INTERFACE

royalsocietypublishing.org/journal/rsif

Research

Cite this article: Dehn KI, Maiello G, Hartmann FT, Morgenstern Y, Hawkins SJ, Offner T, Walter J, Hassenklöver T, Manzini I, Fleming RW. 2025 Human shape perception spontaneously discovers the biological origin of novel, but natural, stimuli. *J. R. Soc. Interface* **22**: 20240931.
https://doi.org/10.1098/rsif.2024.0931

Received: 28 December 2024
Accepted: 24 April 2025

Subject Category:
Life Sciences–Engineering interface

Subject Areas:
systems biology, computational biology, biomimetics

Keywords:
three-dimensional shape perception, perceptual organization, visual similarity, generative models, biological cell classification

Author for correspondence:
Guido Maiello
e-mail: g.maiello@soton.ac.uk

†Kira Isabel Dehn and Guido Maiello are joint first authors.

THE ROYAL SOCIETY PUBLISHING

Human shape perception spontaneously discovers the biological origin of novel, but natural, stimuli

Kira Isabel Dehn[1,†], Guido Maiello[3,†], Frieder Tom Hartmann[1], Yaniv Morgenstern[5], Sara Joy Hawkins[4], Thomas Offner[6], Joshua Walter[2], Thomas Hassenklöver[2], Ivan Manzini[2,7] and Roland W. Fleming[1,7]

[1]Department of Psychology, and [2]Department of Animal Physiology and Molecular Biomedicine, Justus Liebig University Giessen, Giessen, Hessen, Germany
[3]School of Psychology, and [4]School of Biological Sciences, University of Southampton, Southampton, England, UK
[5]Erasmus School of Social and Behavioural Sciences, Erasmus University Rotterdam, Rotterdam, Zuid-Holland, The Netherlands
[6]Georg August University of Göttingen, Göttingen, Lower Saxony, Germany
[7]Center for Mind, Brain and Behavior (CMBB), Marburg, Hessen, Germany

GM, 0000-0001-6625-2583; FTH, 0000-0003-4448-5612; YM, 0000-0001-6902-667X; SJH, 0000-0002-7086-7262; TO, 0000-0002-0228-2545; JW, 0009-0005-0812-8396; TH, 0000-0002-9895-1263; IM, 0000-0002-3575-9637; RWF, 0000-0001-5033-5069

Humans excel at categorizing objects by shape. This facility involves identifying shape features that objects have in common with other members of their class and relies—at least in part—on semantic/cognitive constructs. For example, plants sprout branches, fish grow fins, shoes are moulded to our feet. Can humans parse shapes according to the processes that give shapes their key characteristics, even when such processes are hidden? To answer this, we investigated how humans perceive the shape of cells from the olfactory system of *Xenopus laevis* tadpoles. These objects are novel to most humans yet occur in nature and cluster into classes following their underlying biological function. We reconstructed three-dimensional (3D) cell models through 3D microscopy and photogrammetry, then conducted psychophysical experiments. Human participants performed two tasks: they arranged 3D-printed cell models by similarity and rated them along eight visual dimensions. Participants were highly consistent in their arrangements and ratings and spontaneously grouped stimuli to reflect the cell classes, unwittingly revealing the underlying processes shaping these forms. Our findings thus demonstrate that human perceptual organization mechanisms spontaneously parse the biological systematicities of never-before-seen, natural shapes. Integrating such human perceptual strategies into automated systems may enhance morphology-based analysis in biology and medicine.

1. Introduction

What makes a shoe a shoe and not a fish? Whenever we look at a novel object, we effortlessly recognize what group or class it belongs to based on our experience with previous objects. This facility involves identifying key features of the object's three-dimensional (3D) shape that it has in common with other members of its class. These shared shape features are probably due to the shared processes that generated the objects in the first place. For example, plants sprout branches and leaves, fish grow a tails and fins to propel them through water, and shoes are moulded into a shape that matches the structure of human feet. Given that humans are able to parse

shape features according to their physical causes [1–7], this suggests that we may be able to rely on visual 3D shape perception to intuit the shared origins of objects. Yet, these abilities rely—at least in part—on cognitive and semantic processes and constructs [8]. What happens instead when we are presented with completely novel objects of unknown origin? Remarkably, a wealth of previous research has demonstrated that humans can classify novel objects from a few samples [9–13]. For example, Morgenstern *et al*. [11,12] have shown that humans perform such 'one-shot' categorization of novel object classes by relying both on simple heuristics as well as deeper analyses of shape characteristics, which may involve inferring the causal origin of an object's form. However, most of this previous research has used artificially generated stimuli, where the generation processes were constructed ad hoc by the experimenters. This is problematic, as it pre-assumes which shape features make up a class. Yet, natural objects can have extremely complex causal origins, leading to an almost unlimited range of possible diagnostic features for organizing them into classes. In this context, biological entities make particularly fascinating objects of perception, because they have richly structured shapes that result from exquisitely complex morphogenetic processes. It thus remains unknown whether the human visual system is able to recover the underlying biological origin of objects it has never seen before. This question is challenging to answer, however, since humans have extensive experience with both natural and artificial generative processes in our everyday environment. We thus turned to an environment alien to most: the microscopic realm of cells (figure 1).

More specifically, we investigated how humans perceive and parse the 3D shape of cells from the olfactory epithelium of tadpoles of the African clawed frog *Xenopus laevis* [14]. These cells belong to different classes depending on their biological purpose (figure 1A). The olfactory epithelium contains three main cell types: *basal cells*, *supporting cells* and *olfactory sensory neurons* [15]. Basal cells—found close to the basal lamina—are central to regenerative processes [16–18] in the vertebrate olfactory system and may thus undergo mitosis (i.e. cell division). These *dividing cells* may then develop into either olfactory sensory neurons or supporting cells. Immature cells that have just undergone mitosis and have yet to develop clearly distinguishing features are classed as *indeterminate*. To obtain 3D reconstructions of cells belonging to each of these classes, we acquired two-photon microscopy images of individual cells [19–21] and input these into 3D photogrammetry software u-shape3D [22], which created 3D mesh models of each cell (figure 1B). We then 3D printed these 3D cell models at a scaling factor of 1 : 1000 (figure 1C). Finally, we used these 3D-printed objects and 3D renderings as stimuli in human behavioural experiments (figure 1D,E).

In order to classify the cells according to their underlying classes, biologists often visually assess microscopy images of the cells [23]. These visual classifications primarily rely on identifying the stereotypical morphology of each cell class [14,24]. For example, basal cells tend to be small, round and bulbous. Supporting cells and olfactory sensory neurons are instead elongated as they extend from the basal lamina towards the nasal cavity. Furthermore, olfactory sensory neurons have stereotypical axonal and dendritic projections, whereas supporting cells exhibit a conical shape related to their structural purpose [14,15]. Dividing cells are often easily identifiable as they appear to be made of two body segments. Such manual, visually based classifications can be an important step in neurobiological investigations; yet, they are time consuming and prone to human error [25]. It would thus greatly benefit neurobiological science if the process of morphology-based cell classification could be improved and/or automated [26].

The purpose of this study was thus twofold, spanning two very distinct disciplines. Our first aim was to test whether human perceptual organization mechanisms parse 3D shapes according to the processes that give shapes their key characteristics, even when these processes are completely hidden. Additionally, we sought to leverage this human facility to develop and improve morphology-based cell classification systems. To achieve these goals, we designed two behavioural tasks. We designed a multi-arrangement task [27] (figure 1D) in which we asked participants to arrange 3D-printed objects by how similar they appeared in shape. This task was meant to reveal how naive observers would group novel stimuli, and whether such groupings might reflect the biological origin of the stimuli. We then designed a rating task (figure 1E) which probed eight hand-selected visual feature dimensions (figure 1F). This tested whether we could define ad hoc shape features that were related to human similarity judgements and that could also be used to differentiate the cell classes.

2. Results

2.1. Experiment 1: multi-arrangement task

In Experiment 1, we used a multi-arrangement task (figure 1D) to investigate whether naive observers would spontaneously group novel stimuli, and whether such groupings might reflect the biological origin of the stimuli. We asked participants to arrange 3D-printed objects by similarity. All participants were tested with the same set of 30 stimuli (6 per class). Participants performed this multi-arrangement task once at the beginning of the experiment (first session) and once at the end (second session). Between these two sessions, participants performed a rating task (figure 1E, discussed below).

2.1.1. Participants spontaneously grouped novel stimuli, and groupings were related to the biological cell classes

To visualize participant arrangements, we aligned multi-arrangement data using Procrustes analysis. Specifically, we rigidly translated and rotated the stimulus arrangements so they would match as closely as possible across sessions and participants, regardless of their orientation on the workbench. Upon first visual inspection, how participants arranged the stimuli seemed consistent across sessions and broadly corresponded to the underlying ground truth cell classes (figure 2A). To measure how similarly participants grouped the stimuli, we computed representational dissimilarity matrices (RDMs) [27] for each session

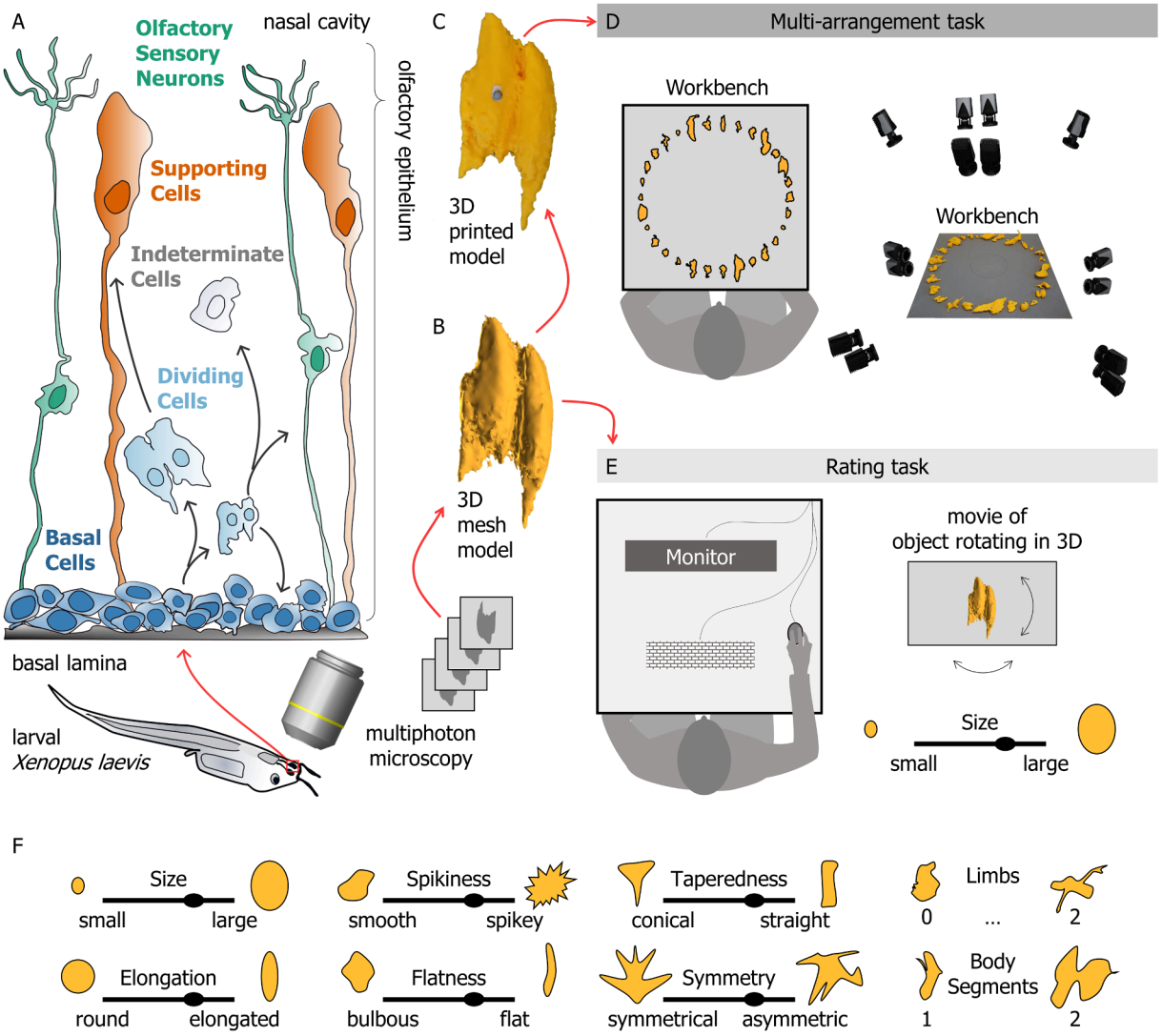

Figure 1. Using novel but natural stimuli to probe human 3D shape perception. (A) Experimental stimuli were 3D reconstructions of cells from the olfactory epithelium of *Xenopus laevis tadpoles*. The cell types reconstructed included basal cells, dividing cells, olfactory sensory neurons, supporting cells and a set of indeterminate cells that could not be decisively assigned to any of the previous types. We reconstructed 30 individual stimuli, 6 stimuli per class. (B) For each of these cell types, 3D mesh model reconstructions were obtained through 3D microscopy. (C) These 3D mesh models were then 3D printed, and a spherical retroreflective marker was glued onto each object. (D) During the multi-arrangement task, participants were seated at a workbench imaged from multiple angles using passive marker tracking cameras. The experimenter placed the 3D-printed stimuli in random order in a circle on the workbench. Participants were asked to rearrange the stimuli by placing them anywhere on the workbench according to their shape similarity. The final object positioning was reconstructed and recorded using a position-tracking system. (E) In the rating task, participants viewed stimulus videos on a computer monitor. Videos were renderings of each object rotating in depth. Participants rated the stimuli along each of eight hand-selected feature dimensions, specified in (F).

and for every participant. RDMs were constructed by computing pairwise Euclidian distances between every possible stimulus combination, normalized to the maximum distance between stimuli. The patterns of correlation between RDMs across sessions and participants represent the agreement between these different arrangements. The RDMs from one example participant (figure 2B) illustrate the high internal consistency across sessions ($r = 0.94$, $p < 0.001$, for the example participant in figure 2B). The similarity of the example participant's RDMs to the group average RDM (figure 2C, left) also illustrates how different participants produced similar arrangements. These robust within- and between-participant agreements held true across our sample (figure 2D, mean within-participant $r = 0.87$, $p < 0.001$, mean between-participant $r = 0.66$, $p < 0.001$).

Individual participants thus relied on specific strategies when arranging the stimuli, and these strategies were similar across participants. This suggests that there may be a common set of features that participants used to judge similarities between 3D shapes. These features might in turn be related to the biological origin of the stimuli. If this were true, participants should have grouped the stimuli lawfully according to their underlying cell classes. To test this, we computed the dissimilarity matrix for perfect ground truth classification and compared it with the average RDMs of all participants from experiment 1. As shown in figure 2C, the group average participant RDM (left) exhibits some—albeit weak—structure along the diagonal which aligns with the cell class RDM (right), giving rise to the weak but significant correlation between participant arrangements and biological cell class (figure 2D, $r = 0.29$, $p < 0.001$). This correlation was clearly not as strong as the between-subject correlation, but this is possibly because the pure cell class RDM was not designed to capture the patterns of similarity across cell classes. For example, it is possible to construct a more complex cell class model by making biologically plausible assumptions that: (i) basal cells are more similar to dividing cells than to other cells (since dividing cells basal cells undergoing mitosis), (ii) it is more

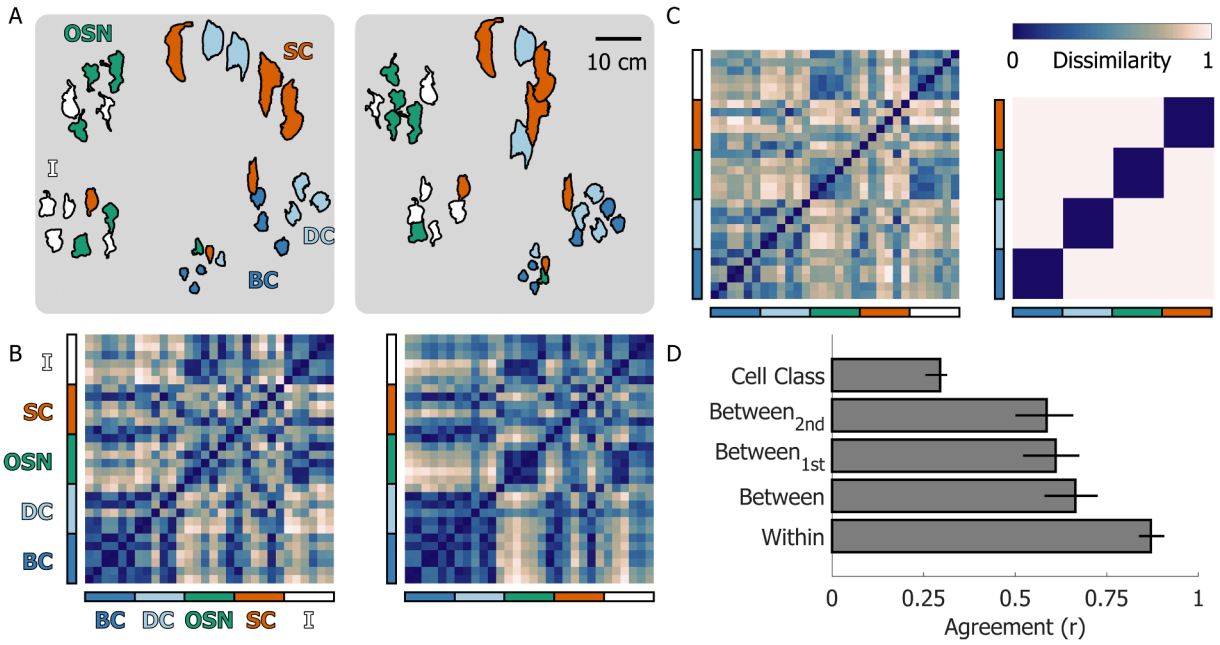

Figure 2. Multi-arrangement results. (A) Multi-arrangement data for one example participant in the first (left) and second sessions (right) of the multi-arrangement task from Experiment 1. Data from the two sessions were aligned using Procrustes analysis. Different colours represent different cell classes: (BC) basal cells; (DC) dividing cells; (OSN) olfactory sensory neurons; (SC) supporting cells, (I) indeterminate cells. (B) Representational dissimilarity matrices (RDMs) for the first (left) and second sessions (right) from the same example participant. Dissimilarity was defined as the normalized Euclidian distance between stimulus pairs. Coloured bars along the *x*- and *y*-axis of the RDMs colour-code the cell classes. (C) Mean RDM averaged across participants and sessions (left) compared with the ground truth classification RDM derived from the cell classes. Note that we excluded indeterminate cells, as these may belong to different biological cell classes. (D) Within and between-participant agreement in the multi-arrangement task and agreement with the ground truth cell classes. Bars represent the mean across participants; error bars represent the 95% bootstrapped confidence intervals of the mean.

likely for dividing cells to be turning into supporting cells rather than olfactory sensory neurons, and (iii) that indeterminate cells are most likely to be immature olfactory sensory neurons. The correlation between the group average participant RDM and this more complex biological cell class model increases to $r = 0.49$ ($p < 0.001$). However, we report the simple cell class model in our primary analyses since any more complex model involves making post hoc decisions about the relative distances between cell classes.

Between the first and second multi-arrangement sessions, participants preformed a 30 min rating task in which they rated the same stimuli along a set of visual feature dimensions. We wondered whether the rating task might have induced participants to change strategies between sessions and arrange the stimuli according to the same set of rated dimensions. If this were the case, we might expect between-participant agreement to be stronger in the second session compared with the first session. Therefore, we calculated between-participant agreement for the first ($r = 0.61$, $p < 0.001$) and second sessions ($r = 0.58$, $p < 0.001$) separately (figure 2D) and observed that between-participant agreements did not differ significantly across sessions ($t(15) = 0.975$, $p = 0.345$). We thus found no evidence that the rating task influenced participant arrangement strategies.

2.2. Experiments 1 and 2: rating task

We designed a computer-based rating task which probed eight hand-selected visual feature dimensions. Participants viewed movie renderings of individual 3D shape stimuli and rated each stimulus along each dimension. Stimulus objects were the same as those used in the multi-arrangement task. We selected the feature dimensions to span a perceptual space that would plausibly cover the 3D shape characteristics of the cell stimuli. In experiment 1, participants performed this rating task between the first and second multi-arrangement sessions. We tested whether participants could reliably rate the stimuli along these feature dimensions, and whether performance at the rating task was as reliable as participant arrangements. We also assessed whether feature ratings were related to how participants arranged the stimuli in the multi-arrangement task. In experiment 2, we verified whether a different set of participants could perform feature ratings through vision alone, i.e. without having ever touched or manipulated the 3D objects. Finally, we used principal component analysis (PCA) to examine the structure of the feature space and tested whether the selected feature dimensions were related to the ground truth biological cell class.

2.2.1. Participants could reliably rate novel stimuli along hand-selected feature dimensions

Figure 1F lists the eight hand-selected feature dimensions we investigated. Participant ratings qualitatively aligned with the feature dimensions. For example, the stimuli increase in size if we organize them according to average participant 'size' ratings (figure 3A). If instead we organize the stimuli according to the average participant 'spikiness' ratings, we can observe how the surface texture visually changes from smooth to spiky (figure 3B). Furthermore, participants agreed with one another on each of the feature dimensions (figure 3C), meaning that different participants gave similar ratings to each of the stimulus

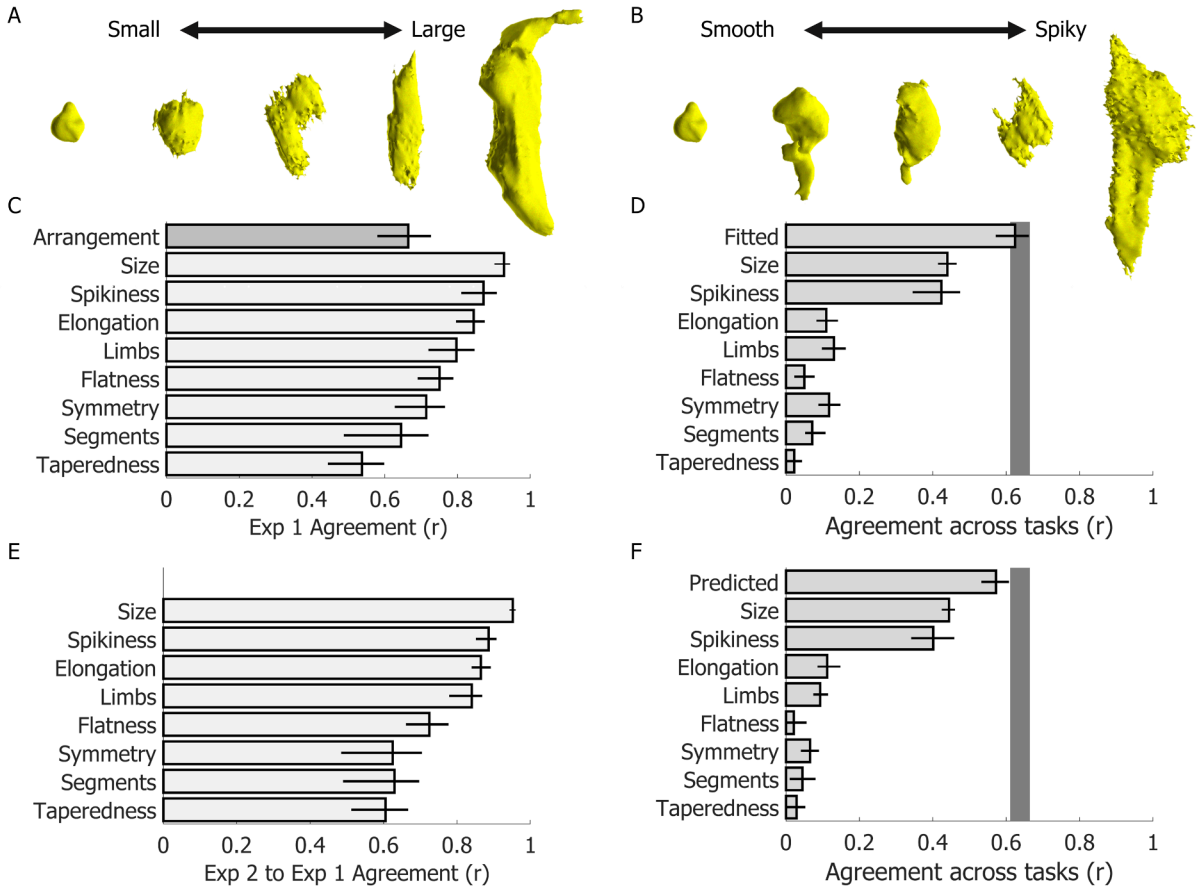

Figure 3. Between-participant agreement within and across experiments and tasks. (A,B) Stimuli ranked 1st, 7th, 15th, 22nd and 30th from the average participant ratings of the 'size' and 'spikiness' dimensions. (C) Between-participant agreement for each of the rating dimensions in experiment 1. As reference, the darker grey bar displays the between-participant agreement for the same participants in the multi-arrangement task. (D) Agreement between rating and multi-arrangement data in experiment 1. (E) Agreement between participant ratings across experiments 1 and 2. (F) Agreement between rating and multi-arrangement data across experiments 1 and 2. Bars represent the mean across participants; error bars are the 95% confidence intervals of the mean. The grey-shaded region in D and F represents the noise ceiling (i.e. the upper and lower estimates of the between-participant agreement in the multi-arrangement task, corresponding to the upper bound of possible agreement across tasks).

objects. Indeed, the between-participant level of agreement for 7/8 feature dimensions was as good as or better than the between-participant agreement in the multi-arrangement task.

2.2.2. Perceptual ratings almost perfectly predicted how participants arranged stimuli in the multi-arrangement task

To assess whether feature ratings were related to how participants arranged the stimuli in the multi-arrangement task, we computed RDMs for each rating dimension and for every participant. We then correlated these RDMs with the average RDM from the multi-arrangement task. We further fit a linear model regressing individual participant rating RDMs onto the average multi-arrangement RDM. Figure 3D shows that 'size' and 'spikiness' are the dimensions that best explain the participant arrangements. It is perhaps interesting to note that these were also the rating dimensions that participants agreed on most strongly. Furthermore, a simple linear combination of the feature ratings could reliably explain participant arrangements. Indeed, the best-fitting model correlated with participant arrangements as strongly as the between-participant agreement in the multi-arrangement task (which here we take as a measure of the noise ceiling). We cross-validated the model using a different set of participants in experiment 2 (see below).

2.2.3. Participants could rate the stimuli using vision alone

We wondered whether haptically manipulating the printed 3D models could have influenced participant behaviour in the rating task. To test this, we thus ran a second experiment. We asked a separate set of participants to perform only the feature rating task, without ever physically interacting with the physical 3D stimuli. We then correlated the rating data of individual participants from experiment 2 to the average ratings across participants from experiment 1. The strong and significant correlations between participants and across experiments (figure 3E) suggest that the observers' behaviour on the rating task was not influenced by the multi-arrangement task. Furthermore, we took the coefficients of the model fitted to the experiment 1 data, which were based on ratings from experiment 1, and applied them to participant ratings from experiment 2. This provided cross-validated predictions of participant arrangements from experiment 1. These predictions could explain

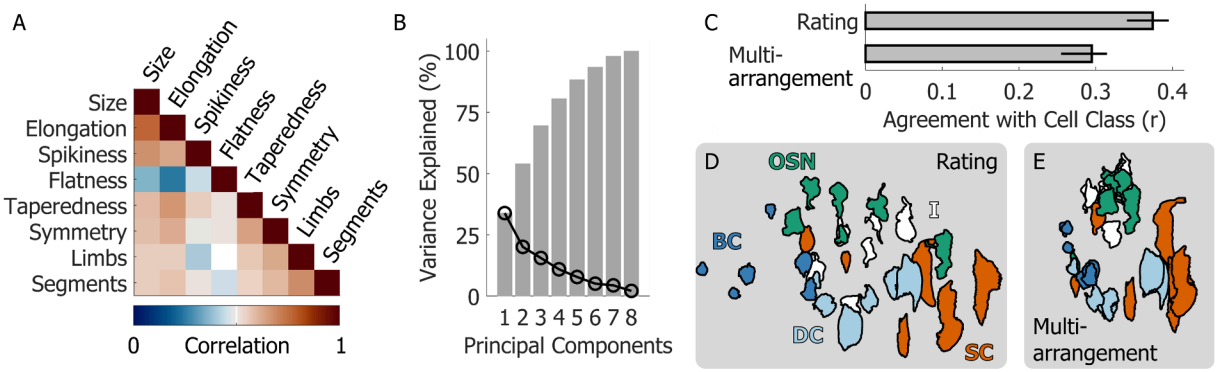

Figure 4. The structure of the rating data. (A) Multi-collinearity matrix showing the correlations between the eight feature dimensions probed by the rating task. (B) Scree plot displaying the variance explained (black curve) and cumulative variance explained (grey bars) the principal components of the rating data. (C) Agreement of the multi-arrangement and rating data with the ground truth cell classes. (D) The average rating data projected in the first two principal components separate the known cell classes. (E) The average multi-arrangement data are more clustered but do not separate cell class as clearly.

participant arrangements from experiment 1 nearly as accurately as the participant ratings from experiment 1 itself (figure 3F). This suggests that visual shape perception was sufficient to explain how participants arranged the stimuli in experiment 1.

2.2.4. Rating data spanned a multi-dimensional space that linearly separated objects into their underlying cell classes

We next tested to what extent the rating dimensions we chose were independent and examined the structure of the perceptual space they spanned. Figure 4A shows overall moderate-to-low correlations between the feature dimensions, whereas figure 4B shows that nearly all eight PCA dimensions are necessary to account for the variance in the rating data. These results suggest that the feature dimensions we selected were at least partly independent. We next fit a multiple linear regression model to attempt to predict cell class from the PCA transformed rating data. The fitted model significantly correlated with cell class, also slightly better than the multi-arrangement data (figure 4C). Indeed, cell class appeared to be visually separable when projected onto the first two PCA dimensions (figure 4D). Furthermore, cell class in PCA space appeared to be more separable but less clustered than in the multi-arrangement data (figure 4E), which explains why the correlation between feature rating data and cell class was still relatively low. However, these patterns suggested that a linear classifier should be able to tease apart cell class from both rating and multi-arrangement data.

2.3. Experiments 1 and 3: recovering cell class from multi-arrangement versus rating data in naive observers and experts

One of the goals of this project was to test whether we could leverage human visual 3D shape perception to classify cell stimuli into their underlying biological classes. We were also interested in whether participants need to be familiar with the underlying cell classes to provide data that could tease these classes apart. For this reason, we ran a final experiment 3 in which we asked a group of five expert biologists—all highly familiar with these specific cell types—to perform the same tasks performed by naive participants in experiment 1. In previous analyses, cell class appeared to be—at least to some extent—linearly separable from the multi-arrangement and rating data. We, therefore, used support vector machine classifiers to attempt to recover cell class from multi-arrangement and rating data in both naive participants and experts.

2.3.1. Support vector machine classifiers recovered cell class across tasks and participant groups

We used a cross-validation procedure to gauge whether support vector machine classifiers could predict cell class from multi-arrangement and rating data provided by naive participants (experiment 1) and expert biologists (experiment 3). For each data type and participant group, we trained classifiers on 20 cell stimuli (five stimuli per each cell class, excluding the indeterminate cells) and tested whether the classifiers correctly predicted the cell class of the four left-out cells. We repeated this procedure for all possible combinations of four left-out cells. Each iteration, we also queried the classifiers to predict the class of the indeterminate cells.

The average human data, projected in two-dimensional (2D) space, appeared qualitatively similar across tasks and participant groups (left subpanels of figure 5A–D). Cells belonging to the same class were near one another, and the spatial relations between cell classes were preserved. For this reason, cross-validated classification accuracy was well above chance across tasks and participant groups (figure 5E). The cross-validated confusion matrices (right subpanels of figure 5A–D) show which cell classes were most easily confused with one another. Basal cells (dark blue) and olfactory sensory neurons (green) were rarely misclassified. Dividing cells (light blue) and supporting cells (orange) were instead more easily confused with each other. This makes sense from a biological viewpoint, as dividing cells may be turning into supporting cells. Furthermore, across all datasets the indeterminate cells (white) were predominantly classified as olfactory sensory neurons. This suggests that in our stimulus set the category of the indeterminate cells consisted predominantly of immature olfactory sensory neurons, which already

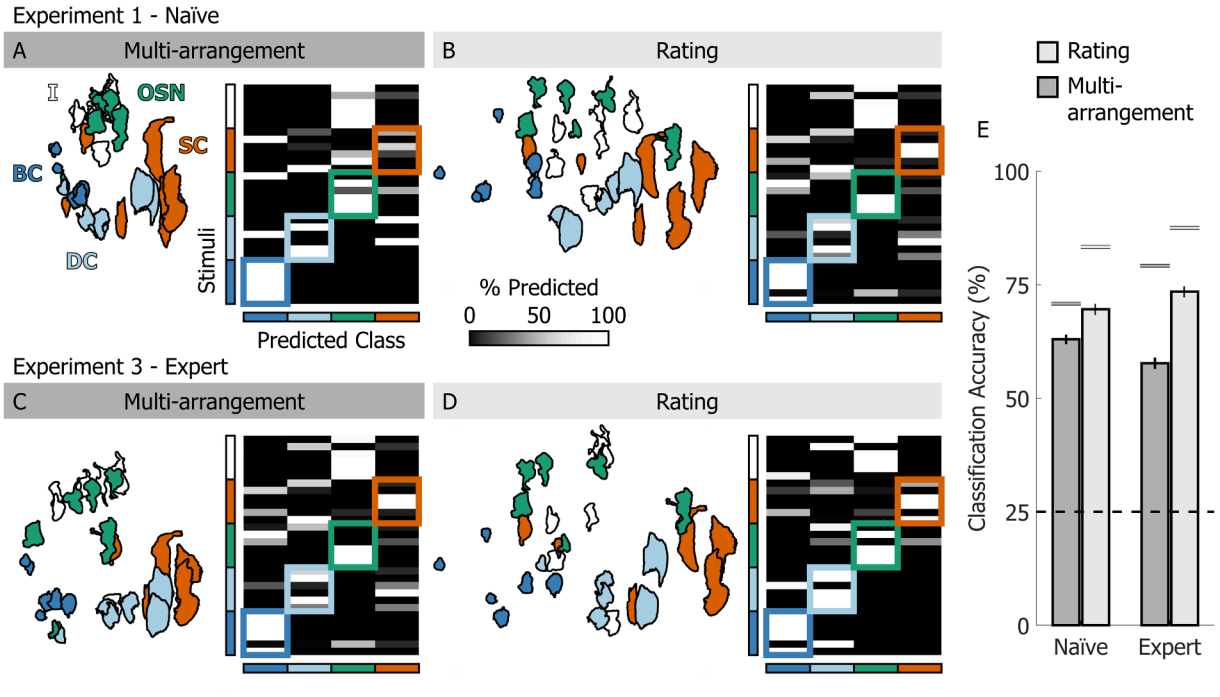

Figure 5. Classification analyses (A) Left: multi-arrangement data from naive participants from experiment 1, averaged across participants. Right: cross-validated confusion matrix for a support vector machine trained to classify cell class from the multi-arrangement data. (B) Left: rating data from naive participants from experiment 1, projected in PCA space and averaged across participants. Right: cross-validated confusion matrix for a support vector machine trained to classify cell class from the PCA-projected rating data. (C,D) As A, B, except for expert biologists from experiment 3. (E) Cross-validated classification accuracy for all support vector machine classifiers trained on different data and participants. Bars are means, error bars are 95% confidence intervals. Dashed-line represents chance performance; grey lines show best-achievable classification accuracy.

displayed shape features similar to those of the mature cells, such as the characteristic axonal and dendritic protrusions. This possibility is corroborated by the fact that expert biologists grouped 5/6 of the indeterminate cells with the olfactory sensory neurons in the multi-arrangement task (figure 5C).

For both experts and naive participants, cross-validated classification accuracy for the classifiers trained on the rating data was higher compared with the classifiers trained on the multi-arrangement data (figure 5E). This was in part due to the higher dimensionality of the rating data. However, this pattern held true when we performed the same analyses using only the first two PCA dimensions of the rating data. Furthermore, it is striking that the classification accuracy obtained from the naive participants' data was nearly as high as that obtained from the experts' data. These results suggest that a feature rating task may be preferable to a spatial arrangement task for classifying cell stimuli into their underlying biological classes. Furthermore, no knowledge of the origin of the stimuli is required to rate visual shape features that can distinguish the different cell classes.

3. Discussion

Humans are able to visually identify and categorize objects based on, among other things, the objects' perceived 3D shape [1–13]. Here, we investigated this ability by using a unique set of stimuli that have a common biological origin but are unknown to most: 3D models of cells from the olfactory system of *Xenopus laevis* tadpoles. We asked human participants to manually sort 3D-printed versions of these stimuli and rate animated renderings along a set of predetermined feature dimensions. Human arrangements and ratings were highly reliable within and across participants and tasks. Participants' arrangements clustered systematically, and indeed biological cell class was linearly decodable from both arrangement and rating data.

3.1. Can humans leverage generative models to infer the causal origin of never-before-seen shapes?

Previous studies have demonstrated that humans are able to make precise categorization decisions based on a small number of examples [9–13]. Such 'one-shot' categorization of novel object classes may involve inferring the causal origin of the objects' form [11,12]. However, previous research has demonstrated that these abilities may rely in part on cognitive and semantic processes and constructs [8]. For example, Spröte *et al*. [4] used both familiar and unfamiliar 2D shapes that appeared to be 'complete' or 'bitten' and asked participants to place dots to identify their perceived symmetry axis. Participant responses for 'complete' and 'bitten' objects were similar, suggesting participants could identify and discount the process that generated the object deformations (i.e. the 'biting'). To study object categorization independently of such higher level cognitive constructs (such as the notion of an object part being 'bitten-off'), some researchers have used discriminative tasks with purely artificially generated stimuli and artificially designed generative processes [28]. Discriminative tasks such as object categorization, however, have a crucial problem when it comes to pre-selected stimuli. The experimenter determines the categories or, in

the case of artificial stimuli, the relevant features and/or generative processes that determine 'ground truth' category labels. Then, participants typically compare artificial objects to one another, thus constraining the investigations to the experimenters' expectations—reflected in the selected features and/or artificial generative processes.

Tiedemann *et al*. [13] in part overcame these limitations using a drawing task in which participants viewed 2D 'exemplar' shapes and then created their own 'variations' of these shapes belonging to the same class. The participant-drawn 'variations' were not mere copies of the 'exemplars' yet were strikingly similar across participants and shared distinctive features. Participants thus behaved as if they were using common generative models to parse the 'exemplars' into categories. However, even in this study the 'exemplars' were artificially generated by the experimenters, and the results were potentially constrained by the participants' drawing skills. These findings could thus still reflect higher level cognitive processes shared between participants and experimenters.

Our experimental design addressed this potential issue by asking participants to arrange—according to the perceived 3D shape similarity—novel real-world objects which fell into unbiased ground truth categories. In line with Tiedemann *et al*. [13], here we found that participants produced robust and consistent arrangements internally as well as between each other. Some bias potentially remained, as the object categories tested were still selected by the experimenters, meaning that the choice of categories might reflect experimenter biases regarding visual distinctiveness or class separability. Nevertheless, participant judgements broadly agreed with the underlying object classes, even though we did not explicitly ask participants to uncover such classes. Given that these classes arose from different natural generative processes, our findings corroborate the notion that humans can recover such generative processes and may use generative models to analyse novel shapes.

3.2. A high-dimensional shape feature space

Our results thus demonstrate that humans can visually tease apart the underlying biological origin of never-before-seen stimuli. What substrate does this ability rely on? Previous literature has proposed that processes of perceptual organization, such as the decomposition of shapes into different parts according to their geometric properties [29–34], are a key aspect of segmenting and representing shapes. Specifically, humans might represent objects as points in a high-dimensional feature space in which objects that are close together belong to the same class, and objects that are far apart belong to different classes [35–38]. To operationalize this hypothesis, Morgenstern *et al*. [11] developed ShapeComp, an image-computable model—based on over 100 2D shape descriptors—capable of accurately predicting human shape similarity judgements. These authors then used ShapeComp to study how humans categorize novel object shapes. When participants were asked to categorize novel objects into different classes and were given multiple exemplars of each class, participants relied on a fixed set of ShapeComp features to perform the task. When participants were given just one exemplar per class instead (i.e. one-shot categorization), they relied on a more sophisticated strategy that involved flexibly reweighting features on a stimulus-by-stimulus basis. This suggested that the feature space is more important than the specific features it is composed of. Our results are consistent with, and substantially extend, these previous findings obtained on 2D shape silhouettes derived from animal shapes. Notably, our arrangement task was neither a one-shot categorization task, nor could participants rely on multiple class exemplars as the stimuli were novel and we did not explicitly tell participants that the stimuli belonged to different classes. Nevertheless, we were able to define a set of ad hoc visual features that were predictive of participant shape similarity arrangements. This set of hand-selected features was also sufficient to tease apart the stimulus classes. Interestingly, the two most predictive features were 'size' and 'spikiness', even though these features differ conceptually—'size' is an atypical feature because shape is typically thought of as a scale invariant, while 'spikiness' is a more localized feature among the ones included (such as elongation, limbs or symmetry). This underscores the flexible and context-sensitive nature of human visual perception, suggesting that observers dynamically reweight shape features based on their discriminative relevance within a particular stimulus set. Therefore, our results demonstrate that our participants' ability to organize novel stimuli according to their hidden underlying causal origin relies on a highly flexible feature space that can be reweighted to reflect the statistical properties of the stimulus set.

However, we recognize that by using predefined, hand-selected dimensions, we introduce the risk of experimenter-induced biases, somewhat conflicting with our aim to minimize such biases. Hebart *et al*. [39] provide an elegant data-driven alternative, in which relevant perceptual dimensions emerge directly from similarity judgements, which are subsequently validated through explicit feature ratings. Due to the relatively small size of our stimulus set, we could not reliably adopt this unbiased approach in the current study. Future research utilizing larger stimulus sets and more extensive sampling could use such fully data-driven methods to rigorously identify and validate the precise perceptual dimensions underlying human shape perception.

3.3. Bringing together feature-based and generative models of shape perception

Spröte *et al*. [4] suggest that the underlying generative processes are a key component to structure and organize perceived shape and a possible strategy to make assumptions of how other members of the same class might look. Fleming and Schmidt [28] instead propose that it is the objects' statistical features that are responsible for their correct identification, and not the generative process itself. Taken together with this previous research, our findings suggest that both could be true. Specifically, in our stimulus set the members of each cell class share a common growth process that probably accounts for the specific shape features shared by members of a same cell class. Mental generative models could thus operate directly in a shape feature space, by first uncovering the statistical feature regularities within a stimulus set, and then reweighting the feature space to reflect these regularities. Then, tasks requiring quick judgements of novel shapes (e.g. one-shot object categorization) could be performed by assessing the plausibility of items and trajectories within this reweighted shape feature space.

Note, however, that while our findings are consistent with generative models playing a role in visual perception, our study was not explicitly designed to differentiate between generative and discriminative perceptual mechanisms. One promising approach to explicitly test these mechanisms could involve leveraging deep learning [40–58]. For example, one could train discriminative deep neural networks [41,43–45,51,52,55,56] (e.g. convolutional neural networks optimized explicitly for classification) as well as generative deep learning models [46–50,53,54,57,58] (e.g. generative adversarial networks) on novel stimulus categories such as the biological shapes used here. After training, examining the representational structure in the latent spaces of these models would clarify which perceptual strategy—generative, discriminative or a combination of both—better aligns with human similarity judgements. Such a computational approach, which would require larger training sets than the stimulus set available for the current study, would nevertheless enable future research to quantitatively dissociate or integrate these perceptual mechanisms in human visual cognition.

3.4. Biological cell classification

Our findings relating to the visual analysis of cell shape may also be of interest to the research field of biological cell classification. To give just a few examples, cell shape guides the formation of healthy cardiac chambers [59], dictates whether individual cells grow or die [60] and can help predict disease progression [61]. The analysis of cell morphology is thus of great interest across biology and medicine, and although researchers have been developing automated methods to conduct cell shape analysis, most previous attempts have focused on 2D images [62,63]. The u-shape3D [22] pipeline used in this work to reconstruct 3D cell shapes from two-photon microscopy images of individual cells is one of the first tools for cellular 3D morphometry, but this field is still in its infancy. Our work suggests that leveraging the computational principles of human visual 3D shape processing may be a way forward. Specifically, simple machine learning approaches could be used to recover cell class from human shape feature ratings, or from stimulus-computable quantities that model these.

Examining the spatial organization of participant arrangements and ratings in figure 5, is even more intriguing. For example, we can note that basal cells are adjacent to dividing cells, and dividing cells are in turn adjacent and overlapping with supporting cells. These patterns appear to reflect the underlying biology (figure 1A), as dividing cells are in fact basal cells undergoing mitosis to replenish other cell types in the olfactory epithelium [14]. Olfactory sensory neurons are—for the most part—clustered away from other cell types. This is probably due to their characteristic axonal and dendritic protrusions, which make them saliently different from the other cell classes [15]. These protrusions develop only after cell division, which is probably why the dividing cells do not overlap with the olfactory sensory neurons. Furthermore, the indeterminate cells seem to be located between supporting cells and olfactory sensory neurons. This reflects the fact that the indeterminate cells are probably immature cells developing into either supporting cells or olfactory sensory neurons. These observations reveal an even deeper insight: we may be able to leverage the principles of human 3D shape processing to map out the life cycle and developmental trajectories of cells within their environment.

3.5. Why are the data not perfectly clustered?

In our study, participants had no prior knowledge of the biological processes that shaped the objects and were asked to arrange the stimuli purely based on their visual similarity. While the natural cell classes were not perfectly mirrored in the similarity arrangements, as we note above, more nuanced patterns emerged in the data (figure 5). This only partial alignment with biological classification is thus not surprising given that many of these cell classes share a common morphogenetic history. For example, dividing cells can develop into supporting cells or olfactory sensory neurons (figure 1A). Such shared developmental pathways may have made it difficult for participants to fully differentiate between distinct cell classes based solely on shape. Moreover, when selecting the stimuli, we focused on cell class but not on the cells' specific degree of maturity. Immature cells, or even cells in the process of dividing, may introduce ambiguity into the perceived shape features, making it more challenging to identify their function. This overlap in statistical features across developing or dividing cells probably blurred the boundaries between classes in the participant arrangements. Additionally, while experts performed slightly better, their data were not significantly more discriminative of cell class than the data from naive participants. This may be due to the fact that experts, although trained to classify cells, were still arranging stimuli based on visual similarity, rather than explicitly grouping them into known biological categories.

3.6. Future directions

Our study opens up many avenues for future research. For example, further investigation is needed to determine the role of tactile manipulation and exploration [64,65] in constructing internal generative models of 3D shape. Additionally, participants in our study might have been able to group stimuli more reliably according to their cell class if they had been explicitly asked to uncover these underlying groups. In particular, expert biologists usually use more than just shape features when classifying cells, by taking into account a cell's location and orientation in the olfactory system to determine its type and function [14]. Incorporating these contextual cues in future experiments might improve classification accuracy for both experts and naive participants. Furthermore, biologists often observe cells through various imaging systems and in different environments, which probably influences their classification process in ways that were not captured by our task. On the other hand, we acknowledge that our selection of visually distinctive stimulus classes facilitated participants' spontaneous discovery of the underlying biological categories. Our findings thus demonstrate human perceptual strategies specifically in scenarios where morphological

regularities are visually apparent. Nonetheless, this study provides a proof of concept that human 3D shape perception can be leveraged to classify and map out different cell classes in biological systems.

Moving forward, several strategies could be used to enhance this approach. First, crowdsourcing [66] could be used to gather a broader set of perceptual data, possibly revealing additional common strategies for grouping unfamiliar biological shapes. Second, biologically inspired computer vision systems [50,53,67–70] could be developed or adapted to replicate human strategies for parsing 3D shapes and distinguishing between subtle biological features. These systems could eventually surpass human capabilities in tasks like cell classification by incorporating contextual data or learning to recognize features that are not immediately obvious to the human eye.

In conclusion, by focusing on objects that exist in nature but with which humans have limited experience—i.e. cellular structures—we were able to demonstrate that human visual perception can successfully categorize novel 3D shapes based on their underlying biological origins. Our approach reveals the flexibility of the human perceptual system in organizing unfamiliar objects according to shape features, even in the absence of contextual information. These findings provide a new understanding of how humans visually process 3D shape and suggests practical applications in fields like biology and medicine, where leveraging human perceptual strategies could aid in automating tasks such as cell classification and morphological analysis.

4. Methods

4.1. Participants

Participants were students and staff from Justus Liebig University Giessen. All participants had normal or corrected to normal vision and participated in the experiments for course credit or financial compensation (at a rate of 8 Euro h^{-1}). We recruited 16 naive participants for experiment 1 (12 female, mean (range) age 23 (18–27) years), 15 naive participants for experiment 2 (9 female, age 26 (20-–35) years), and 5 'expert' participants for experiment 3 (4 female, age 28 (25–32) years). Expert participants were current or past members of the Animal Physiology and Molecular Biomedicine research group at the Institute of Animal Physiology of Justus Liebig University Giessen. All participants provided written informed consent. All procedures were approved by the local ethics committee of Justus Liebig University Giessen and adhered to the tenants of the sixth revision of the Declaration of Helsinki (2008).

4.2. Stimuli

Experiments included two types of behavioural tasks, a multi-arrangement task and a rating task. The stimuli used in all experiments and tasks were 3D reconstructions of real cells from the olfactory system of *Xenopus laevis* tadpoles. These reconstructions were generated by the team of expert biologists led by Professor Ivan Manzini at the Department of Animal Physiology and Molecular Biomedicine at Justus Liebig University Giessen. All animal procedures were performed following the guidelines of laboratory animal research of the Institutional Care and Use Committee of the Justus Liebig University of Giessen (649_M; GI 15/7 Nr. G 2/2019).

The cell reconstructions could belong to different cell types found in the olfactory epithelium of *Xenopus* tadpoles. We selected 30 cell stimuli for our experiments, such that they would be equally distributed into four known cell classes plus one indeterminate set. The four known cell classes were basal cells, dividing cells, olfactory sensory neurons and supporting cells. Cells were classified as belonging to these classes through imaging investigations performed by the expert biologists. Cells belonging to the indeterminate set were those that could not be reliably classified by the experts into any of the known groups. These indeterminate cells might have been, for example, immature versions of olfactory sensory neurons or supporting cells.

Cell reconstructions were obtained through 3D microscopy. Specifically, 3D-microscopy images were processed using u-shape3D [22], a software package to reconstruct and analyse 3D cell morphologies. The reconstructions were created as 3D triangulated mesh models. The mesh models were further processed for rendering and 3D printing using Blender 2.91 software (The Blender Foundation, Amsterdam, The Netherlands). We removed non-manifold edges, edges with no length, faces or face corners with no area and small disconnected elements. Larger disconnected elements were instead manually reconnected to the cell's main body.

For the multi-arrangement task, the stimuli were 3D-printed versions of the cell reconstructions. The reconstructions were 3D printed out of a light, yellow plastic at a scaling factor of 1 : 1000 (1 µm = 1 cm). For the rating task, we created video renderings of each mesh model. Each video sequence displayed one mesh model rotating back and forth along frontal and vertical axes at approximately 120° s^{-1}. Video sequences were all 30 s long and were created with MATLAB R2020b software. Mesh models in print-ready '.stl' format and video renderings for all stimuli employed in our experiments will be made available from our open-access data repository (doi: 10.5281/zenodo.14046922).

4.3. Apparatus

Experiments were programmed in Python 3.7.0 and MATLAB R2020b. In the multi-arrangement task, participants were asked to arrange the full set of 3D-printed stimulus objects on a workbench (60 × 60 cm). The position of the stimuli arranged on the workbench was captured with high-precision 3D tracking rig [71] using hardware and software produced by motion capture

company Qualisys (Qualisys AB, Sweden). The stimuli were imaged from multiple angles by eight tracking cameras (Qualisys Miqus) and six video cameras (Qualisys Miqus Video) arranged on a square frame surrounding the workspace (figure 1D). To capture the position of the stimuli, we glued one passive reflective marker on top of each 3D-printed cell (figure 1C).

In the rating task, participants viewed stimulus videos on an Asus VG248QE monitor (24″, resolution = 1920 × 1080 pixel) at 60 Hz, positioned at a distance of 40 cm from the observers (figure 1E). The monitor subtended 67 × 41° of visual angle, and the object size ranged from 1 to 10° of visual angle.

4.4. Procedure

In experiment 1, naive participants performed a first session of a multi-arrangement task, followed by a rating task, followed by a second session of the multi-arrangement task. In experiment 2, naive participants performed only a rating task, which was identical to the rating task performed by participants in experiment 1. In experiments 1 and 2, participants were kept naive to the objects' origin and were never told that the objects belonged to distinct categories. In experiment 3, a set of expert participants preformed the same set of tasks and sessions as naive participants in experiment 1.

4.4.1. Multi-arrangement task

Each session, the experimenter initially positioned the 3D-printed stimuli in random order in a circle on the workbench. Stimulus order was pseudo-randomly generated by the experimental script. Participants were asked to freely arrange the stimuli anywhere on the table surface according to their perceived similarity. This meant that similar objects should be placed spatially close together whilst dissimilar objects should be placed far apart. No other constraints on spatial configurations were imposed, and participants had unlimited time to perform these arrangements. Once a participant was satisfied with the stimulus arrangement, the experimenter recorded the final stimulus positioning measured by the Qualisys tracking system using an ad hoc script written in Python. The identity of the stimuli was manually annotated by the experimenter off-line.

4.4.2. Rating task

Participants performed a computer-based rating task in which they were shown video sequences of rotating 3D stimuli. Each trial, a participant had to rate one of the stimuli along one of eight feature dimensions. Participants rated each stimulus only once for each dimension, for a total of (30 × 8) 240 trials. Participants had unlimited time to perform these ratings. The order of the stimuli was randomized across participants, but the rating dimensions were queried in the same succession for each stimulus object. Participants made their assessment either by using the mouse to adjust a slider or by typing in a number on the keyboard. To ensure participants understood the rating dimensions, prior to the start of the experiment they were shown exemplary cartoon drawings for each dimension (figure 1F).

4.5. Analyses

We used MATLAB R2020b software to analyse the data.

4.5.1. Analysis of the multi-arrangement data

First, we used Procrustes analysis to align and visualize the arrangements of the 30 cell models for each participant in each condition. Specifically, we translated and rotated the arrangement data (without rescaling but allowing reflection) so they were best-aligned across participants and sessions. Then, representational dissimilarity matrices for each participant and each session were constructed from the arrangement data by computing the pairwise Euclidean distances between the 30 objects. All analyses were performed using only the upper triangular portions of the dissimilarity matrices, excluding the diagonal.

4.5.1.1. Multi-arrangement task, within-participant agreement

For each participant, we computed two dissimilarity matrices, one for the first session $\mathrm{RDM}_{p,s1}$ and one for the second session $\mathrm{RDM}_{p,s2}$. We computed the participant's average dissimilarity matrix across the first and second session as

$$\mathrm{RDM}_{p,Av} = \frac{\mathrm{RDM}_{p,s1} + \mathrm{RDM}_{p,s2}}{2}.$$

Then, we calculated the Pearson correlation between this average dissimilarity matrix and the participant's first- and second-session dissimilarity matrices,

$$W_{p,s1} = \mathrm{corr}(\mathrm{RDM}_{p,s1}, \mathrm{RDM}_{p,Av}), \quad W_{p,s2} = \mathrm{corr}(\mathrm{RDM}_{p,s2}, \mathrm{RDM}_{p,Av}).$$

The average of these two correlations was taken as the within-participant agreement,

$$W_p = \frac{W_{p,s1} + W_{p,s2}}{2}.$$

4.5.1.2. Multi-arrangement task, between-participant agreement

We first computed the average dissimilarity matrix across all N participants as

$$\mathrm{RDM}_{Av} = \frac{1}{N}\sum_{p=1}^{N} \mathrm{RDM}_{p,Av}.$$

Then, the between-participant agreement we calculated as the Pearson correlation between participants' average dissimilarity matrices $\mathrm{RDM}_{p,Av}$ and the average dissimilarity matrix of all participants,

$$B_p = \mathrm{corr}(\mathrm{RDM}_{p,Av}, \mathrm{RDM}_{Av}).$$

4.5.1.3. Multi-arrangement task, within-session, between-participant agreement

To compute single-session between-participant agreement, we first calculated the single-session average dissimilarity matrices across all N participants as

$$\mathrm{RDM}_{Av,s1} = \frac{1}{N}\sum_{p=1}^{N} \mathrm{RDM}_{p,s1}, \quad \mathrm{RDM}_{Av,s2} = \frac{1}{N}\sum_{p=1}^{N} \mathrm{RDM}_{p,s2}.$$

Then, single-session between-participant agreements were computed as

$$B_{p,s1} = \mathrm{corr}(\mathrm{RDM}_{p,s1}, \mathrm{RDM}_{Av,s1}), \quad B_{p,s2} = \mathrm{corr}(\mathrm{RDM}_{p,s2}, \mathrm{RDM}_{Av,s2}).$$

4.5.1.4. Multi-arrangement task, agreement with cell class

The dissimilarity matrix for cell class RDM_{CC} was constructed as a binary matrix, where cells belonging to the same class had 0 dissimilarity, and cells belonging to different classes had a dissimilarity of 1. The agreement between participant arrangements and cell class was taken as the correlation between the cell class dissimilarity matrix and the average dissimilarity matrix across participants,

$$\mathrm{CC} = \mathrm{corr}(\mathrm{RDM}_{CC}, \mathrm{RDM}_{Av}).$$

4.5.2. Analysis of the rating task data

4.5.2.1. Rating task, between-participant agreement in experiment 1

For each of the eight feature dimensions d, we first computed the average ratings across participants for all 30 stimuli. Between-participant agreement was taken as the correlation between individual participant ratings $Rt_{p,d}$ and the average rating across participants $Rt_{Av,d}$,

$$B_{p,d} = \mathrm{corr}(Rt_{p,d}, Rt_{Av,d}).$$

4.5.2.2. Between-task agreement within experiment 1

We computed dissimilarity matrices for each rating dimension and for every participant $\mathrm{RDM}^{Rt}_{p,d}$. To compute between-task agreement, we then correlated these dissimilarity matrices with the average dissimilarity matrix from the multi-arrangement task,

$$B^{MA-Rt}_{p,d} = \mathrm{corr}(\mathrm{RDM}^{Rt}_{p,d}, \mathrm{RDM}_{Av}).$$

Next, we fit a linear model regressing individual participant rating dissimilarity matrices onto the average multi-arrangement dissimilarity matrix. The model was defined as

$$\mathrm{RDM}_{Av} = \sum_{d=1}^{8} \beta_d \mathrm{RDM}^{Rt}_{p,d} + \delta,$$

where β_d are linear coefficients and δ is the intercept.

4.5.2.3. Rating task agreement across experiments 1 and 2

Between-participant agreement across experiments 1 and 2 was taken as the correlation between individual participant ratings in experiment 2, $Rt^{E2}_{p,d}$ and the average rating across participants in Experiment 1,

$$B_{p,d}^{E1-E2} = \mathrm{corr}\left(Rt_{p,d}^{E2}, Rt_{Av,d}^{E1}\right).$$

4.5.2.4. Between-task agreement across experiments 1 and 2

We computed dissimilarity matrices for each rating dimension and for every participant from experiment 2 $\mathrm{RDM}_{p,d}^{Rt,E2}$. To compute between-task agreement across experiments 1 and 2, we correlated these dissimilarity matrices with the average dissimilarity matrix from the multi-arrangement task in experiment 1,

$$B_{p,d}^{MA-Rt,E1-E2} = \mathrm{corr}\left(\mathrm{RDM}_{p,d}^{Rt,E2}, \mathrm{RDM}_{Av}^{MA,E1}\right).$$

Finally, we used the coefficients fitted above to predict, from the rating data in experiment 2, the multi-arrangement dissimilarity matrix from experiment 1.

4.5.2.5. Principal component analysis of the rating task data

We assessed the multi-collinearity of the rating dimensions by computing the pairwise correlations between all the eight rating features. We then performed PCA to recover eight new dimensions or principal components, each one orthogonal (and thus uncorrelated) to the others, composed of a linear combination of the original features. To prevent biased results, we standardized the scaling of the two dimensions which were assessed by assigning a discrete number (*main body parts* and *number of limbs*) to the same [0 1] range in which the other six dimensions were expressed.

4.5.2.6. Agreement between rating data and cell class

To investigate whether rating data aligned with the underlying biological cell class, we computed per-participant RDMs for each PCA dimension, and then fit a linear model regressing individual participant dissimilarity matrices onto the cell class dissimilarity matrix.

4.5.3. Classification analyses

Our final analyses tested whether it was possible to recover the underlying biological cell classes of our stimuli from human perceptual arrangements and judgements of 3D shape. We trained support vector machine classifiers to classify the participant multi-arrangement and rating data into the correct underlying cell class. We performed these analyses on the data from experiments 1 (naives) and 3 (experts). Classification accuracy was computed using the cross-validation procedure described in the §2.

Ethics. All procedures were approved by the local ethics committee of Justus Liebig University Giessen and adhered to the tenets of the sixth revision of the Declaration of Helsinki (2008).

Data accessibility. Data and analysis scripts are publicly available from the Zenodo database [72].

Declaration of AI use. We have not used AI-assisted technologies in creating this article.

Authors' contributions. K.I.D.: conceptualization, data curation, formal analysis, investigation, methodology, project administration, software, validation, visualization, writing—original draft, writing—review and editing; G.M.: conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, software, supervision, validation, visualization, writing—original draft, writing—review and editing; F.T.H.: conceptualization, methodology, software, writing—review and editing; Y.M.: conceptualization, methodology, writing—review and editing; S.J.H.: conceptualization, methodology, resources, writing—review and editing; T.O.: conceptualization, methodology, resources, writing—review and editing; J.W.: conceptualization, methodology, resources, writing—review and editing; T.H.: conceptualization, methodology, resources, writing—review and editing; I.M.: conceptualization, funding acquisition, methodology, resources, writing—review and editing; R.W.F.: conceptualization, funding acquisition, methodology, project administration, resources, supervision, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. This research was supported by the DFG (SFB-TRR-135-TP-C1: 'Cardinal Mechanisms of Perception' and project PA 3723/1-1). Research Cluster 'The Adaptive Mind' funded by the Excellence Program of the Hessian Ministry of Higher Education, Science, Research and Art. An ERC Advanced Grant (ERC-2022-AdG-101098225 'STUFF') to R.W.F.

References

1. Phillips F, Fleming RW. 2020 The *Veiled Virgin* illustrates visual segmentation of shape by cause. *Proc. Natl Acad. Sci. USA* **117**, 11735–11743. (doi:10.1073/pnas.1917565117)
2. Spröte P, Fleming RW. 2016 Bent out of shape: the visual inference of non-rigid shape transformations applied to objects. *Vis. Res.* **126**, 330–346. (doi:10.1016/j.visres.2015.08.009)
3. Schmidt F, Spröte P, Fleming RW. 2016 Perception of shape and space across rigid transformations. *Vis. Res.* **126**, 318–329. (doi:10.1016/j.visres.2015.04.011)
4. Spröte P, Schmidt F, Fleming RW. 2016 Visual perception of shape altered by inferred causal history. *Sci. Rep.* **6**, 36245. (doi:10.1038/srep36245)
5. Schmidt F, Fleming RW. 2016 Visual perception of complex shape-transforming processes. *Cogn. Psychol.* **90**, 48–70. (doi:10.1016/j.cogpsych.2016.08.002)
6. Schmidt F, Fleming RW. 2018 Identifying shape transformations from photographs of real objects. *PLoS One* **13**, e0202115. (doi:10.1371/journal.pone.0202115)

7. Schmidt F, Phillips F, Fleming RW. 2019 Visual perception of shape-transforming processes: 'shape scission'. *Cognition* **189**, 167–180. (doi:10.1016/j.cognition.2019.04.006)
8. Schmidt F, Kleis J, Morgenstern Y, Fleming RW. 2020 The role of semantics in the perceptual organization of shape. *Sci. Rep*. **10**, 22141. (doi:10.1038/s41598-020-79072-w)
9. Feldman J. 1992 Constructing perceptual categories. In *Proc. 1992 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp. 244–250. Champaign, IL: IEEE Comput. Soc. Press. (doi:10.1109/CVPR.1992.223268)
10. Feldman J. 1997 The structure of perceptual categories. *J. Math. Psychol*. **41**, 145–170. (doi:10.1006/jmps.1997.1154)
11. Morgenstern Y, Schmidt F, Fleming RW. 2019 One-shot categorization of novel object classes in humans. *Vis. Res*. **165**, 98–108. (doi:10.1016/j.visres.2019.09.005)
12. Morgenstern Y, Schmidt F, Fleming RW. 2020 A dataset for evaluating one-shot categorization of novel object classes. *Data Brief* **29**, 105302. (doi:10.1016/j.dib.2020.105302)
13. Tiedemann H, Morgenstern Y, Schmidt F, Fleming RW. 2022 One-shot generalization in humans revealed through a drawing task. *eLife* **11**, e75485. (doi:10.7554/elife.75485)
14. Manzini I. 2015 From neurogenesis to neuronal regeneration: the amphibian olfactory system as a model to visualize neuronal development *in vivo*. *Neural Regen. Res*. **10**, 872. (doi:10.4103/1673-5374.158334)
15. Manzini I, Schild D, Di Natale C. 2022 Principles of odor coding in vertebrates and artificial chemosensory systems. *Physiol. Rev*. **102**, 61–154. (doi:10.1152/physrev.00036.2020)
16. Hawkins SJ, Weiss L, Offner T, Dittrich K, Hassenklöver T, Manzini I. 2017 Functional reintegration of sensory neurons and transitional dendritic reduction of mitral/tufted cells during injury-induced recovery of the larval *Xenopus* olfactory circuit. *Front. Cell. Neurosci*. **11**, 380. (doi:10.3389/fncel.2017.00380)
17. Hawkins SJ, Gärtner Y, Offner T, Weiss L, Maiello G, Hassenklöver T, Manzini I. 2024 The olfactory network of larval *Xenopus laevis* regenerates accurately after olfactory nerve transection. *Eur. J. Neurosci*. **60**, 3719–3741. (doi:10.1111/ejn.16375)
18. Weiss L, Segoviano Arias P, Offner T, Hawkins SJ, Hassenklöver T, Manzini I. 2021 Distinct interhemispheric connectivity at the level of the olfactory bulb emerges during *Xenopus laevis* metamorphosis. *Cell Tissue Res*. **386**, 491–511. (doi:10.1007/s00441-021-03527-3)
19. Offner T, Daume D, Weiss L, Hassenklöver T, Manzini I. 2020 Whole-brain calcium imaging in larval *Xenopus*. *Cold Spring Harb. Protoc.* **2020**, pdb.prot106815. (doi:10.1101/pdb.prot106815)
20. Offner T, Weiss L, Daume D, Berk A, Inderthal TJ, Manzini I, Hassenklöver T. 2023 Functional odor map heterogeneity is based on multifaceted glomerular connectivity in larval *Xenopus* olfactory bulb. *iScience* **26**, 107518. (doi:10.1016/j.isci.2023.107518)
21. Weiss L, Offner T, Hassenklöver T, Manzini I. 2018 Dye electroporation and imaging of calcium signaling in *Xenopus* nervous system. In *Xenopus* (ed. K Vleminckx), pp. 217–231, vol. **1865**. New York, NY: Springer. (doi:10.1007/978-1-4939-8784-9_15)
22. Driscoll MK, Welf ES, Jamieson AR, Dean KM, Isogai T, Fiolka R, Danuser G. 2019 Robust and automated detection of subcellular morphological motifs in 3D microscopy images. *Nat. Methods* **16**, 1037–1044. (doi:10.1038/s41592-019-0539-z)
23. Hansen A, Reiss JO, Gentry CL, Burd GD. 1998 Ultrastructure of the olfactory organ in the clawed frog, *Xenopus laevis*, during larval development and metamorphosis. *J. Comp. Neurol*. **398**, 273–288. (doi:10.1002/(sici)1096-9861(19980824)398:23.0.co;2-y)
24. Schwob JE. 2002 Neural regeneration and the peripheral olfactory system. *Anat. Rec*. **269**, 33–49. (doi:10.1002/ar.10047)
25. Brunyé TT, Drew T, Weaver DL, Elmore JG. 2019 A review of eye tracking for understanding and improving diagnostic interpretation. *Cogn. Res*. **4**, 7. (doi:10.1186/s41235-019-0159-2)
26. Gu Y, Zhang AC, Han Y, Li J, Chen C, Lo YH. 2019 Machine learning based real-time image-guided cell sorting and classification. *Cytometry A* **95**, 499–509. (doi:10.1002/cyto.a.23764)
27. Kriegeskorte N, Mur M. 2012 Inverse MDS: inferring dissimilarity structure from multiple item arrangements. *Front. Psychol*. **3**, 245. (doi:10.3389/fpsyg.2012.00245)
28. Fleming RW, Schmidt F. 2019 Getting 'fumpered': classifying objects by what has been done to them. *J. Vis.* **19**, 15. (doi:10.1167/19.4.15)
29. Richards W, Hoffman DD. 1985 Codon constraints on closed 2D shapes. *Comput. Vis. Graph. Image Process*. **31**, 265–281. (doi:10.1016/0734-189x(85)90031-3)
30. Hoffman DD, Singh M. 1997 Salience of visual parts. *Cognition* **63**, 29–78. (doi:10.1016/s0010-0277(96)00791-3)
31. Siddiqi K, Kimia BB. 1995 Parts of visual form: computational aspects. *IEEE Trans. Pattern Anal. Mach. Intell*. **17**, 239–251.
32. Biederman I. 1987 Recognition-by-components: a theory of human image understanding. *Psychol. Rev*. **94**, 115–147. (doi:10.1037//0033-295x.94.2.115)
33. Singh M, Seyranian GD, Hoffman DD. 1999 Parsing silhouettes: the short-cut rule. *Percept. Psychophys*. **61**, 636–660. (doi:10.3758/bf03205536)
34. Marr D, Nishihara HK. 1978 Representation and recognition of the spatial organization of three-dimensional shapes. *Proc. R. Soc. Lond. B* **200**, 269–294. (doi:10.1098/rspb.1978.0020)
35. Shepard RN. 1987 Toward a universal law of generalization for psychological science. *Science* **237**, 1317–1323. (doi:10.1126/science.3629243)
36. Tenenbaum JB, Griffiths TL. 2001 Generalization, similarity, and Bayesian inference. *Behav. Brain Sci*. **24**, 629–640. (doi:10.1017/s0140525x01000061)
37. Hegdé J, Bart E, Kersten D. 2008 Fragment-based learning of visual object categories. *Curr. Biol*. **18**, 597–601. (doi:10.1016/j.cub.2008.03.058)
38. Kromrey S, Maestri M, Hauffen K, Bart E, Hegdé J. 2010 Fragment-based learning of visual object categories in non-human primates. *PLoS One* **5**, e15444. (doi:10.1371/journal.pone.0015444)
39. Hebart MN, Zheng CY, Pereira F, Baker CI. 2020 Revealing the multidimensional mental representations of natural objects underlying human similarity judgements. *Nat. Hum. Behav*. **4**, 1173–1185. (doi:10.1038/s41562-020-00951-3)
40. Jozwik KM, Kriegeskorte N, Storrs KR, Mur M. 2017 Deep convolutional neural networks outperform feature-based but not categorical models in explaining object similarity judgments. *Front. Psychol*. **8**, 1726. (doi:10.3389/fpsyg.2017.01726)
41. Kubilius J, Bracci S, Op de Beeck HP. 2016 Deep neural networks as a computational model for human shape sensitivity. *PLoS Comput. Biol*. **12**, e1004896. (doi:10.1371/journal.pcbi.1004896)
42. O'Connell TP, Bonnen T, Friedman Y, Tewari A, Tenenbaum JB, Sitzmann V, Kanwisher N. 2023 Approaching human 3D shape perception with neurally mappable models. arXiv. (doi:10.48550/arXiv.2308.11300)
43. Baker N, Elder JH. 2022 Deep learning models fail to capture the configural nature of human shape perception. *iScience* **25**, 104913. (doi:10.1016/j.isci.2022.104913)
44. Kalfas I, Vinken K, Vogels R. 2018 Representations of regular and irregular shapes by deep convolutional neural networks, monkey inferotemporal neurons and human judgments. *PLoS Comput. Biol*. **14**, e1006557. (doi:10.1371/journal.pcbi.1006557)
45. Heinke D, Wachman P, van Zoest W, Leek EC. 2021 A failure to learn object shape geometry: implications for convolutional neural networks as plausible models of biological vision. *Vis. Res*. **189**, 81–92. (doi:10.1016/j.visres.2021.09.004)
46. Salakhutdinov R. 2015 Learning deep generative models. *Annu. Rev. Stat. Appl.* **2**, 361–385. (doi:10.1146/annurev-statistics-010814-020120)
47. Radford A, Metz L, Chintala S. 2015 Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv. (doi:10.48550/arXiv.1511.06434)
48. Testolin A, Zorzi M. 2016 Probabilistic models and generative neural networks: towards a unified framework for modeling normal and impaired neurocognitive functions. *Front. Comput. Neurosci*. **10**, 73. (doi:10.3389/fncom.2016.00073)

49. Storrs KR, Anderson BL, Fleming RW. 2021 Unsupervised learning predicts human perception and misperception of gloss. *Nat. Hum. Behav*. **5**, 1402–1417. (doi:10.1038/s41562-021-01097-6)
50. Maiello G. 2023 Unsupervised learning in biological brains. *Nat. Rev. Psychol*. **2**, 201–201. (doi:10.1038/s44159-023-00166-z)
51. Baker N, Lu H, Erlikhman G, Kellman PJ. 2018 Deep convolutional networks do not classify based on global object shape. *PLoS Comput. Biol*. **14**, e1006613. (doi:10.1371/journal.pcbi.1006613)
52. Erdogan G, Jacobs RA. 2017 Visual shape perception as Bayesian inference of 3D object-centered shape representations. *Psychol. Rev*. **124**, 740–761. (doi:10.1037/rev0000086)
53. Morgenstern Y, Hartmann F, Schmidt F, Tiedemann H, Prokott E, Maiello G, Fleming RW. 2021 An image-computable model of human visual shape similarity. *PLoS Comput. Biol*. **17**, e1008981. (doi:10.1371/journal.pcbi.1008981)
54. Morgenstern Y, Storrs KR, Schmidt F, Hartmann F, Tiedemann H, Wagemans J, Fleming RW. 2024 High-level aftereffects reveal the role of statistical features in visual shape encoding. *Curr. Biol.* **34**, 1098–1106. (doi:10.1016/j.cub.2023.12.039)
55. Geirhos R, Rubisch P, Michaelis C, Bethge M, Wichmann FA, Brendel W. 2018 ImageNet-trained CNNs are biased towards texture; increasing shape bias improves accuracy and robustness. arXiv. (doi:10.48550/arXiv.1811.12231)
56. Jarvers C, Neumann H. 2024 Teaching deep networks to see shape: lessons from a simplified visual world. *PLoS Comput. Biol*. **20**, e1012019. (doi:10.1371/journal.pcbi.1012019)
57. Wu J, Zhang C, Xue T, Freeman WT, Tenenbaum JB. 2016 Learning a probabilistic latent space of object shapes via 3D generative-adversarial modeling. In *30th Conf. on Neural Information Processing Systems*, Barcelona, Spain, pp. 82–90. Red Hook, NY: Curran Associates Inc. (doi:10.48550/arXiv.1610.07584)
58. Yildirim I, Siegel MH, Soltani AA, Chaudhuri SR, Tenenbaum JB. 2023 Perception of 3D shape integrates intuitive physics and analysis-by-synthesis. *Nat. Hum. Behav*. **8**, 320–335. (doi:10.1038/s41562-023-01759-7)
59. Auman HJ, Coleman H, Riley HE, Olale F, Tsai HJ, Yelon D. 2007 Functional modulation of cardiac form through regionally confined cell shape changes. *PLoS Biol*. **5**, e53. (doi:10.1371/journal.pbio.0050053)
60. Chen CS, Mrksich M, Huang S, Whitesides GM, Ingber DE. 1997 Geometric control of cell life and death. *Science* **276**, 1425–1428. (doi:10.1126/science.276.5317.1425)
61. Pfau M *et al*. 2019 Prognostic value of shape-descriptive factors for the progression of geographic atrophy secondary to age-related macular degeneration. *Retina* **39**, 1527–1540. (doi:10.1097/iae.0000000000002206)
62. Nowak DA, Berner J, Herrnberger B, Kammer T, Grön G, Schönfeldt-Lecuona C. 2009 Continuous theta-burst stimulation over the dorsal premotor cortex interferes with associative learning during object lifting. *Cortex* **45**, 473–482. (doi:10.1016/j.cortex.2007.11.010)
63. Driscoll MK, Danuser G. 2015 Quantifying modes of 3D cell migration. *Trends Cell Biol*. **25**, 749–759. (doi:10.1016/j.tcb.2015.09.010)
64. Wijntjes MWA, Volcic R, Pont SC, Koenderink JJ, Kappers AML. 2009 Haptic perception disambiguates visual perception of 3D shape. *Exp. Brain Res*. **193**, 639–644. (doi:10.1007/s00221-009-1713-9)
65. Norman JF, Norman HF, Clayton AM, Lianekhammy J, Zielke G. 2004 The visual and haptic perception of natural object shape. *Percept. Psychophys*. **66**, 342–351. (doi:10.3758/bf03194883)
66. Schmidt F, Hebart MN, Schmid AC, Fleming RW. 2025 Core dimensions of human material perception. *Proc. Natl Acad. Sci. USA* **122**, e2417202122. (doi:10.1073/pnas.2417202122)
67. Maiello G, Chessa M, Bex PJ, Solari F. 2020 Near-optimal combination of disparity across a log-polar scaled visual field. *PLoS Comput. Biol*. **16**, e1007699. (doi:10.1371/journal.pcbi.1007699)
68. Chessa M, Maiello G, Bex PJ, Solari F. 2016 A space-variant model for motion interpretation across the visual field. *J. Vis*. **16**, 12. (doi:10.1167/16.2.12)
69. Rideaux R, Storrs KR, Maiello G, Welchman AE. 2021 How multisensory neurons solve causal inference. *Proc. Natl Acad. Sci. USA* **118**, e2106235118. (doi:10.1073/pnas.2106235118)
70. Storrs KR, Maiello G. 2020 A model for neural network modeling in neuroscience. *J. Neurosci*. **40**, 7010–7012. (doi:10.1523/jneurosci.1205-20.2020)
71. Hartmann F, Maiello G, Rothkopf CA, Fleming RW. 2023 Estimation of contact regions between hands and objects during human multi-digit grasping. *J. Vis. Exp.* 64877. (doi:10.3791/64877)
72. Maiello G. 2025 Data to 'Human shape perception spontaneously discovers the biological origin of novel, but natural, stimuli' (v. 2.0). Zenodo. (doi:10.5281/zenodo.15039941)