parvocellular inputs, is particularly sensitive to fine-grained local visual features^{118,119}. Thus, an emerging perspective is that a complete object representation, one that includes both global and local information, arises via interactions of the dorsal and ventral pathways, though additional evidence is needed to support this conjecture^{14,120}.

All told, the low visual acuity and early development of magnocellular cells in infants might provide an important developmental constraint for bootstrapping their shape-perception abilities. Thus, although one might think that lower-resolution information imposes a limitation on the infant visual system, as noted above, studies suggest that blurry vision early in neonates might be adaptive. Low acuity might help infants to filter irrelevant background information and support the formation of holistic shape percepts.

Together, these studies show that infants are particularly sensitive to shape information from the earliest days of life and, like adults, show evidence of shape constancy and a global precedence effect. Existing evidence suggests that these early shape-processing abilities are bootstrapped by infant-specific biological constraints such as low visual acuity and neural pathways driven by magnocellular input.

Viewpoint invariance

An important signature of object recognition is the ability to recognize an object as the same object across different viewing conditions, such as changes in orientation, size and illumination. This process is usually referred to as viewpoint-invariant object recognition. Despite the term 'invariance,' even adult humans rarely achieve perfect invariance¹²¹, and often fail in contexts where the object is overly self-occluded such as when foreshortened¹²². Nevertheless, humans demonstrate remarkable tolerance to changes in orientation, including viewing objects at orientations they have never seen before. Here, we focus on the mechanisms that support tolerance for viewpoints that change the appearance of an object's shape on the retina, namely rotations in depth.

A prerequisite for identifying a depth-rotated object is the ability to perceive depth from cues such as binocular disparity and to form 3D shape representations. Electrophysiology studies with monkeys have shown that ocular dominance columns and sensitivity to binocular disparity in early visual cortex are already present at birth^{123,124} (Box 1). Likewise, newborns are sensitive to the depth properties of objects¹²⁵

and show a visual preference for real objects over photographs of the same objects (which lack stereoscopic depth cues)¹²⁶. However, it is not until approximately 4 months of age that infants consistently demonstrate the ability to represent 3D shape from binocular disparity^{127,128}, motion^{129–131}, or line junctions and shading^{56,57}. At 4 months of age, infants' perception of 3D shape becomes remarkably precise, enabling them to discriminate between visually matched possible and impossible figures¹³². At 6 months of age, they can generalize an object's 3D shape across different types of pictorial depth cue¹³³.

Although infants do not show mature depth perception abilities until 4–6 months of age, they already demonstrate the ability to accomplish viewpoint-invariant recognition. Indeed, by at least 3 months of age, infants can identify 3D objects in a viewpoint-invariant manner that enables them to recognize objects from orientations not previously encountered. Using a variation of the habituation paradigm, researchers presented infants with an image sequence consisting of a single object presented at different orientations on each trial. At test, infants viewed the same object at an orientation that had not been displayed previously. Infants did not dishabituate to the familiar object in a novel orientation, but looked longer at a visually matched novel object, indicating their recognition of the familiar object^{134,135}. Another study used an operant learning procedure to train 3-month-old infants to kick their legs in the presence of an object presented from one viewpoint. Both immediately after training and also after a 24-h delay, infants kicked when shown the trained object presented from a novel viewpoint but not when shown a novel object, suggesting that they had recognized the object from its novel viewpoint¹³⁶.

The neural basis for infants' viewpoint-invariant object representations is unclear. Data from older children suggests that the lateral occipital cortex displays invariance to changes in size at an earlier point in development (5–10 years of age) than changes in orientation (11–17 years of age), suggesting a late emergence of 3D shape in the ventral pathway⁴⁹. Other studies with 5–10-year-old children have found that the response profiles of lateral occipital cortex and posterior parietal cortex were correlated with the ability to recognize objects across different viewpoints⁵⁰. Similarly, representations of 3D shape and depth information have been documented in lateral occipital cortex and posterior parietal cortex in adult human and non-human primates^{137–140}. The specific contributions of each pathway to 3D shape perception might depend in part on the type of depth information available. For example, in adults, 3D shape from shading might be computed primarily in the ventral pathway, whereas 3D shape from motion or disparity is computed in the dorsal pathway^{17,137,138,141}. This distinction between dorsal and ventral representations of depth might explain how infants and young children accomplish viewpoint-invariance behaviourally, despite immature representations of 3D shape in their ventral pathway.

Early studies suggested that 3D shape perception and viewpoint-invariant object recognition in adults was accomplished by interpolating between previously seen views of an object^{121,142}. However, by 3 months of age, infants have been shown to generalize object identity to viewpoints that could not be interpolated from previous experience^{134,136}.

Studies with newborn chickens (chicks) have suggested that very little object experience is needed to accomplish viewpoint-invariant object recognition. For example, in one study, newborn chicks were reared in a controlled virtual-reality chamber containing only a single object that rotated through a limited set of viewpoints¹⁴³. Because chicks quickly imprint on and follow around the first object they see, researchers can test their recognition performance by examining

whether they follow novel versions of the imprinted object. The chicks were presented with either the imprinted object from a novel orientation or a visually matched novel object (Fig. 2d). The chicks preferred to follow the imprinted object shown from a novel viewpoint even when the novel object had greater image-level similarity to the originally imprinted object viewpoint. Subsequent research has demonstrated that experience with object rotations is unnecessary, so that chicks form invariant object representations when imprinted to an object presented from a single viewpoint¹⁴⁴. These controlled-rearing studies provide a powerful tool with which to explore how much visual experience is needed for an organism to form 3D shape representations, as well as insight on which evolutionarily preserved structures might be sufficient to accomplish viewpoint-invariant recognition.

Although extensive viewpoint experience might not be needed for viewpoint-invariant recognition, motion is an important visual cue for infants when forming 3D shape representations. For instance, 3-month-old infants are better able to generalize to novel viewpoints of an object when they are familiarized with smoothly rotating objects (60° of rotation) rather than static images of the object from different viewpoints^{130,131,145}. Moreover, control-rearing studies with chicks show that chicks are only able to accomplish viewpoint-invariant object recognition if exposed to objects that exhibit smooth continuous motion rather than temporally discontinuous static object images¹⁴⁶.

Beyond viewpoint invariance, motion might be a foundational perceptual grouping cue crucial for shape perception. For instance, 2-month-old infants are only able to perceive a complete object through partial occlusion when the parts of the object move synchronously but not when they are static^{64–67}. It is not until 5 months of age that infants can perceptually group object parts that were occluded into a single object from static images alone². Likewise, patients recovering from prolonged visual deprivation overwhelmingly rely on motion cues to parse visual scenes^{102,147,148}. Consistent with these behavioural data, motion-selective areas (such as area MT) are some of the earliest regions of the visual system to mature, providing an early neural mechanism by which motion might support the representation of shape^{149–152}.

Altogether, these data suggest that viewpoint-invariant object recognition is robust from early in development, and that infants might require little experience with different object views to identify objects across variations in viewpoint. Instead, domain-general experience with smoothly moving objects helps infants to develop the ability to construct robust 3D shape representations.

Object categorization

Another critical visual recognition ability is object categorization. Like viewpoint-invariant recognition, object categorization requires a shape representation that is not dependent on visual features specific to any single projection of the object (Fig. 1b). However, whereas viewpoint-invariant recognition is the process of identifying an object as remaining the same across different viewing conditions (constancy), object categorization is the process of grouping perceptually dissimilar objects on the basis of a shared property. We consider the evidence from behavioural tasks and patterns of neural selectivity to understand the development of object categorization.

Behavioural categorization

Categorization is central to the mental representation of objects. Grouping objects into distinct categories provides a low-dimensional structure by which humans can understand the similarity between objects and quickly make inferences about the functional properties of

novel objects. For instance, grouping poodles and Australian shepherds as objects in the category dog can help one to infer that poodles can bark because Australian shepherds also do. Adults can readily form object categories from visual properties or abstract rules. However, most object categories are represented at the basic level of categorization, in which objects have similar global shapes but varying local features^{153,154}. Subordinate-level categorization or recognition of individual items in the category is also important, enabling one to pick out a specific object such as one's own dog.

Infants demonstrate the ability to categorize objects from early in development¹⁵⁵. For instance, with exposure to just three category exemplars whose local features vary, newborns can categorize a novel category exemplar (with never-before-seen local features) on the basis of shape properties such as closure (categorized as 'O'-shaped versus 'X'-shaped figures)¹⁵⁶. By 3 months of age, infants can categorize more complex shapes (such as squares, circles and triangles) by their global form despite variations in their contours and local shape properties^{157,158}.

For naturalistic images such as animals and furniture, infants show a coarse superordinate-level to basic-level trajectory. At 2 months of age, infants form a category for mammals that includes dogs and cats, but excludes birds and furniture. By 3 months of age, infants show evidence of basic-level category boundaries^{159,160}. Moreover, infants' representations of naturalistic categories is relatively abstract. For example, if they are habituated to naturalistic photographs of dogs, they will generalize to new dogs presented as silhouettes that lack colour and texture information³⁷.

Infants are also sensitive to the distributional properties of category exemplars. For instance, the inclusivity of infants' categories varies with a category's variability, such that infants demonstrate more exclusive category boundaries for objects that have less visual variability among exemplars (such as cats) than those with greater visual variability (such as dogs)^{159,161,162}. This sensitivity can result in uneven, or asymmetric, category boundaries wherein infants are more likely to classify a novel cat as a member of the category 'dog', than a novel dog as member of the category 'cat'. As in adults, infants' category boundaries can be altered experimentally by manipulating the variability of habituated exemplars. Infants are better at categorizing novel objects whose features correspond to the central tendency of a category (the prototype) than exemplars near the category boundary: they would be more likely to categorize a golden retriever as a dog than a pug^{158,163}. Also like adults, infants' recognition of an object prototype improves with exposure to a greater number of category exemplars and with longer delays between training and test¹⁵⁸. The formation of object prototypes might serve as a crucial mechanism for rapid object learning more generally: deep neural network models that emphasize object prototypes demonstrate better performance on few-shot categorization tasks and also reveal infant-like signatures of categorization, such as category asymmetries¹⁶⁴.

There is evidence that by 6 months of age, infants, like adults^{10,165}, might already be able to accomplish one-shot categorization⁶⁹. Specifically, with exposure to just one example, 6-month-old infants will extend category membership to an exemplar with novel local features⁶⁹ (Fig. 4). Infants' performance on this task was best matched by a computational model that represents global shape (known as the shape skeleton^{59,166}) rather than the visual features learned by deep neural networks, further suggesting that category learning relies on a global representation of shape in infancy. By 4 years of age, children are adept at using shape information to accomplish one-shot categorization

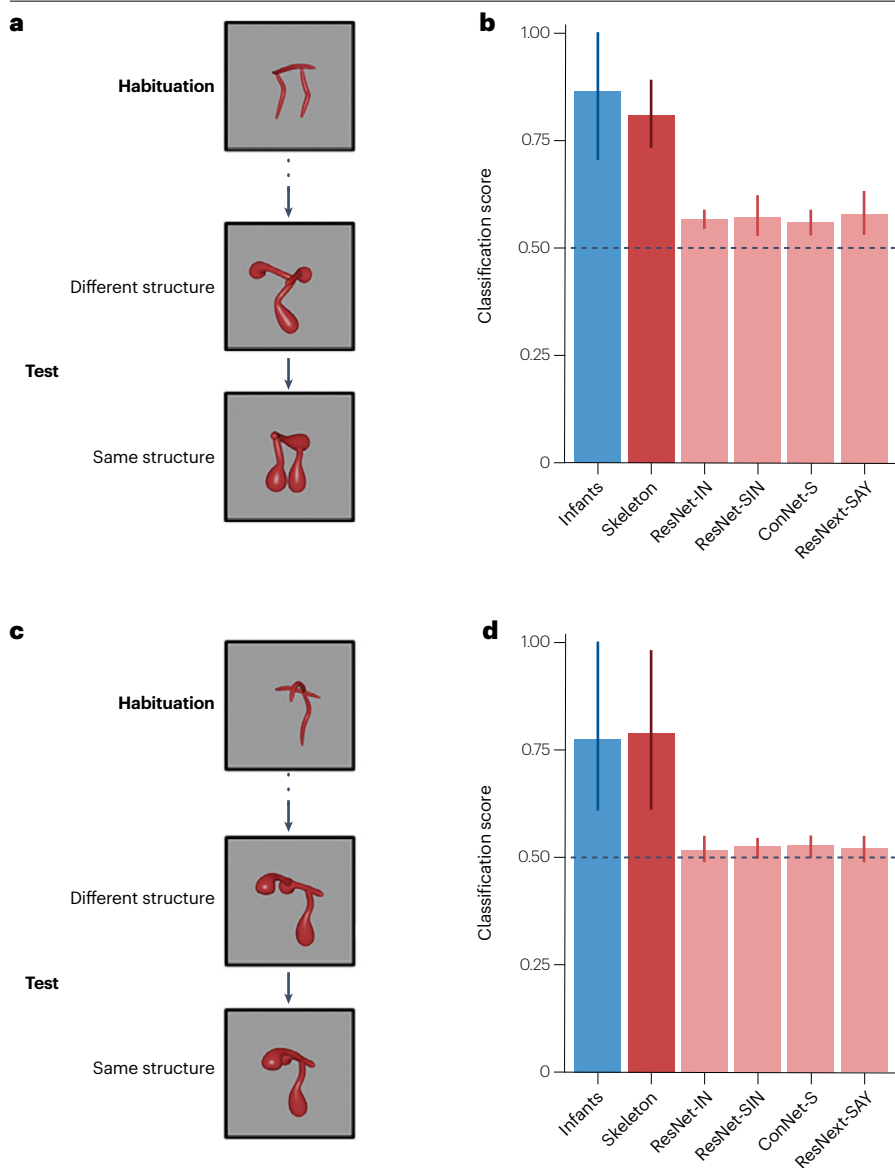


Fig. 4 | One-shot object categorization in human infants. a, b, Illustration of the experimental procedure administered to infants. In two experiments, infants were habituated to one object and then tested with objects that did or did not match the target in global shape structure. Both types of test object differed in their local features from the habituation object. In the first experiment (top), the familiar and novel test objects had qualitatively different structures. In the second experiment, the familiar and novel test object were better visually matched. **c, d,** Infant one-shot categorization performance relative to a skeletal model of global shape (Skeleton), deep neural networks trained on naturalistic images from ImageNet (ResNet-IN and ConNet-S), images with greater amounts of variability from Stylized ImageNet (ResNet-SIN) and infant head-cam videos (ResNext-SAY). Infant performance was best matched by the skeleton model, suggesting sensitivity to global shape. Error bars depict bootstrapped confidence intervals (95% CI). Dotted line indicates chance performance (0.5). Adapted from ref. 69, CC BY 4.0 (<https://creativecommons.org/licenses/by/4.0/>).

across variations in colour and texture and use this information to learn new object labels quickly^{31,167–169}. This finding suggests that by at least 6 months of age, infants seem to have acquired inductive biases that guide categorization for novel objects.

Studies using head-mounted cameras (headcams) to investigate the visual statistics (or ‘visual diet’) of the environment of infants have suggested that infants accomplish few-shot categorization without extensive experience with different objects. These studies show that infants’ visual experience in the first year of life is dominated by about ten unique objects and only three faces on average^{170–172}. However, despite the sparsity of objects in the visual world of infants, their object categorization abilities can match or surpass the object recognition abilities of many state-of-the-art deep neural networks^{25,69,173}, which are often trained on several thousands of exemplars from hundreds of categories. As noted above, infants can learn novel categories from just a few exemplars⁶⁹ (Fig. 4) and, by 4 years of age, show greater

robustness to image distortions than state-of-the-art deep neural networks^{30,85,173,174}.

Infants’ visual experience with the few objects in their environment is extensive and seemingly biased to encourage category learning¹⁷⁵. Specifically, infants’ object viewpoints are not normally distributed around the objects but are biased to planar orientations that maximally display the object shape structure^{176,177}. For instance, if an infant is holding a toy intended to represent a dog, they will often rotate the toy to select views that best depict the dog’s shape, such as from the side. Thus, although infants are exposed to fewer total objects than a typical deep neural network, the visual experience they receive is comprised of a much larger sample of highly diagnostic viewpoints per object. These findings are consistent with a growing literature that suggests that infants ‘curate’ their own experience to support visual learning by adjusting where they fixate or how they manipulate objects^{175,178,179}.

Training deep neural networks using infant headcam videos or the visual diet of other newborn animals (such as chicks) leads to above-chance object recognition performance^{180,181} and reasonable fits between the internal representations of the models and the multivariate response of the adult ventral visual pathway (though the fits are weaker than with other models)¹⁸². Indeed, deep neural networks are better at object categorization if they are trained with frames from infants' headcam videos than with frames from adult headcam videos¹⁸³. Nevertheless, even with naturalistic visual experience, these models still fall short of the object-recognition abilities of human infants, young children and other juvenile animals on many recognition tasks, such as categorizing on the basis of shape^{69,173,184}. These findings suggest that deep neural networks might still be missing important architectural features or appropriate learning objectives found in human infants.

Neural category selectivity

Although infants show sophisticated categorization abilities, category representations in the ventral visual pathway are slow to develop. The adult occipitotemporal cortex exhibits regions of selectivity to familiar categories such as faces, bodies, places¹³ and objects¹⁸⁵. Selectivity is typically measured as a statistically significantly greater neural response to one category (such as faces) than to another presented category (such as objects). Extensive visual experience is important for the development of category selective regions in the ventral pathway. For instance, the visual word form area, a region selectively involved in perceiving written text, develops only after children or adults learn to recognize letters or read words^{186,187}. Likewise, control-rearing studies in monkeys show that category selectivity for faces develops only following extensive experience with faces. Monkeys deprived of face experience showed selectivity for other experienced categories (such as hands), but not for faces¹⁸⁸ (Fig. 5). It seems likely that neural selectivity would emerge for any category with which humans gain extensive experience, especially during childhood¹⁸⁹. Indeed, intensively training monkeys to discriminate between exemplars of a synthetically

generated object category led to the emergence of category selective responses to those objects in juvenile monkeys, but not in adults who were similarly trained¹⁹⁰.

The occipitotemporal cortex of infants exhibits a protracted developmental trajectory with selectivity to faces arising earlier (2–6 months of age) than to other categories (such as bodies and landscapes around 6–15 months of age)^{191–194}. However, fully adult-like selective category representations in occipitotemporal cortex might not mature until adolescence^{46,47}, or even young adulthood¹⁹⁵. Although functional selectivity for different categories in infant occipitotemporal cortex is still developing, category information can be accurately decoded from the distributed pattern of infants' neural responses^{196,197}. This result suggests that the mechanisms supporting object recognition in infancy might be different, and more distributed, than those in adults. Indeed, comparisons between infants, the representations of the adult ventral visual pathway and deep neural networks, show that, in young infants (4–6 months of age), categorization abilities are better explained by the representations of adult early visual cortex and early layers of deep neural network models (which represent early visual cortex-like simple visual features)^{197,198} than by adult occipitotemporal cortex representations or later layers of deep neural network models (which are predictive of adult occipitotemporal cortex representations)^{23,199}. Consistent with this behavioural pattern of development, the early visual areas (V1–V3) are among the earliest regions of the brain to develop and display signatures of mature functioning at birth^{200–203} (Box 1). However, with increased age, infants' behavioural categorization performance becomes increasingly better described by higher-level layers of the models¹⁹⁸, with the category representations of 18-month-olds well predicted by the multivariate response of adult occipitotemporal cortex²⁰⁴.

Beyond the ventral pathway, categorization on the basis of global shape might be partly supported by the dorsal visual pathway¹⁴. The dorsal pathway receives coarse, low-spatial-frequency information from magnocellular cells and does not have the fidelity to represent the fine-grained details of local features. Moreover, the dorsal pathway

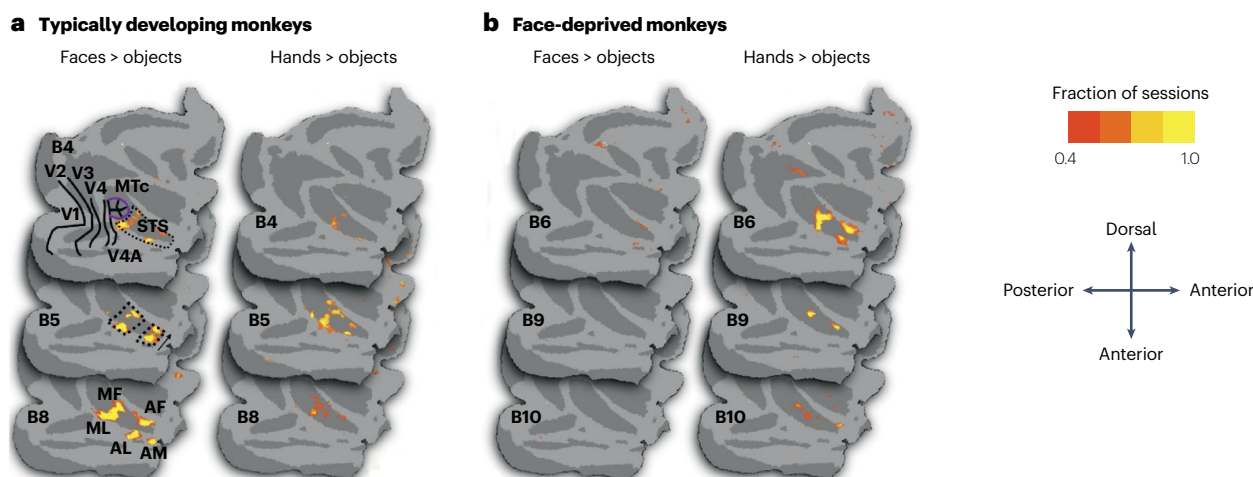


Fig. 5 | Experience is necessary for the development of category-selectivity. Images depict flattened cortical maps. Selectivity (greater response to faces or hands than objects) is shown in yellow. **a**, Monkeys reared with typical visual experience show normal selectivity for both faces and hands. **b**, By contrast, monkeys reared in the absence of face experience show normal selectivity

for hands, but not faces. B4–B10, monkey subjects 1–6; V1–V4, visual areas 1–4; V4A, anterior visual area 4; MTC, caudal middle temporal area; STS, superior temporal sulcus; MF, middle fundus; AF, anterior fundus; ML, middle lateral; AL, anterior lateral; AM, anterior middle. Reprinted from ref. 188, Springer Nature Limited.

is heavily involved in computing the spatial arrangement of object features, enabling representation of an object’s global shape structure¹⁸. In adults, object category information can be decoded from the multivariate response of dorsal regions involved in computing the spatial arrangements of features. These regions might mediate global shape representations in occipitotemporal cortex^{18,120} and contribute to representations of categories like faces^{19,119}. Combined with accumulating evidence that the dorsal pathway also develops earlier than the ventral pathway^{149,150,205} (Box 1), these findings suggest that object recognition might primarily be supported by the dorsal pathway in early development, but that the ventral pathway takes an increasingly prominent role in later development.

Together, these findings suggest that infants exhibit robust categorization abilities from early in life. Unlike conventional machine learning models, infants’ ability to categorize new objects does not depend on extensive experience with many objects: instead, infants ‘curate’ their visual experience with the few objects in their environment to learn diagnostic object properties such as global shape. However, despite the sophistication of infants’ categorization abilities, category-selective regions of the ventral pathway are not yet fully developed and categorization in infancy might instead be supported by early visual areas and the dorsal pathway. Thus, a recurring pattern in human development is that many object-recognition abilities are present from early in development, but might be supported by different neural mechanisms than in adulthood and might require extended time to reach their mature state.

Innate constraints on object recognition

Although visual experience is necessary for infants’ visual abilities to mature fully, several innate constraints, such as blurry vision and sensitivity to motion, guide infants’ abilities to form holistic 3D shape

representations^{206,207}. In this section, we take a broad view of the degree to which innate mechanisms guide specific object-recognition abilities (Table 1).

In the domain of categorization, early sensitivity to cues such as curvilinearity and rectilinearity might support rudimentary distinctions between animate and inanimate objects, such as dogs and furniture^{208–210}. Newborns also display innate sensitivities for visual properties that are correlated with important object categories such as faces, including symmetry and top-heaviness (top-heavy stimuli have more elements at the top than the bottom)^{211–213}. Although there is not yet a complete understanding of the neural correlates of these visual features, some of these properties might be represented by subcortical regions^{214–216} and might account for infants’ bias for face-like stimuli at birth^{217–219}. Structural and anatomical properties such as retinotopy and connectivity^{223,220} might also guide visual development. For example, category representations for items that are typically foveated (such as manipulable objects, faces and words) typically develop along occipitotemporal cortex regions with preferential connectivity to foveal cortex in early visual cortex^{206,207,221}.

Computational studies have begun to explore how such visual biases might arise in the newborn visual system. One study revealed that deep neural networks with randomly initialized weights (and therefore no predefined functional selectivity) nevertheless display rudimentary object-recognition abilities, and that their internal neuron-like units exhibited sensitivities to categories such as faces^{222,223}. With a large Gaussian distribution of random weights, some units will exhibit sensitivity to simple features like curvilinearity and top-heaviness by chance. This finding raises the possibility that object-recognition abilities and neural responses in infants might also arise from a visual system with little predefined functional selectivity. Moreover, these simulations can shed light on the potential processes by which category

Table 1 | Innate and experiential contributions to the development of object recognition

Visual function	Finding	Implications
Shape perception	Newborns can discriminate simple two-dimensional shapes and exhibit shape constancy ³⁶	Rudimentary shape-perception abilities might be innate and sufficient to form holistic shape representations
	Newborns exhibit a global-precedence effect ⁶⁸	
	Most perceptual grouping abilities do not develop until several months after birth ⁸⁹	The ability to form holistic shape representations is bootstrapped by blurry vision early in development
	Patients with sight restored late in development exhibit poor perceptual grouping ¹⁰²	
	Low visual acuity early in deep neural network training improves holistic perception ¹⁰⁵	
	The magnocellular pathway develops earlier than the parvocellular pathway ¹¹²	
Viewpoint-invariance	Ocular dominance columns are present at birth ¹²³	Infants might be born with the ability to perceive some three-dimensional shape properties and recognize objects from novel viewpoints
	Newborns are sensitive to some depth properties ¹²⁵	
	Newborn chicks and young infants can recognize objects from novel viewpoints ¹⁴³	
	Infants are better at recognizing depth-rotated objects after familiarization with moving objects ¹³¹	Experience with smooth continuous motion might be critical for forming three-dimensional shape representations
	Experience with smooth motion is needed for chicks to accomplish viewpoint-invariant recognition ¹⁴⁶	
Object categorization	Infants experience very few object classes in the first year of life ¹⁷⁰	The infant visual system is already equipped with visual representations to accomplish object recognition at birth
	Infants can accomplish one-shot categorization ⁶⁹	
	Newborns exhibit an innate bias for properties such as symmetry and top-heaviness ²¹²	
	Randomly initialized deep neural networks are capable of coarse categorical judgements ²²³	The visual representations of the neonate visual system are an emergent property of domain-general evolutionary pressures
	Visual experience is needed to develop category-selective regions ¹⁸⁸	

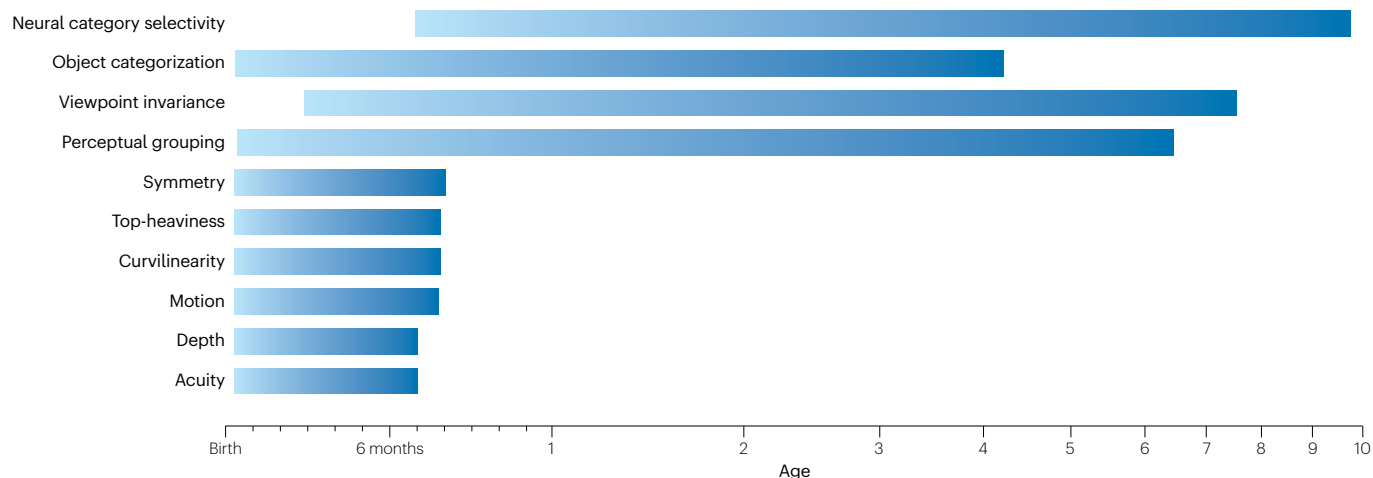


Fig. 6 | Development of visual functions. The development of visual functions in months and years. Each shaded gradient represents a different visual function. Lighter shades of the gradient indicate when each function can initially be

measured, whereas darker shades indicate when it begins to show adult-like signatures.

selectivity develops. Specifically, model units whose features happened to be similar to an object category were later co-opted to represent that category following training. For example, spuriously face-like model units became more selective for faces with visual experience, whereas units spuriously similar to other object categories became selective for those categories^{223,224}. This finding parallels the ‘cortical-recycling’ theory of development, which posits that category representations in the occipitotemporal cortex develop by co-opting, or recycling, neurons with similar functional profiles^{225–227}.

Together, these findings illustrate how the human visual system is not a ‘blank slate’ at birth, but is already equipped with many biases and constraints that support object recognition. These innate biases might not have evolved from pressures to accomplish any one specific object-recognition task, such as face perception, but might, instead, be an emergent property of a visual system evolved to adapt to an individual’s postnatal visual environment. Looking forwards, a key goal of developmental and evolutionary neuroscience is to uncover the overarching learning objectives that drive the development of the visual system²²⁸.

Summary and future directions

We have documented the developmental origins, milestones, and constraints that lead to mature object recognition (Fig. 6). Specifically, we examined how infants develop the capability to construct holistic 3D shape representations that support object recognition across changes in viewpoints and categorize objects across variations among exemplars.

Humans demonstrate traces of sophisticated object-recognition abilities from early in infancy. However, despite these abilities, the ventral visual pathway, which is the primary pathway underlying object recognition in adults, does not reach maturity until at least adolescence. Instead, these abilities might be bootstrapped by infant-specific adaptations, such as low visual acuity and self-curated visual experience, that emphasize an object’s global shape. Moreover, object recognition in infancy might be accomplished using visual properties represented by early visual areas and complemented by input from the dorsal pathway.

Finally, although innate and early developing constraints can support object recognition, extensive experience with objects is needed for recognition abilities and their neural correlates to mature.

Although psychologists, neuroscientists and computer scientists have made immense progress in understanding the processes by which humans reach maturity, there remain many open questions that are challenging to test with developmental populations. These include questions regarding biology (including about pruning^{229,230}, recurrence⁷⁵ and visual acuity¹⁰⁵), which typically require invasive methods or complicated neuroimaging techniques, and questions regarding the nature of experience (such as the necessity of experience with faces^{231–233}) that are impossible or unethical to test in humans. Indeed, as is evident from much of this Review, there are few studies that examine the neural and computational principles that guide object-recognition development.

To address these limitations, the developmental research programme should be complemented by hypothesis-driven computational modelling to explore the developmental constraints that lead to human-like behaviours and neural responses at different developmental timepoints. Indeed, because researchers are rarely able to experimentally manipulate different developmental factors, biologically plausible computational models provide an ideal testbed with which to explore developmental questions. Researchers can ‘control-rear’ computational models to have different architectural constraints, learning objectives or visual diets, and can then compare the behaviour and internal representations of the models to empirical data from human children. Such an approach could shed light on fundamental questions regarding the organization of the visual system at birth, its underlying visual representations, and the innate and experiential contributions to object recognition.

Several methodological developments are needed for this research approach to be successful. First, developmental researchers and computational scientists need to develop behavioural tasks and evaluation methods that enable human children and models to be tested on equal footing²³⁴. Object-recognition models are typically evaluated using large-scale classification tasks contrasting thousands of categories – a

task which no child (or even adult) could feasibly perform in a laboratory setting²³⁵. Instead, models should be tested using metrics that approximate human behaviours, such as looking time²³⁶. Second, researchers need more and higher-quality task-based neuroimaging studies with children. Movie watching holds particular promise in this regard because movies are rich in content, enabling researchers to ask many different questions. Importantly, movies are also engaging for infants²³⁷ and young children^{238,239}, holding their attention for long periods of time during neuroimaging scans. The same videos could be presented to models, enabling researchers to explore what kinds of mechanism lead to child-like neural responses in the models. Finally, and perhaps most importantly, developmental researchers need to formalize their theories in computationally tractable ways that can be implemented and tested with quantitative models. Although purely conceptual theories (such as box-and-arrow ‘models’) are important, computational specification of developmental theories provides concrete evidence of which developmental processes are important for an ability to develop and leads to a more replicable developmental science.

Thus, integrating classic behavioural approaches from developmental psychology with the modern tools of cognitive and computational neuroscience will reveal a deeper understanding of the developmental mechanisms that lead to mature object recognition.

Published online: 22 December 2023

References

- Feldman, J. What is a visual object? *Trends Cogn. Sci.* **7**, 252–256 (2003).
- Spelke, E. S. Principles of object perception. *Cogn. Sci.* **14**, 29–56 (1990).
- Grill-Spector, K. & Kanwisher, N. Visual recognition: as soon as you know it is there, you know what it is. *Psychol. Sci.* **16**, 152–160 (2005).
- Thorpe, S., Fize, D. & Marlot, C. Speed of processing in the human visual system. *Nature* **381**, 520–522 (1996).
- Keyser, C., Xiao, D.-K., Földiák, P. & Perrett, D. I. The speed of sight. *J. Cogn. Neurosci.* **13**, 90–101 (2001).
- Shepard, R. N. Toward a universal law of generalization for psychological science. *Science* **237**, 1317–1323 (1987).
- Tenenbaum, J. B. & Griffiths, T. L. Generalization, similarity, and Bayesian inference. *Behav. Brain. Sci.* **24**, 629–640 (2001).
- Morgenstern, Y., Schmidt, F. & Fleming, R. W. One-shot categorization of novel object classes in humans. *Vis. Res.* **165**, 98–108 (2019).
- Lake, B., Salakhutdinov, R., Gross, J. & Tenenbaum, J. One shot learning of simple visual concepts. In *Proc. Ann. Meet. Cogn. Sci. Soc. (CogSci)*, 2011.
- Tiedemann, H., Morgenstern, Y., Schmidt, F. & Fleming, R. W. One-shot generalization in humans revealed through a drawing task. *eLife* **11**, e75485 (2022).
- Marr, D. & Nishihara, H. K. Representation and recognition of the spatial organization of three-dimensional shapes. *Proc. R. Soc. Lond. Ser. B* **200**, 269–294 (1978).
- Hubel, D. H. & Wiesel, T. N. Receptive fields of single neurones in the cat's striate cortex. *J. Physiol.* **148**, 574–591 (1959).
- Kanwisher, N. & Dilks, D. D. in *The New Visual Neuroscience* (eds Chalupa, L. & J. Werner, J.) 733–748 (MIT Press, 2012).
- Ayzenberg, V. & Behrmann, M. Does the brain's ventral visual pathway compute object shape? *Trends Cogn. Sci.* **26**, 1119–1132 (2022).
- Freud, E., Behrmann, M. & Snow, J. C. What does dorsal cortex contribute to perception? *Open Mind* **4**, 40–56 (2020).
- Welchman, A. E. The human brain in depth: how we see in 3D. *Annu. Rev. Vis. Sci.* **2**, 345–376 (2016).
- Van Dromme, I. C., Premereur, E., Verhoef, B.-E., Vanduffel, W. & Janssen, P. Posterior parietal cortex drives inferotemporal activations during three-dimensional object vision. *PLoS Biol.* **14**, e1002445 (2016).
- Ayzenberg, V. & Behrmann, M. The dorsal visual pathway represents object-centered spatial relations for object recognition. *J. Neurosci.* **42**, 4693–4710 (2022).
- Zachariou, V., Nikas, C. V., Safiullah, Z. N., Gotts, S. J. & Ungerleider, L. G. Spatial mechanisms within the dorsal visual pathway contribute to the configural processing of faces. *Cereb. Cortex* **27**, 4124–4138 (2017).
- Kar, K., Kubilius, J., Schmidt, K., Issa, E. B. & DiCarlo, J. J. Evidence that recurrent circuits are critical to the ventral stream's execution of core object recognition behavior. *Nat. Neurosci.* **22**, 974–983 (2019).
- Bar, M. et al. Top-down facilitation of visual recognition. *Proc. Natl Acad. Sci. USA* **103**, 449–454 (2006).
- Zamir, A. R. et al. Taskonomy: disentangling task transfer learning. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)* 3712–3722 (IEEE, 2018).
- Blauch, N. M., Behrmann, M. & Plaut, D. C. A connectivity-constrained computational account of topographic organization in high-level visual cortex. *Proc. Natl Acad. Sci. USA* **119**, e2112566119 (2022).
- Doerig, A. et al. The neuroconnectionist research programme. *Nat. Rev. Neurosci.* **24**, 431–450 (2023).
- Zador, A. M. A critique of pure learning and what artificial neural networks can learn from animal brains. *Nat. Commun.* **10**, 3770 (2019).
- Geirhos, R. et al. Partial success in closing the gap between human and machine vision. *Adv. Neural Inf. Process. Syst.* **34**, 23885–23899 (2021).
- Biederman, I. Recognition-by-components: a theory of human image understanding. *Psychol. Rev.* **94**, 115–147 (1987).
- Mervis, C. B. & Rosch, E. Categorization of natural objects. *Annu. Rev. Psychol.* **32**, 89–115 (1981).
- Elder, J. H. & Velisavljević, L. Cue dynamics underlying rapid detection of animals in natural scenes. *J. Vision* **9**, <https://doi.org/10.1167/9.7.7> (2009).
- Geirhos, R. et al. ImageNet-trained CNNs are biased towards texture; increasing shape bias improves accuracy and robustness. Preprint at *arXiv* <https://arxiv.org/abs/1811.12231> (2018).
- Landau, B., Smith, L. B. & Jones, S. S. The importance of shape in early lexical learning. *Cogn. Dev.* **3**, 299–321 (1988).
- Wagemans, J. et al. Identification of everyday objects on the basis of silhouette and outline versions. *Perception* **37**, 207–244 (2008).
- Biederman, I. & Ju, G. Surface versus edge-based determinants of visual recognition. *Cogn. Psychol.* **20**, 38–64 (1988).
- Fant, R. L. Visual experience in infants: decreased attention to familiar patterns relative to novel ones. *Science* **146**, 668–670 (1964).
- Slater, A., Morison, V. & Rose, D. Perception of shape by the new-born baby. *Br. J. Dev. Psychol.* **1**, 135–142 (1983).
- Slater, A. & Morison, V. Shape constancy and slant perception at birth. *Perception* **14**, 337–344 (1985).
- Quinn, P. C., Eimas, P. D. & Tarr, M. J. Perceptual categorization of cat and dog silhouettes by 3- to 4-month-old infants. *J. Exp. Child. Psychol.* **79**, 78–94 (2001).
- Behrmann, M., Peterson, M. A., Moscovitch, M. & Suzuki, S. Independent representation of parts and the relations between them: evidence from integrative agnosia. *J. Exp. Psychol. Hum. Percept. Perform.* **32**, 1169–1184 (2006).
- Grill-Spector, K., Kourtzi, Z. & Kanwisher, N. The lateral occipital complex and its role in object recognition. *Vis. Res.* **41**, 1409–1422 (2001).
- Emerson, L. L., Crosswhite, S. L., Richards, J. E. & Aslin, R. N. The lateral occipital cortex is selective for object shape, not texture/color, at six months. *J. Neurosci.* **37**, 3698–3703 (2017).
- Aslin, R. N. & Mehler, J. Near-infrared spectroscopy for functional studies of brain activity in human infants: promise, prospects, and challenges. *J. Biomed. Opt.* **10**, 1083–3668 (2005).
- Wilcox, T. & Biondi, M. fNIRS in the developmental sciences. *Wiley Interdisc. Rev. Cogn. Sci.* **6**, 263–283 (2015).
- Wilcox, T. et al. Hemodynamic changes in the infant cortex during the processing of featural and spatiotemporal information. *Neuropsychologia* **47**, 657–662 (2009).
- Wilcox, T., Hawkins, L. B., Hirshkowitz, A. & Boas, D. A. Cortical activation to object shape and speed of motion during the first year. *NeuroImage* **99**, 129–141 (2014).
- Wilcox, T., Bortfeld, H., Woods, R., Wruck, E. & Boas, D. A. Hemodynamic response to featural changes in the occipital and inferior temporal cortex in infants: a preliminary methodological exploration. *Dev. Sci.* **11**, 361–370 (2008).
- Golarai, G. et al. Differential development of high-level visual cortex correlates with category-specific recognition memory. *Nat. Neurosci.* **10**, 512–522 (2007).
- Scherf, K. S., Behrmann, M., Humphreys, K. & Luna, B. Visual category-selectivity for faces, places and objects emerges along different developmental trajectories. *Dev. Sci.* **10**, F15–F30 (2007).
- Freud, E., Plaut, D. C. & Behrmann, M. Protracted developmental trajectory of shape processing along the two visual pathways. *J. Cogn. Neurosci.* **31**, 1589–1597 (2019).
- Nishimura, M., Scherf, K. S., Zachariou, V., Tarr, M. J. & Behrmann, M. Size precedes view: developmental emergence of invariant object representations in lateral occipital complex. *J. Cogn. Neurosci.* **27**, 474–491 (2015).
- Dekker, T., Mareschal, D., Sereno, M. I. & Johnson, M. H. Dorsal and ventral stream activation and object recognition performance in school-age children. *NeuroImage* **57**, 659–670 (2011).
- Wilcox, T., Stubbs, J., Hirshkowitz, A. & Boas, D. A. Functional activation of the infant cortex during object processing. *NeuroImage* **62**, 1833–1840 (2012).
- Laurence, S. F. & Huttenlocher, J. The representation of geometric cues in infancy. *Infancy* **13**, 103–127 (2008).
- Dillon, M. R., Izard, V. & Spelke, E. S. Infants' sensitivity to shape changes in 2D visual forms. *Infancy* **25**, 618–639 (2020).
- Slater, A., Mattock, A., Brown, E. & Bremner, J. G. Form perception at birth: revisited. *J. Exp. Child. Psychol.* **51**, 395–406 (1991).
- Cohen, L. B. & Younger, B. A. Infant perception of angular relations. *Infant. Behav. Dev.* **7**, 37–47 (1984).
- Bhatt, R. S. & Waters, S. E. Perception of three-dimensional cues in early infancy. *J. Exp. Child. Psychol.* **70**, 207–224 (1998).
- Kavšek, M., Jonas, A. & Granrud, C. E. Infants' sensitivity to pictorial depth cues: a review and meta-analysis of looking studies. *Infant. Behav. Dev.* **35**, 109–128 (2012).

58. Biederman, I. & Cooper, E. E. Priming contour-deleted images: evidence for intermediate representations in visual object recognition. *Cogn. Psychol.* **23**, 393–419 (1991).
59. Ayzenberg, V. & Lourenco, S. F. Skeletal descriptions of shape provide unique perceptual information for object recognition. *Sci. Rep.* **9**, 9359 (2019).
60. Wagemans, J. et al. A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure-ground organization. *Psychol. Bull.* **138**, 1172–1217 (2012).
61. Quinn, P. C., Bhatt, R. S., Brush, D., Grimes, A. & Sharpnack, H. Development of form similarity as a Gestalt grouping principle in infancy. *Psychol. Sci.* **13**, 320–328 (2002).
62. Schmidt, H. & Spelke, E. Gestalt relations and object perception in infancy. *Infant. Behav. Dev.* **7**, 319 (1984).
63. Farroni, T., Valenza, E., Simion, F. & Umiltà, C. Configural processing at birth: evidence for perceptual organisation. *Perception* **29**, 355–372 (2000).
64. Johnson, S. P. & Aslin, R. N. Perception of object unity in 2-month-old infants. *Dev. Psychol.* **31**, 739–745 (1995).
65. Slater, A., Johnson, S. P., Brown, E. & Badenoch, M. Newborn infant's perception of partly occluded objects. *Infant. Behav. Dev.* **19**, 145–148 (1996).
66. Kellman, P. J. & Spelke, E. S. Perception of partly occluded objects in infancy. *Cogn. Psychol.* **15**, 483–524 (1983).
67. Slater, A. et al. Newborn and older infants' perception of partly occluded objects. *Infant. Behav. Dev.* **13**, 33–49 (1990).
68. Cassia, V. M., Simion, F., Milani, I. & Umiltà, C. Dominance of global visual properties at birth. *J. Exp. Psychol. Gen.* **131**, 398 (2002).
69. Ayzenberg, V. & Lourenco, S. Perception of an object's global shape is best described by a model of skeletal structure in human infants. *eLife* **11**, e74943 (2022).
70. Ghim, H.-R. & Eimas, P. D. Global and local processing by 3- and 4-month-old infants. *Percept. Psychophys.* **43**, 165–171 (1988).
71. von der Heydt, R. Figure-ground organization and the emergence of proto-objects in the visual cortex. *Front. Psychol.* **6**, 1695 (2015).
72. Lee, T. S., Mumford, D., Romero, R. & Lamme, V. A. The role of the primary visual cortex in higher level vision. *Vis. Res.* **38**, 2429–2454 (1998).
73. Wokke, M. E., Vandenbroucke, A. R., Scholte, H. S. & Lamme, V. A. Confuse your illusion: feedback to early visual cortex contributes to perceptual completion. *Psychol. Sci.* **24**, 63–71 (2013).
74. Tang, H. et al. Recurrent computations for visual pattern completion. *Proc. Natl Acad. Sci. USA* **115**, 8835–8840 (2018).
75. Kietzmann, T. C. et al. Recurrence is required to capture the representational dynamics of the human visual system. *Proc. Natl Acad. Sci. USA* **116**, 21854–21863 (2019).
76. Schrimpf, M. et al. Brain-score: which artificial neural network for object recognition is most brain-like? Preprint at *bioRxiv* <https://doi.org/10.1101/407007> (2018).
77. Batardière, A. et al. Early specification of the hierarchical organization of visual cortical areas in the macaque monkey. *Cereb. Cortex* **12**, 453–465 (2002).
78. Burkhalter, A. Development of forward and feedback connections between areas V1 and V2 of human visual cortex. *Cereb. Cortex* **3**, 476–487 (1993).
79. 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).
80. Burkhalter, A., Bernardo, K. L. & Charles, V. Development of local circuits in human visual cortex. *J. Neurosci.* **13**, 1916–1931 (1993).
81. Smyser, C. D. et al. Longitudinal analysis of neural network development in preterm infants. *Cereb. Cortex* **20**, 2852–2862 (2010).
82. Nagy, Z., Westerberg, H. & Klingberg, T. Maturation of white matter is associated with the development of cognitive functions during childhood. *J. Cogn. Neurosci.* **16**, 1227–1233 (2004).
83. Emberson, L. L., Richards, J. E. & Aslin, R. N. Top-down modulation in the infant brain: learning-induced expectations rapidly affect the sensory cortex at 6 months. *Proc. Natl Acad. Sci. USA* **112**, 9585–9590 (2015).
84. Kouider, S. et al. A neural marker of perceptual consciousness in infants. *Science* **340**, 376–380 (2013).
85. Ayzenberg, V. & Lourenco, S. Young children outperform feed-forward and recurrent neural networks on challenging object recognition tasks. *J. Vis.* **20**, 310–310 (2020).
86. Káldy, Z. & Kovács, I. Visual context integration is not fully developed in 4-year-old children. *Perception* **32**, 657–666 (2003).
87. Kovács, I. Human development of perceptual organization. *Vis. Res.* **40**, 1301–1310 (2000).
88. Kovács, I., Kozma, P., Fehér, A. & Benedek, G. Late maturation of visual spatial integration in humans. *Proc. Natl Acad. Sci. USA* **96**, 12204–12209 (1999).
89. Scherf, K. S., Behrmann, M., Kimchi, R. & Luna, B. Emergence of global shape processing continues through adolescence. *Child. Dev.* **80**, 162–177 (2009).
90. Atkinson, J. Human visual development over the first 6 months of life. A review and a hypothesis. *Hum. Neurobiol.* **3**, 61–74 (1983).
91. Banks, M. S. & Salapatek, P. Infant pattern vision: a new approach based on the contrast sensitivity function. *J. Exp. Child. Psychol.* **31**, 1–45 (1981).
92. Atkinson, J., Braddick, O. & Braddick, F. Acuity and contrast sensitivity of infant vision. *Nature* **247**, 403–404 (1974).
93. Brown, A. M. & Yamamoto, M. Visual acuity in newborn and preterm infants measured with grating acuity cards. *Am. J. Ophthalmol.* **102**, 245–253 (1986).
94. Sokol, S. Measurement of infant visual acuity from pattern reversal evoked potentials. *Vis. Res.* **18**, 33–39 (1978).
95. Newport, E. L. Maturation constraints on language learning. *Cogn. Sci.* **14**, 11–28 (1990).
96. Elman, J. L. Learning and development in neural networks: the importance of starting small. *Cognition* **48**, 71–99 (1993).
97. Bjorklund, D. F. The role of immaturity in human development. *Psychol. Bull.* **122**, 153–169 (1997).
98. Lickliter, R. Premature visual stimulation accelerates intersensory functioning in bobwhite quail neonates. *Dev. Psychobiol.* **23**, 15–27 (1990).
99. Lickliter, R. & Hellewell, T. B. in *Developmental Time and Timing* (eds Lurkewitz, G. & Devenny, D. A.) 105–124 (Lawrence Erlbaum, 1993).
100. Kenny, P. A. & Turkewitz, G. Effects of unusually early visual stimulation on the development of homing behavior in the rat pup. *Dev. Psychobiol.* **19**, 57–66 (1986).
101. Harlow, H. F. The development of learning in the rhesus monkey. *Sci. Prog.* **12**, 239–269 (1959).
102. Ostrovsky, Y., Meyers, E., Ganesh, S., Mathur, U. & Sinha, P. Visual parsing after recovery from blindness. *Psychol. Sci.* **20**, 1484–1491 (2009).
103. Le Grand, R., Mondloch, C. J., Maurer, D. & Brent, H. P. Early visual experience and face processing. *Nature* **410**, 890–890 (2001).
104. Ellemberg, D., Lewis, T. L., Maurer, D., Brar, S. & Brent, H. P. Better perception of global motion after monocular than after binocular deprivation. *Vis. Res.* **42**, 169–179 (2002).
105. Vogelsang, L. et al. Potential downside of high initial visual acuity. *Proc. Natl Acad. Sci. USA* **115**, 11333–11338 (2018).
106. Jang, H. & Tong, F. Convolutional neural networks trained with a developmental sequence of blurry to clear images reveal core differences between face and object processing. *J. Vis.* **21**, 6 (2021).
107. Jinsi, O., Henderson, M. M. & Tarr, M. J. Early experience with low-pass filtered images facilitates visual category learning in a neural network model. *PLoS One* **18**, e0280145 (2023).
108. Wang, W., Zhou, T., Chen, L. & Huang, Y. A subcortical magnocellular pathway is responsible for the fast processing of topological properties of objects: a transcranial magnetic stimulation study. *Hum. Brain Res.* **44**, 1617–1628 (2022).
109. Felleman, D. J. & Van Essen, D. C. Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* **1**, 1–47 (1991).
110. Rakic, P., Barlow, H. B. & Gaze, R. M. Prenatal development of the visual system in rhesus monkey. *Phil. Trans. R. Soc. Lond. B* **278**, 245–260 (1977).
111. Kogan, C. S., Zangenehpour, S. & Chaudhuri, A. Developmental profiles of SMI-32 immunoreactivity in monkey striate cortex. *Dev. Brain Res.* **119**, 85–95 (2000).
112. Hammarrenger, B. et al. Magnocellular and parvocellular developmental course in infants during the first year of life. *Doc. Ophthalmol.* **107**, 225–233 (2003).
113. Arsenovic, A., Ischebeck, A. & Zaretskaya, N. Dissociation between attention-dependent and spatially specific illusory shape responses within the topographic areas of the posterior parietal cortex. *J. Neurosci.* **42**, 8125–8135 (2022).
114. Grassi, P. R., Zaretskaya, N. & Bartels, A. A generic mechanism for perceptual organization in the parietal cortex. *J. Neurosci.* **38**, 7158–7169 (2018).
115. Riddoch, M. J. et al. A tale of two agnosias: distinctions between form and integrative agnosia. *Cogn. Neuropsychol.* **25**, 56–92 (2008).
116. Romei, V., Driver, J., Schyns, P. G. & Thut, G. Rhythmic TMS over parietal cortex links distinct brain frequencies to global versus local visual processing. *Curr. Biol.* **21**, 334–337 (2011).
117. Zaretskaya, N., Anstis, S. & Bartels, A. Parietal cortex mediates conscious perception of illusory gestalt. *J. Neurosci.* **33**, 523–531 (2013).
118. Guo, C. et al. Adversarially trained neural representations are already as robust as biological neural representations. In *Proc. Int. Conf. Machine Learning* Vol. 162 (eds Chaudhuri, K. et al.) 8072–8081 (Proc. Machine Learning Research, 2022).
119. Waidmann, E. N., Koyano, K. W., Hong, J. J., Russ, B. E. & Leopold, D. A. Local features drive identity responses in macaque anterior face patches. *Nat. Commun.* **13**, 5592 (2022).
120. Ayzenberg, V., Simmons, C. & Behrmann, M. Temporal asymmetries and interactions between dorsal and ventral visual pathways during object recognition. *Cereb. Cortex Comm.* **4**, tgad003 (2022).
121. Tarr, M. J. & Bülthoff, H. H. Image-based object recognition in man, monkey and machine. *Cognition* **67**, 1–20 (1998).
122. Humphrey, G. K. & Jolicoeur, P. An examination of the effects of axis foreshortening, monocular depth cues, and visual field on object identification. *Q. J. Exp. Psychol.* **46**, 137–159 (1993).
123. 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).
124. Chino, Y. M., Smith, E. L. III, Hatta, S. & Cheng, H. Postnatal development of binocular disparity sensitivity in neurons of the primate visual cortex. *J. Neurosci.* **17**, 296–307 (1997).
125. Slater, A., Mattock, A. & Brown, E. Size constancy at birth: newborn infants' responses to retinal and real size. *J. Exp. Child. Psychol.* **49**, 314–322 (1990).
126. Slater, A., Morison, V. & Rose, D. New-born infants' perception of similarities and differences between two- and three-dimensional stimuli. *Br. J. Dev. Psychol.* **2**, 287–294 (1984).
127. Jandó, G. et al. Early-onset binocularity in preterm infants reveals experience-dependent visual development in humans. *Proc. Natl Acad. Sci. USA* **109**, 11049–11052 (2012).
128. Fox, R., Aslin, R. N., Shea, S. L. & Dumais, S. T. Stereopsis in human infants. *Science* **207**, 323–324 (1980).

129. Hirshkowitz, A. & Wilcox, T. Infants' ability to extract three-dimensional shape from coherent motion. *Infant. Behav. Dev.* **36**, 863–872 (2013).
130. Kellman, P. J. & Short, K. R. Development of three-dimensional form perception. *J. Exp. Psychol. Hum. Percept. Perform.* **13**, 545 (1987).
131. Kellman, P. J. Perception of three-dimensional form by human infants. *Percept. Psychophys.* **36**, 353–358 (1984).
132. Shuwairi, S. M., Albert, M. K. & Johnson, S. P. Discrimination of possible and impossible objects in infancy. *Psychol. Sci.* **18**, 303–307 (2007).
133. Tsuruhara, A., Sawada, T., Kanazawa, S., Yamaguchi, M. K. & Yonas, A. Infant's ability to form a common representation of an object's shape from different pictorial depth cues: a transfer-across-cues study. *Infant Behav. Dev.* **32**, 468–475 (2009).
134. Mash, C., Arterberry, M. E. & Bornstein, M. H. Mechanisms of visual object recognition in infancy: five-month-olds generalize beyond the interpolation of familiar views. *Infancy* **12**, 31–43 (2007).
135. Ruff, H. A. Infant recognition of the invariant form of objects. *Child. Dev.* **49**, 293–306 (1978).
136. Kraebel, K. S. & Gerhardstein, P. C. Three-month-old infants' object recognition across changes in viewpoint using an operant learning procedure. *Infant Behav. Dev.* **29**, 11–23 (2006).
137. Georgieva, S., Peeters, R., Kolster, H., Todd, J. T. & Orban, G. A. The processing of three-dimensional shape from disparity in the human brain. *J. Neurosci.* **29**, 727–742 (2009).
138. Georgieva, S. S., Todd, J. T., Peeters, R. & Orban, G. A. The extraction of 3D shape from texture and shading in the human brain. *Cereb. Cortex* **18**, 2416–2438 (2008).
139. Orban, G. A. The extraction of 3D shape in the visual system of human and nonhuman primates. *Annu. Rev. Neurosci.* **34**, 361–388 (2011).
140. Yamane, Y., Carlson, E. T., Bowman, K. C., Wang, Z. & Connor, C. E. A neural code for three-dimensional object shape in macaque inferotemporal cortex. *Nat. Neurosci.* **11**, 1352–1360 (2008).
141. Murphy, A. P., Leopold, D. A., Humphreys, G. W. & Welchman, A. E. Lesions to right posterior parietal cortex impair visual depth perception from disparity but not motion cues. *Phil. Trans. R. Soc. B* **371**, 20150263 (2016).
142. Tarr, M. J. & Bülthoff, H. H. Is human object recognition better described by geon structural descriptions or by multiple views? Comment on Biederman and Gerhardstein. *J. Exp. Psychol. Hum. Percept. Perform.* **21**, 1494–1505 (1995).
143. Wood, J. N. Newborn chickens generate invariant object representations at the onset of visual object experience. *Proc. Natl Acad. Sci. USA* **110**, 14000–14005 (2013).
144. Wood, J. N. & Wood, S. M. W. One-shot learning of view-invariant object representations in newborn chicks. *Cognition* **199**, 104192 (2020).
145. Kellman, P. J. & Shipley, T. F. A theory of visual interpolation in object perception. *Cogn. Psychol.* **23**, 141–221 (1991).
146. Wood, J. N. & Wood, S. M. The development of invariant object recognition requires visual experience with temporally smooth objects. *Cogn. Sci.* **42**, 1391–1406 (2018).
147. Ye, J. et al. Resilience of temporal processing to early and extended visual deprivation. *Vis. Res.* **186**, 80–86 (2021).
148. Ben-Ami, S. et al. Human (but not animal) motion can be recognized at first sight — after treatment for congenital blindness. *Neuropsychologia* **174**, 108307 (2022).
149. Bourne, J. A. & Rosa, M. G. 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).
150. Ciesielski, K. T. et al. Maturation changes in human dorsal and ventral visual networks. *Cereb. Cortex* **29**, 5131–5149 (2019).
151. Distler, C., Bachevalier, J., Kennedy, C., Mishkin, M. & Ungerleider, L. Functional development of the corticocortical pathway for motion analysis in the macaque monkey: a ¹⁴C-2-deoxyglucose study. *Cereb. Cortex* **6**, 184–195 (1996).
152. Biagi, L., Tosetti, M., Crespi, S. A. & Morrone, M. C. Development of BOLD response to motion in human infants. *J. Neurosci.* **43**, 3825–3837 (2023).
153. Rosch, E. in *Cognition and Categorization* (eds Rosch, E. & Lloyd, B.) 27–48 (Erlbaum, 1978).
154. Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M. & Boyes-Braem, P. in *Cognitive Psychology: Key Readings* (eds Balota, D. A. & Marsh, E. J.) 448–471 (Psychology Press, 2004).
155. Mareschal, D. & Quinn, P. C. Categorization in infancy. *Trends Cogn. Sci.* **5**, 443–450 (2001).
156. Turati, C., Simion, F. & Zanon, L. Newborns' perceptual categorization for closed and open geometric forms. *Infancy* **4**, 309–325 (2003).
157. Quinn, P. C., Slater, A. M., Brown, E. & Hayes, R. A. Developmental change in form categorization in early infancy. *Br. J. Dev. Psychol.* **19**, 207–218 (2001).
158. Bomba, P. C. & Siqueland, E. R. The nature and structure of infant form categories. *J. Exp. Child. Psychol.* **35**, 294–328 (1983).
159. Quinn, P. C., Eimas, P. D. & Rosenkrantz, S. L. Evidence for representations of perceptually similar natural categories by 3-month-old and 4-month-old infants. *Perception* **22**, 463–475 (1993).
160. Quinn, P. C. & Johnson, M. H. Global-before-basic object categorization in connectionist networks and 2-month-old infants. *Infancy* **1**, 31–46 (2000).
161. Mareschal, D., French, R. M. & Quinn, P. C. A connectionist account of asymmetric category learning in early infancy. *Dev. Psychol.* **36**, 635–645 (2000).
162. Oakes, L. M. & Spalding, T. L. The role of exemplar distribution in infants' differentiation of categories. *Infant. Behav. Dev.* **20**, 457–475 (1997).
163. Quinn, P. C. The categorical representation of visual pattern information by young infants. *Cognition* **27**, 145–179 (1987).
164. Sorscher, B., Ganguli, S. & Sompolinsky, H. Neural representational geometry underlies few-shot concept learning. *Proc. Natl Acad. Sci. USA* **119**, e2200800119 (2022).
165. Feldman, J. The structure of perceptual categories. *J. Math. Psychol.* **41**, 145–170 (1997).
166. Feldman, J. & Singh, M. Bayesian estimation of the shape skeleton. *Proc. Natl Acad. Sci. USA* **103**, 18014–18019 (2006).
167. Landau, B., Smith, L. & Jones, S. Object perception and object naming in early development. *Trends Cogn. Sci.* **2**, 19–24 (1998).
168. Smith, L. B., Jones, S. S. & Landau, B. Naming in young children: a dumb attentional mechanism? *Cognition* **60**, 143–171 (1996).
169. Smith, L. B. Learning to recognize objects. *Psychol. Sci.* **14**, 244–250 (2003).
170. Clerkin, E. M., Hart, E., Rehag, J. M., Yu, C. & Smith, L. B. Real-world visual statistics and infants' first-learned object names. *Phil. Trans. R. Soc. B* **372**, 20160055 (2017).
171. Jayaraman, S., Fausey, C. M. & Smith, L. B. The faces in infant-perspective scenes change over the first year of life. *PLoS One* **10**, e0123780 (2015).
172. Jayaraman, S. & Smith, L. B. Faces in early visual environments are persistent not just frequent. *Vis. Res.* **157**, 213–221 (2019).
173. Tartaglino, A. R., Vong, W. K. & Lake, B. M. A developmentally-inspired examination of shape versus texture bias in machines. Preprint at *arXiv* <https://arxiv.org/abs/1811.12231> (2022).
174. Huber, L. S., Geirhos, R. & Wichmann, F. A. A four-year-old can outperform ResNet-50: out-of-distribution robustness may not require large-scale experience. In *3rd Worksh. on Shared Visual Representations in Human and Machine Intelligence (SVRHM)* (NeurIPS, 2021).
175. Smith, L. B., Jayaraman, S., Clerkin, E. & Yu, C. The developing infant creates a curriculum for statistical learning. *Trends Cogn. Sci.* **22**, 325–336 (2018).
176. Slone, L. K., Smith, L. B. & Yu, C. Self-generated variability in object images predicts vocabulary growth. *Dev. Sci.* **22**, e12816 (2019).
177. James, K. H., Jones, S. S., Smith, L. B. & Swain, S. N. Young children's self-generated object views and object recognition. *J. Cogn. Dev.* **15**, 393–401 (2014).
178. Perez, J. & Feigenson, L. Violations of expectation trigger infants to search for explanations. *Cognition* **218**, 104942 (2022).
179. Stahl, A. E. & Feigenson, L. Observing the unexpected enhances infants' learning and exploration. *Science* **348**, 91–94 (2015).
180. Lee, D., Gujarathi, P. & Wood, J. N. Controlled-rearing studies of newborn chicks and deep neural networks. In *3rd Worksh. on Shared Visual Representations in Human and Machine Intelligence (SVRHM)* (NeurIPS, 2021).
181. Orhan, E. A., Gupta, P. V. & Lake, B. M. Self-supervised learning through the eyes of a child. In *Advances in Neural Information Processing Systems* 116 (NeurIPS, 2021).
182. Zhuang, C. et al. Unsupervised neural network models of the ventral visual stream. *Proc. Natl Acad. Sci. USA* **118**, e2014196118 (2021).
183. Bamach, S., Crandall, D. J., Smith, L. B. & Yu, C. in *Joint Int. Conf. on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)* 290–295 (IEEE, 2017).
184. Pak, D., Lee, D., Wood, S. M. & Wood, J. N. A newborn embodied Turing test for view-invariant object recognition. Preprint at *arXiv* <https://arxiv.org/abs/2306.05582> (2023).
185. Rajalingham, R. & DiCarlo, J. J. Reversible inactivation of different millimeter-scale regions of primate IT results in different patterns of core object recognition deficits. *Neuron* **102**, 493–505. e495 (2019).
186. Dehaene, S. et al. How learning to read changes the cortical networks for vision and language. *Science* **330**, 1359–1364 (2010).
187. Saygin, Z. M. et al. Connectivity precedes function in the development of the visual word form area. *Nat. Neurosci.* **19**, 1250–1255 (2016).
188. Arcaro, M. J., Schade, P. F., Vincent, J. L., Ponce, C. R. & Livingstone, M. S. Seeing faces is necessary for face-domain formation. *Nat. Neurosci.* **20**, 1404–1412 (2017).
189. Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P. & Gore, J. C. Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. *Nat. Neurosci.* **2**, 568–573 (1999).
190. Srihasam, K., Vincent, J. L. & Livingstone, M. S. Novel domain formation reveals proto-architecture in inferotemporal cortex. *Nat. Neurosci.* **17**, 1776–1783 (2014).
191. Kosakowski, H. L. et al. Selective responses to faces, scenes, and bodies in the ventral visual pathway of infants. *Curr. Biol.* **32**, 265–274.e5 (2021).
192. Deen, B. et al. Organization of high-level visual cortex in human infants. *Nat. Commun.* **8**, 13995 (2017).
193. Powell, L. J., Deen, B. & Saxe, R. Using individual functional channels of interest to study cortical development with fNIRS. *Dev. Sci.* **21**, e12595 (2018).
194. Yan, X. et al. When do visual category representations emerge in infants' brains? Preprint at *bioRxiv* <https://doi.org/10.1101/2023.05.11.539934> (2023).
195. Germaine, L. T., Duchaine, B. & Nakayama, K. Where cognitive development and aging meet: face learning ability peaks after age 30. *Cognition* **118**, 201–210 (2011).
196. Cohen, M. A. et al. Representational similarity precedes category selectivity in the developing ventral visual pathway. *NeuroImage* **197**, 565–574 (2019).
197. Xie, S. et al. Visual category representations in the infant brain. *Curr. Biol.* **32**, 5422–5432. e5426 (2022).
198. Kiat, J. E. et al. Linking patterns of infant eye movements to a neural network model of the ventral stream using representational similarity analysis. *Dev. Sci.* **25**, e13155 (2021).

199. Yamins, D. L. et al. Performance-optimized hierarchical models predict neural responses in higher visual cortex. *Proc. Natl Acad. Sci. USA* **111**, 8619–8624 (2014).
200. Gao, W. et al. Temporal and spatial evolution of brain network topology during the first two years of life. *PLoS One* **6**, e25278 (2011).
201. Gogtay, N. et al. Dynamic mapping of human cortical development during childhood through early adulthood. *Proc. Natl Acad. Sci. USA* **101**, 8174–8179 (2004).
202. Keunen, K., Counsell, S. J. & Benders, M. J. N. L. The emergence of functional architecture during early brain development. *NeuroImage* **160**, 2–14 (2017).
203. Natu, V. S. et al. Infants' cortex undergoes microstructural growth coupled with myelination during development. *Commun. Biol.* **4**, 1191 (2021).
204. Spriet, C., Abassi, E., Hochmann, J.-R. & Papeo, L. Visual object categorization in infancy. *Proc. Natl Acad. Sci. USA* **119**, e2105866119 (2022).
205. Ayzenberg, V., Granovetter, M. C., Robert, S., Patterson, C. & Behrmann, M. Differential functional reorganization of ventral and dorsal visual pathways following childhood hemispherectomy. *Dev. Cogn. Neurosci.* **64**, 101323 (2023).
206. Kamps, F. S., Hendrix, C. L., Brennan, P. A. & Dilks, D. D. Connectivity at the origins of domain specificity in the cortical face and place networks. *Proc. Natl Acad. Sci. USA* **117**, 6163–6169 (2020).
207. Hasson, U., Levy, I., Behrmann, M., Hendler, T. & Malach, R. Eccentricity bias as an organizing principle for human high-order object areas. *Neuron* **34**, 479–490 (2002).
208. Yetter, M. et al. Curvilinear features are important for animate/inanimate categorization in macaques. *J. Vis.* **21**, 3 (2021).
209. Yue, X., Robert, S. & Ungerleider, L. G. Curvature processing in human visual cortical areas. *NeuroImage* **222**, 117295 (2020).
210. Ponce, C. R., Hartmann, T. S. & Livingstone, M. S. End-stopping predicts curvature tuning along the ventral stream. *J. Neurosci.* **37**, 648–659 (2017).
211. Cassia, V. M., Valenza, E., Simion, F. & Leo, I. Congruency as a nonspecific perceptual property contributing to newborns' face preference. *Child. Dev.* **79**, 807–820 (2008).
212. Cassia, V. M., Turati, C. & Simion, F. Can a nonspecific bias toward top-heavy patterns explain newborns' face preference? *Psychol. Sci.* **15**, 379–383 (2004).
213. Turati, C., Simion, F., Milani, I. & Umiltà, C. Newborns' preference for faces: what is crucial? *Dev. Psychol.* **38**, 875–882 (2002).
214. Johnson, M. H. Subcortical face processing. *Nat. Rev. Neurosci.* **6**, 766–774 (2005).
215. Hafed, Z. M. & Chen, C.-Y. Sharper, stronger, faster upper visual field representation in primate superior colliculus. *Curr. Biol.* **26**, 1647–1658 (2016).
216. Versace, E., Damini, S. & Stancher, G. Early preference for face-like stimuli in solitary species as revealed by tortoise hatchlings. *Proc. Natl Acad. Sci. USA* **117**, 24047–24049 (2020).
217. Johnson, M. H., Dziurawiec, S., Ellis, H. & Morton, J. Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition* **40**, 1–19 (1991).
218. Reid, V. M. et al. The human fetus preferentially engages with face-like visual stimuli. *Curr. Biol.* **27**, 1825–1828.e1823 (2017).
219. Simion, F., Valenza, E., Umiltà, C. & Barba, B. D. Preferential orienting to faces in newborns: a temporal–nasal asymmetry. *J. Exp. Psychol. Hum. Percept. Perform.* **24**, 1399 (1998).
220. Arcaro, M. J. & Livingstone, M. S. On the relationship between maps and domains in inferotemporal cortex. *Nat. Rev. Neurosci.* **22**, 573–583 (2021).
221. Gomez, J., Barnett, M. & Grill-Spector, K. Extensive childhood experience with Pokémon suggests eccentricity drives organization of visual cortex. *Nat. Hum. Behav.* **3**, 611–624 (2019).
222. Xu, S., Zhang, Y., Zhen, Z. & Liu, J. The face module emerges from domain-general visual experience: a deprivation study on deep convolution neural network. *Front. Comput. Neurosci.* **15**, 626259 (2020).
223. Baek, S., Song, M., Jang, J., Kim, G. & Paik, S.-B. Face detection in untrained deep neural networks. *Nat. Commun.* **12**, 7328 (2021).
224. Hannagan, T., Agrawal, A., Cohen, L. & Dehaene, S. Emergence of a compositional neural code for written words: recycling of a convolutional neural network for reading. *Proc. Natl Acad. Sci. USA* **118**, e2104779118 (2021).
225. Nordt, M. et al. Cortical recycling in high-level visual cortex during childhood development. *Nat. Hum. Behav.* **5**, 1686–1697 (2021).
226. Dehaene, S. & Cohen, L. Cultural recycling of cortical maps. *Neuron* **56**, 384–398 (2007).
227. Behrmann, M. & Plaut, D. C. Hemispheric organization for visual object recognition: a theoretical account and empirical evidence. *Perception* **49**, 373–404 (2020).
228. Bakhtiari, S., Mineault, P., Lillicrap, T., Pack, C. & Richards, B. The functional specialization of visual cortex emerges from training parallel pathways with self-supervised predictive learning. *Adv. Neural Inf. Process. Syst.* **34**, 25164–25178 (2021).
229. Zhu, M. & Gupta, S. To prune, or not to prune: exploring the efficacy of pruning for model compression. Preprint at [arXiv https://arxiv.org/abs/1710.01878](https://arxiv.org/abs/1710.01878) (2017).
230. Lu, H. & Erlikhman, G. Enhancement of representational sparsity in deep neural networks can improve generalization. *J. Vis.* **19**, 209b (2019).
231. Yuan, L., Xiang, V., Crandall, D. & Smith, L. Learning the generative principles of a symbol system from limited examples. *Cognition* **200**, 104243 (2020).
232. Smith, L. B. & Slone, L. K. A developmental approach to machine learning? *Front. Psychol.* **8**, 2124 (2017).
233. Blauch, N. M., Behrmann, M. & Plaut, D. C. Computational insights into human perceptual expertise for familiar and unfamiliar face recognition. *Cognition* **208**, 104341 (2020).
234. Stojnić, G., Gandhi, K., Yasuda, S., Lake, B. M. & Dillon, M. R. Commonsense psychology in human infants and machines. *Cognition* **235**, 105406 (2023).
235. Wichmann, F. A. et al. Methods and measurements to compare men against machines. *Electron. Imaging* **2017**, 36–45 (2017).
236. Yermolayeva, Y. & Rakison, D. H. Connectionist modeling of developmental changes in infancy: approaches, challenges, and contributions. *Psychol. Bull.* **140**, 224–255 (2014).
237. Yates, T. S. et al. Neural event segmentation of continuous experience in human infants. *Proc. Natl Acad. Sci. USA* **119**, e2200257119 (2022).
238. Yates, T. S., Ellis, C. T. & Turk-Browne, N. B. Emergence and organization of adult brain function throughout child development. *NeuroImage* **226**, 117606 (2021).
239. Lerner, Y., Scherf, K. S., Katkov, M., Hasson, U. & Behrmann, M. Changes in cortical coherence supporting complex visual and social processing in adolescence. *J. Cogn. Neurosci.* **33**, 2215–2230 (2021).
240. Wilcox, T., Haslup, J. A. & Boas, D. A. Dissociation of processing of featural and spatiotemporal information in the infant cortex. *NeuroImage* **53**, 1256–1263 (2010).
241. Bachevalier, J., Hagger, C. & Mishkin, M. in *Alfred Benzon Symposium Vol. 31 Brain Work And Mental Activity* (eds Lassen, N. A. et al.) 231–240 (Munksgaard, 1991).
242. Arcaro, M. J. & Livingstone, M. S. A hierarchical, retinotopic proto-organization of the primate visual system at birth. *eLife* **6**, e26196 (2017).
243. Ellis, C. T. et al. Retinotopic organization of visual cortex in human infants. *Neuron* **109**, 2616–2626.e2616 (2021).
244. 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).
245. Mohammed, C. P. D. & Khalil, R. Postnatal development of visual cortical function in the mammalian brain. *Front. Syst. Neuro.* **14**, 29 (2020).
246. Rodman, H. R., Scalaidhe, S. & Gross, C. G. Response properties of neurons in temporal cortical visual areas of infant monkeys. *J. Neurophysiol.* **70**, 1115–1136 (1993).
247. Rodman, H. R. Development of inferior temporal cortex in the monkey. *Cereb. Cortex* **4**, 484–498 (1994).
248. Kamps, F. S., Pincus, J. E., Radwan, S. F., Wahab, S. & Dilks, D. D. Late development of navigationally relevant motion processing in the occipital place area. *Curr. Biol.* **30**, 544–550.e543 (2020).
249. Grotheer, M. et al. Human white matter myelination rate slows down at birth. Preprint at [bioRxiv https://doi.org/10.1101/2023.03.02.530800v1](https://doi.org/10.1101/2023.03.02.530800v1) (2023).
250. Ahmad, Z., Behrmann, M., Patterson, C. & Freud, E. Unilateral resection of both cortical visual pathways in a pediatric patient alters action but not perception. *Neuropsychologia* **168**, 108182 (2022).
251. Grinter, E. J., Maybery, M. T. & Badcock, D. R. Vision in developmental disorders: is there a dorsal stream deficit? *Brain Res. Bull.* **82**, 147–160 (2010).
252. Pitcher, D. & Ungerleider, L. G. Evidence for a third visual pathway specialized for social perception. *Trends Cogn. Sci.* **25**, 100–110 (2021).
253. Weiner, K. S. & Gomez, J. Third visual pathway, anatomy, and cognition across species. *Trends Cogn. Sci.* **25**, 548–549 (2021).
254. Braddick, O. & Atkinson, J. Development of human visual function. *Vis. Res.* **51**, 1588–1609 (2011).
255. Dubowitz, L. M. S., De Vries, L., Mushin, J. & Arden, G. B. Visual function in the newborn infant: is it cortically mediated? *Lancet* **327**, 1139–1141 (1986).
256. Ma, Z., Tu, W. & Zhang, N. Increased wiring cost during development is driven by long-range cortical, but not subcortical connections. *NeuroImage* **225**, 117463 (2021).
257. King, A. J., Schnupp, J. W. H., Carlile, S., Smith, A. L. & Thompson, I. D. The development of topographically-aligned maps of visual and auditory space in the superior colliculus. *Prog. Brain Res.* **112**, 335–350 (1996).
258. O'Reilly, R. C., Russin, J. L., Zolfaghar, M. & Rohrlach, J. Deep predictive learning in neocortex and pulvinar. *J. Cogn. Neurosci.* **33**, 1158–1196 (2021).
259. Sowards, T. V. & Sowards, M. A. Innate visual object recognition in vertebrates: some proposed pathways and mechanisms. *Comp. Biochem. Physiol. A* **132**, 861–891 (2002).
260. Arcaro, M. J., Pinsk, M. A., Chen, J. & Kastner, S. Organizing principles of pulvino-cortical functional coupling in humans. *Nat. Commun.* **9**, 5382 (2018).
261. Arcaro, M. J., Pinsk, M. A. & Kastner, S. The anatomical and functional organization of the human visual pulvinar. *J. Neurosci.* **35**, 9848–9871 (2015).
262. Baizer, J. S., Desimone, R. & Ungerleider, L. G. Comparison of subcortical connections of inferior temporal and posterior parietal cortex in monkeys. *Vis. Neuro.* **10**, 59–72 (1993).
263. Gattass, R., Galkin, T. W., Desimone, R. & Ungerleider, L. G. Subcortical connections of area V4 in the macaque. *J. Comp. Neurol.* **522**, 1941–1965 (2014).
264. Ungerleider, L. G., Galkin, T. W., Desimone, R. & Gattass, R. Subcortical projections of area V2 in the macaque. *J. Cogn. Neurosci.* **26**, 1220–1233 (2014).
265. Webster, M. J., Bachevalier, J. & Ungerleider, L. G. Connections of inferior temporal areas TEO and TE with parietal and frontal cortex in macaque monkeys. *Cereb. Cortex* **4**, 470–483 (1994).
266. Mercuri, E. et al. Basal ganglia damage and impaired visual function in the newborn infant. *Arch. Dis. Child. Fet. Neonat.* **77**, F111–F114 (1997).
267. Blumberg, M. S. & Adolph, K. E. Protracted development of motor cortex constrains rich interpretations of infant cognition. *Trends Cogn. Sci.* <https://doi.org/10.1016/j.tics.2022.12.014> (2023).

268. Güçlü, U. & van Gerven, M. A. Deep neural networks reveal a gradient in the complexity of neural representations across the ventral stream. *J. Neurosci.* **35**, 10005–10014 (2015).
269. Conwell, C., Prince, J. S., Alvarez, G. A. & Konkle, T. What can 5.17 billion regression fits tell us about artificial models of the human visual system? In *3rd Worksh. on Shared Visual Representations in Human and Machine Intelligence (SVRHM)* (NeurIPS, 2021).
270. Baker, N., Lu, H., Erlikhman, G. & Kellman, P. J. Deep convolutional networks do not classify based on global object shape. *PLoS Comp. Biol.* **14**, e1006613 (2018).

Author contributions

The authors contributed equally to all aspects of the article.

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature Reviews Psychology* thanks Cameron Ellis, Peter Gerhardstein, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© Springer Nature America, Inc. 2023