The Influence of Shape on Human Categorization

Vorgelegt von

Henning Tiedemann

Dissertation

zur Erlangung des Doktorgrades der

Naturwissenschaften (Dr. rer. nat.) der

Justus-Liebig-Universität Gießen

Fachbereich 06, Psychologie und Sportwissenschaften

Otto-Behaghel-Straße 10F, 35394 Gießen

Table of contents

Abstract

An object's shape is one of the most salient and informative features that it possesses. Indeed, an object's shape usually defines what it is. Most of the time shape is all that is needed to rapidly assign an object to a category, in turn giving us a wealth of information about the object. This interaction of shape and categorization is the central topic of this dissertation.

Study 1 (Tiedemann, Morgenstern, Schmidt & Fleming, 2022) investigates the problem of one-shot learning – the human ability to infer categories from just one example – using a drawing-based generative task. We show that robust and varied perceptual categories can be inferred from viewing just one object. The part structure as well as highly distinctive individual parts of shapes were the main driving force in the creation of these categories.

Study 2 (Tiedemann, Schmidt & Fleming, 2022) focusses on the effect of a shape's part structure on superordinate categorization, specifically the category-boundary between animals and plants. The results demonstrate that subtle differences in curvedness and symmetry of parts, alongside the existence or absence of higher order parts can determine the superordinate category of objects.

Together, theses studies show that human perceptual categories, both known and novel, are heavily influenced and delineated by specific shape features, chief among them an object's part structure. Crucially, the specifics of part structure are a key ingredient in the human ability to learn categories quickly, with few examples.

Introduction

Categorization is an essential skill we possess to make sense of our surroundings: By grouping new objects with previously encountered ones, we are able to infer relevant characteristics in a fast and economical manner. Does it pose danger? Is it edible? Can it be used as a tool? These questions can usually be answered at a glance, such as with the variety of objects shown in **Fig 1**. While some of our category-knowledge is available in early age (14-month-old children appreciate the difference between living and artificial things; Mandler & McDonough, 1998) it is still subject to constant change through experience, such as when we encounter novel objects, which necessitates creating new categories. Categorization is guided in large part through vision. Consider what information is immediately available by looking at the objects in **Fig 1**: Observers can estimate an object's size, its color, whether it is alive or not (Schmidt, Hegele & Fleming, 2017)—in itself a primitive form of categorization—which way it is oriented (Cohen & Singh, 2006), aspects of its history (such as identifying missing parts; Spröte, Schmidt & Fleming, 2016) and much more. All this offers a wealth of information ready to be used to assign it a category, helping us, in turn, to quickly gain a deeper understanding of the object at hand.

Fig 1. Examples of objects from different categories. While visual characteristics of objects tend to vary tremendously, we are usually able to categorize them at a glance. This is achieved in large part through an objects shape (All images taken from pexels.com).

Categorization

Learning categories is an essential skill, as such it is not surprising that this facet of the human mind has been investigated intensively and for many decades. A large body of work concerns itself with potential strategies to learn categories, perceptual or otherwise. Early theories proposed a category-learning mechanism which centered around the comparison to previously encountered objects. One approach hypothesized the existence of ‘prototypes’ – an average, ‘ideal’ specimen representing a category – which newly encountered objects are compared to (e.g., Rosch, 1973). An example of this is the prototypical face, which is never actually experienced, but only derived from seen ones, which was shown to be erroneously remembered as having been encountered before (Bruce, Doyle, Dench & Burton, 1991). In contrast, exemplar theories assume that novel objects are compared to each encountered exemplar of potential categories, with the category with the highest similarity winning (e.g., Estes, 1986). What these and later theories (e.g., Shepard, 1987) have in common is a high-dimensional psychological feature-space, sometimes populated either by prototypes of different categories or a wealth of individual objects clustered in categories, in which new objects are located and grouped according to some rule. An important construct in many of these theories is the 'decision boundary' — thresholds within this high-dimensional space marking borders at which categorization changes from one class to another (Ashby & Gott, 1988). Spaces around these boundaries are often seen as probabilistic, with the chance for categorization dropping the further one moves away from a categories center and towards a decision boundary. 'Learning' here is the process of adjusting these boundaries and associated category regions. Importantly, while early theories generally propose one universal categorization method, later research uncovered that the type of task and stimuli can drastically change categorization behavior: Whether known or novel (often called 'artificial') categories are used, in how many dimensions the presented objects vary, and whether observers are experts or novices in the probed stimulus-field (Biederman & Shiffrar,1987), among other factors, determine how humans categorize, sometimes shown through considerably different activation patterns in the brain (for a review, see Ashby & Maddox, 2005).

Modelling human categorization performance has improved drastically in recent years with the machine learning revolution, with deep neural networks performing approximately as well as humans in object classification tasks (Krizhevsky, Sutskever & Hinton, 2012; Simonyan & Zisserman, 2014; Szegedy et al., 2015). While it is still an open question what aspects of these networks work as models of the human brain (Geirhos et al, 2019; Singer, Seeliger & Hebart, 2020; Heinke, Wachman, van Zoest & Leek, 2021) they offer a promising way to create human-like perception-systems in the future. Importantly, however, the most successful neural networks need thousands of examples to learn a new category. Humans, on the other hand, are capable of one-shot learning – the ability to infer a category from one novel piece of data (Lake, Salakhutdinov & Tenenbaum, 2013; Morgenstern, Schmidt & Fleming, 2019). How humans (and even some animals, such as pigeons; Cerella, 1979) are able to achieve this is poorly understood and this ability is as of yet barely replicated in artificial intelligence (however, see Lake, Salakhutdinov & Tenenbaum, 2013; Jadon, 2020).

At first blush, one-shot learning seems impossible: How could the extent of a category in the feature space be estimated from just one object, which in itself cannot show any range within its features at all? Many theories, including most prototype and exemplar theories, posit that categorization is based on heuristic and inductive comparisons between objects where relevant features are chosen ad-hoc depending on what is present. While this strategy is indeed used by humans and is very efficient (Gigerenzer & Gassmaier, 2011), it is unable to explain one-shot learning, where no feature-space can be established, and comparisons could only be made to pre-existing spaces, which might not fit the category at hand.

More sophisticated models posit that instead of inductively extracting relevant features from experienced objects, we make use of certain expectations about how objects in the world tend to come about and vary, which biases us towards certain categorizations (Feldman, 1992; Fei-Fei, Fergus & Perona, 2006; Goodman, Tenenbaum, Feldman & Griffiths, 2008). An example of this is the configuration of body parts in animals, which are, due to highly regular growth processes, almost always resulting in a bilaterally symmetric body plan (Reece et al., 2014), instead of a random arrangement of limbs, as shown in **Fig 2a**. If this regularity is expected by the human mind, it makes it a powerful feature to use in categorizing an object as an animal, thus eliminating the need to exhaustively search other features for a fitting candidate. The role of growth processes as category-defining features will be investigated in Study 2. Feldman (1992, 1997) called category-defining features of this kind “non-accidental” – features that occur regularly in the world yet are unlikely to be the result of chance. That non-accidental features can also be purely statistical, without a real-world basis in e.g., growth processes, was shown by Feldman (1997) with a simple dot and a line (see **Fig 2b**): If both the line and dot are randomly placed in an area, the chance of the dot being placed on the line is extremely small. Instead, the dot and line would be placed in separate places without touching one another most of the time. Thus, if we observe a dot on a line, we would believe that this is a non-accidental feature of that object (dot *and* line) – and its category – instead of a random configuration of two distinct objects. In turn, objects not possessing this feature might not be assigned the same category, even if other, accidental features overlap, such as line-length or color. The existence of these biases makes possible a strategy to solve one-shot learning: Seeing an object and extracting its non-accidental features, we could then synthesize a new object, purely through imagination, that is constant in the non-accidental, yet varies in other features to create a new, visually distinct object that still has most signifiers of the original objects category. This process has been tested successfully by Feldman (1997). Here, participants were presented with an exemplar object – a dot on a line – and were asked to imagine a new object of that category and draw it. Participants mostly created dots on a line as opposed to separate dots and lines, showing that this feature was deemed critical for its category membership. Although these results are impressive, they were drawn from highly simplistic exemplar objects, which could realistically only be varied in very few dimensions. While this nicely illustrated the power of non-accidental features in principle, it did not allow the discovery of what kind of features would be deemed non-accidental by humans in more complex, realistic stimuli, where such features might be difficult to predict. Thus, Study 1 of this dissertation substantially expanded on this method with much more varied and complex exemplar objects and extensive follow-up studies on the generated drawings, to probe category-defining features in more detail.

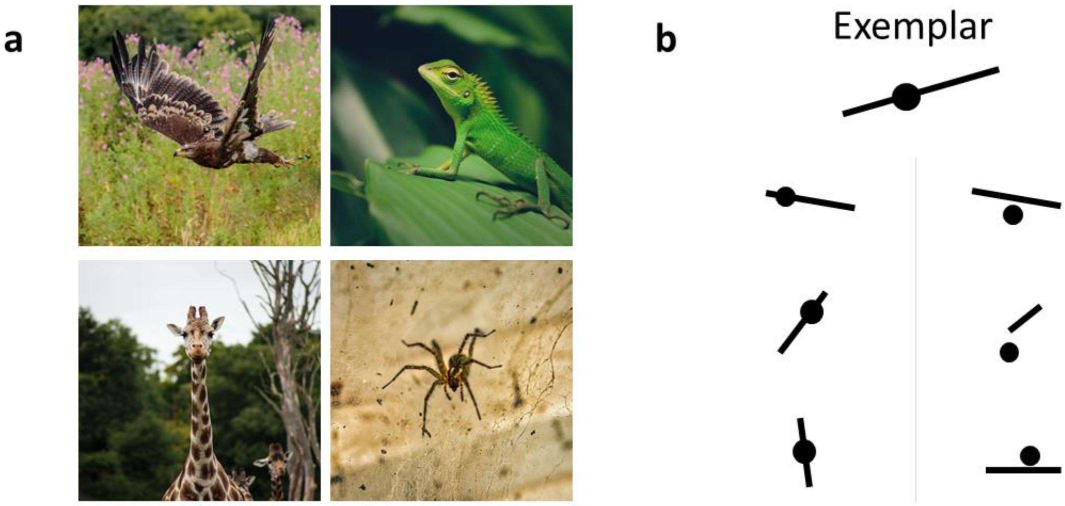

Fig 2. (a) Examples of bilaterally symmetric animals. Almost all mobile organisms share a symmetric growth pattern, thus providing a prime candidate for a category-defining feature (All images taken from pexels.com). **(b)** One-shot categorization with a novel stimulus. Feldman (1997) has shown that observers tend to associate the objects on the left with the exemplar (top) but not those on the right. This is because the conjunction of dot and line was deemed as a category-defining feature, showing that such features need not have a basis in nature.

A Bayesian expansion of the previously mentioned high-dimensional psychological space-theories by Tenenbaum & Griffiths (2001) sheds light on how the sample-size in category-learning (i.e., one- vs many-shot learning) can influence decision boundaries and regions: The more samples of a category are available, the steeper the decision boundary gets, making category-decisions along this boundary more clear-cut. This would mean that in a one-shot scenario we would be much less secure in our grouping with the available exemplar, because of the shallow decision boundary, but we would still group objects with the observed one. The size of category region (i.e., the region in which categorization is highly likely), in turn, is dependent on the witnessed variability of samples, with more variability expanding the region, resulting in more varied objects having a high likelihood of being grouped here. Morgenstern, Schmidt & Fleming (2019) showed empirical support for this theory. In addition, they found that with a varied sample of objects, a fixed set of features explained categorization decisions across categories, whereas in a one-shot scenario each category necessitated a differently weighted set of features. This finding is investigated further in Study 1, in which the generated objects of each category possessed individual distinctive parts that ended up being highly category-defining.

Shape perception

The second major focus of this dissertation is shape perception. One of the most salient visual features of objects is their shape. In fact, in most circumstances it is the single most defining quality an object possesses. Considering the variety of shapes we encounter in our environment, as shown in **Fig *1***, it is hardly surprising that this distinctive visual feature is made ample use of by our mind to recognize and learn object categories (Marr & Nishihara, 1978; Samuelson & Smith, 2005; Landau, Smith & Jones, 1998). Shape perception allows us to gauge when an object will stand or fall over (Samuel & Kerzel, 2011; Cholewiak, Fleming & Singh, 2013) or influences how we grasp an object (Cuijpers, Smeets & Brenner, 2004), among many other things. While 'shape' is an intuitively easy to understand concept, it is far from easy to define and has been defined in several ways, as discussed in Todd and Petrov's (2022) review of shape perception research. While these intricacies of definition persist, the field of shape perception has made much progress in explaining how this part of the human visual system works.

An essential step in analyzing a shape is to segment it into visual parts, giving us vital insights into the function and capabilities of objects (Hoffman & Richards, 1984): **Fig 3a** shows the silhouette of a cat. Immediately and inevitably, we perceive individual parts in concert with the whole object (see **Fig 3b**). We also quickly assign functions to these parts, such as the legs being the means of locomotion. It has been suggested that breaking down an object into parts allows for a more efficient analysis, by creating smaller subproblems as well as offering a higher order object representation with parts acting as symbolic tokens (Todd & Petrov, 2022). Such tokens can represent a variety of things, such as the previously mentioned function (hands, claws and an elephant's trunk could all be represented and treated as a "grasping part", even though they are visually quite distinct). They can also represent more abstract qualities, such as their role in a shape's part structure (a humans and a birds shoulders are connected to different types of limbs (arms vs. wings), yet hold the same place in the objects part-hierarchy), making the identification of equivalent parts across visually diverse categories easier (Schmidt, Kleis, Morgenstern & Fleming, 2020).

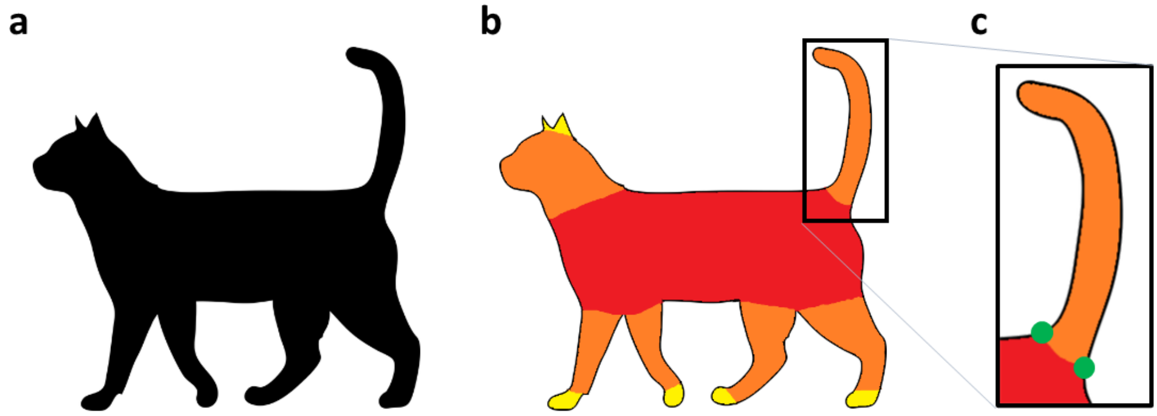

Fig 3. (a) This silhouette of a cat is immediately analyzed in terms of parts, even if presented in this reduced fashion. **(b)** The same object with visual parts delineated. **(c)** Concave point pairs (green circles) are an important visual feature used by the human mind to segment objects into parts.

A first step in any part-based analysis is of course to arrive at a part segmentation (Biederman, 1987; Hoffman & Richards, 1984; Cave & Kosslyn, 1993; Siddiqi, Tresness & Kimia, 1996). Visual part segmentation occurs early and in parallel over the visual field (Baylis & Driver 1995; Wolfe & Bennett, 1997; Hulleman, te Winkel, & Boselie, 2000; Machilsen, B.; Pauwels, M.; Wagemans, J, 2009; Xu & Singh, 2002). Many rules of perceptual part segmentation have been uncovered (e.g., Singh, Seyranian & Hoffman, 1999), chief among them the minima rule (Hoffman & Richards, 1984), in which aligned points of concave discontinuity are identified and used to delineate parts (see **Fig 3c**). The ubiquity of such aligned minima is the result of the transversality regularity: "When two arbitrarily shaped surfaces are made to interpenetrate they always meet in a concave contour of discontinuity of their tangent planes" (From Hoffman & Richards, 1984). Since many growth processes of biological organisms create such a configuration (although, of course, no interpenetration occurs), such as the growth of human legs or twigs on a tree, this correlation between biological part and resultant contour boundary explains why humans have evolved to make ample use of this visual feature.

Importantly, not all parts are treated equally: Certain parts are more salient than others, depending on their relative size and how much they protrude, among other factors (Hoffman & Singh, 1997). Individual parts can act as highly potent category-definers (Tversky, 1989) and even 3-month-old infants have been shown to categorize animals such as cats and dogs using the visual features of their heads as opposed to their bodies or global shape (Quinn, Eimas, Tarr, 2001).

Another critical feature of parts is their spatial relation, which alone can successfully model the perceived similarity of objects (e.g., Latecki & Lakamper, 2000). The importance of part-relations is illustrated by humans suffering from integrative agnosia, a form of visual agnosia that impairs the ability to integrate parts into higher-order objects, which impairs object recognition (Behrmann, Peterson, Moscovitch & Suzuki, 2006). Computationally, different algorithms have been proposed that generate an internal skeleton for a shape and its individual parts, which are often highly similar to the object's perceptual part structure (Feldman & Singh, 2006; Bai & Latecki, 2008; Lowet, Firestone & Scholl, 2018). Given the importance of parts and their relations, many theories of object perception have been formulated in part-centric terms (e.g., Palmer, 1977; Marr & Nishihara, 1978; Tversky & Hemenway, 1984; Biederman, 1987), with every object consisting of a set of primitive base-shapes, immensely reducing the number of possible shape descriptions. The effects of part structure on categorization are investigated in both Study 1 (through a drawing task) and Study 2 (in relation to real-world growth processes).

When it comes to the categorization of shapes, the general categorization problem of which features to use for grouping applies the same way: In theory, there are an infinite number of shape features that could be extracted, compared and used to form category boundaries. For known categories, semantic knowledge can solve this problem. We know that birds have wings, therefore an object with wings is likely to be a bird (Tversky, 1989) – even if how exactly we conclude a part to be a wing is yet unclear. Similarly, more global shape characteristics such as symmetry or curvature, influence our superordinate categorization for animals, plants or minerals (Schmidt, Hegele & Fleming, 2017). For novel categories, however, semantic knowledge is unavailable, yet humans still agree on what features are relevant for new categories even if they exhibit unusual visual characteristics such as technological innovations (Bicycles look quite unlike anything in our evolutionary past, yet we have no problem identifying them). Previous studies also found that there is no universal set of features used in categorizing novel objects (Morgenstern, Schmidt & Fleming, 2019), meaning that depending on the object, different features are deemed relevant. The question now is what these features might be. This brings us back to idea of non-accidental features (Feldman, 1997) – features that are unlikely to have come about by chance and thus tell us something deep about the object and its origin. Because of the special nature of non-accidental features, they might be prime candidates for category defining features.

This problem – which features we deem important depending on the object at hand – is a central question of this dissertation: Study 1 probes highly abstract shapes with a generative task in a one-shot scenario to find which features are deemed category-defining. Study 2 investigates in what way typical biological growth-processes determine our categorization and how the resulting parts and their relations might be candidates for non-accidental features.

Study 1: One-shot generalization in humans revealed through a drawing task

The first study of this dissertation (Tiedemann, Morgenstern, Schmidt & Fleming, 2022) aimed to answer the question how humans are able to learn a novel category from just a single object. Specifically, we probed what shape-characteristics humans deem important and category-defining depending on the presented object. We also aimed to investigate whether relevant features we uncovered fit into Feldman's (1997) idea of non-accidental features.

To do this we employed the generative one-shot categorization method, a method pioneered by Jacob Feldman (1992, 1997). In this task, participants are presented with a single object – an exemplar – and asked to imagine and then draw a new object (or multiple objects) that belongs to the exemplar's category. This generative approach diverges considerably from the vast majority of categorization research, which employed mostly discriminative methods (i.e., tasks in which participants evaluate groups of pre-selected stimuli; Ashby & Maddox, 2005). The downside of such purely discriminative tasks is two-fold. First, since questions of category-membership are asked in the context of a limited number of stimuli, participants can resort to simple strategies in which objects are simply compared to one another 'in a vacuum'. While this is very efficient and can be valid strategy in certain scenarios (such as in scenarios with low predictability, Gigerenzer & Gaissmaier, 2011), it does not reflect categorization in the real-world, in which it is not always clear to which other objects a witnessed object needs to be compared. In fact, comparing objects only to other objects in the vicinity is probably the least human-like way to categorize. Furthermore, these tasks can theoretically be solved by locating and comparing the stimuli in a pre-existing feature space. This means that a purportedly novel category is not treated as such, and no conclusions about novel feature spaces can be drawn.

Second, discriminative tasks necessitate that the range of stimuli is entirely pre-selected, either by the experimenters or by an algorithm chosen by the experimenters. This runs the risk that the resultant categories span a highly artificial feature space, in which learned category boundaries might be in part an artifact of the chosen stimuli and less of natural categorization choices. While this still gives us valuable insight into these specific feature-spaces, it brings into question the generalizability of such findings.

The generative one-shot categorization task solves these problems: Because only a single object is shown, no artificial feature space is established to which observers might be biased. Likewise, since no other objects are presented, the strategy of ad-hoc comparisons is unavailable, in turn forcing the participant to engage more deeply with the object and analyze it in terms of features that are deemed important on its own, potentially even inferring a generative process that might have created the object, which is then imitated. Because of the generative nature of the task – participants are drawing objects as opposed to, say, picking them from a pre-selected list - they are free to choose any number of features they deem relevant for that category to create new objects, showing us without artificial restrictions, what is considered category-defining and what is not.

In the study participants were shown 8 exemplar shapes consecutively and asked to draw 12 new objects per exemplar. Exemplars were chosen to be abstract, to limit the influence of semantic knowledge, and to vary in visual complexity, i.e., in the number of parts. This way 204 drawings per exemplar-category were collected. A variety of follow up experiments using new groups of independent observers and analyses allowed the following new insights into how humans solve one-shot learning:

1. The drawn shapes varied tremendously, with most shapes looking distinctly different from the original exemplar, as shown in **Fig 4a** – a consequence of the unconstrained task that tapped into the creativity of the participants. Crucially, the large majority of drawings were seen as (a) more different from the exemplar than mere copies of exemplars and (b) belonging to the category of the originating exemplar and not another category. This demonstrates that the generative one-shot categorization task is successful in creating varied yet robust categories from just one object.

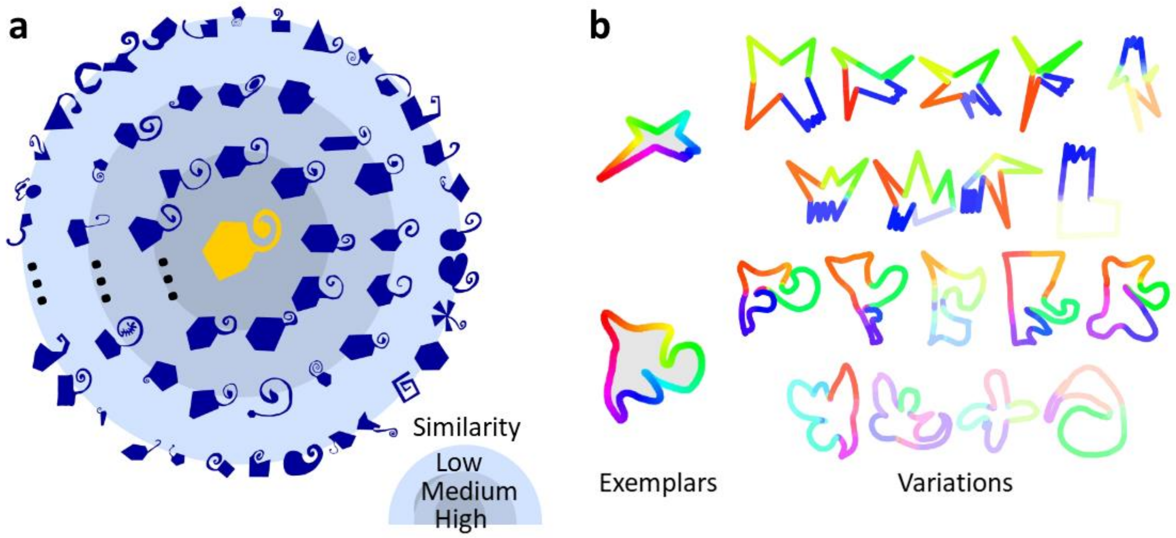

Fig 4. (a) A subset of the drawings for one category, with the exemplar shown in the middle in yellow and drawings in blue. Drawings closer to the center are more similar to the exemplar than those on the outer rings. **(b)** Visualization of perceptual part-correspondence data for two categories. The exemplars part structure was recreated faithfully, even in visually quite different drawings.

2. Drawings tended to retain the exemplar's part structure, namely the number and spatial relations of parts, even if parts themselves were altered considerably (see **Fig 4b**). This shows that, even though the exemplars were novel, abstract objects, participants still determined the part structure to be a relevant feature of that category.

3. Certain parts (both exemplar and drawing parts) were consistently seen as more distinctive than others by independent observers. These distinctive parts were shown to be category-defining by swapping out the original distinctive part with one of another category, in turn often changing the categorical association with it. Notably, this effect was not found when an indistinctive part was swapped with another indistinctive part. This suggests that drawers, when creating a new shape, rate each part in terms of its importance to the object's category and recreates it in such a way that the part retains the critical features that made it category-defining.

Together, these findings give us first insights into generative models that humans employ to solve one-shot learning. Many of the drawn shapes could have been created using a part-based creation strategy: First, the exemplar was segmented into parts, which were analyzed in terms of their distinctiveness. These parts were then transformed visually in a myriad of ways, though crucially so that distinctive parts kept the features that made them distinctive. The altered parts were then ‘reassembled’ into a new object. In this step, parts could be added, removed or substituted and the order of parts could be changed, though this happened rarely. Other strategies are harder to formalize: Some statistical motifs, such as the ‘spikiness’ of exemplars was reproduced in shapes that deviated in terms of part structure, showing that higher level features can also be used to create diverse yet still valid category members.

An important finding is that even in novel, abstract objects, part structure is considered a crucial category feature. An open question is what exact features of part structure are deemed relevant, such as whether certain parts are more or less category-defining depending on their place in the objects part structure (e.g., central versus ‘outer’ parts), and to what extent real-world growth processes and their resultant part structures bias our preferences. This last question—the role of real-world growth regularities in human categorical decisions—is investigated in Study 2.

Study 2: Superordinate categorization based on the perceptual organization of parts

Study 2 of this dissertation (Tiedemann, Schmidt & Fleming, 2022) investigates how an objects part-structure – the spatial arrangement and connections of an object’s parts – influences human categorization. Specifically, the influence of typical real-world growth patterns is tested with regards to categorization choices between animals and plants.

Quickly deciding whether an object belongs to a superordinate category such as plants, animals or artifacts is an essential step for humans to negotiate our surroundings. Especially in situations necessitating fast reactions, such as fight or flight scenarios, it is advantageous to quickly assign an object a broad class such as animal, before spending crucial time assigning more detailed category-memberships. Previous research has shown certain key features relevant for superordinate categorization such as an object’s curvature (Schmidt, Hegele & Fleming, 2017; Long, Störmer & Alvarez, 2017) or characteristic parts (Tversky, 1989). These and many more visual characteristics of objects are the result of their origin, which for most organic objects is their growth process. Biological growth processes tend to follow a limited set of rules (Stevens, 1974), two of which are central to this study: Plants grow from a central body, from which periodically new parts sprout. These new limbs then also sprout new limbs, resulting in a complex part hierarchy, as shown in **Fig 5**. The vast majority of animals, on the other hand, start from a single main body, out of which sets of symmetrical limbs grow (Reece et al, 2014), which themselves rarely sprout new parts. This existence (or lack) of higher order parts is potentially a potent category definer, on account of the ubiquitous adherence of animals and plants to their respective typical growth pattern.

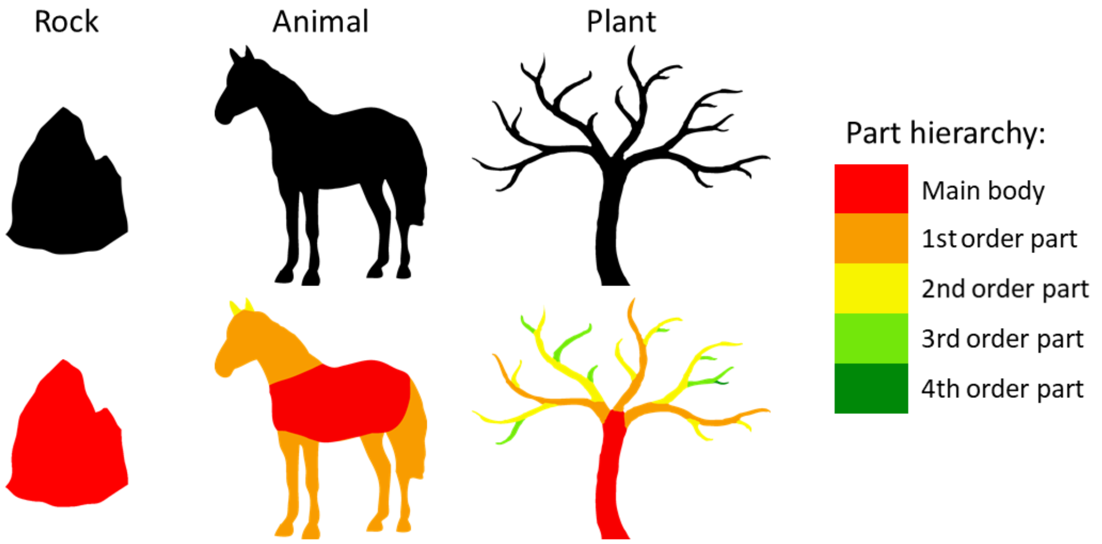

Fig 5. Typical part hierarchies of different superordinate categories. One salient difference between rocks, animals and plants is their part hierarchy, which describes how far removed from the main body parts grow.

The effect of these different part structures was tested in terms of superordinate categorization, specifically, whether objects with different part structures were categorized as plants or animals. For this reason, abstract shapes were created that differed in their part structure but were otherwise quite simple, without having, e.g., typical curvature patterns, textures or parts such as eyes. This way the effect of part structure could be isolated and also continuously varied. In three two alternative forced choice (2AFC) tasks, multiple types of part structure differences were tested: Symmetrical versus asymmetrical part pairs, curved versus straight parts and shapes with sprouting second order limbs versus shapes without second order limbs. Across the board, participants' categorization preferences were in line with biological growth patterns, i.e., objects grown in a typically plant-like way (asymmetrical, relatively straight limbs with second order parts) were judged to be plants as opposed to animals, which were seen to have symmetrical, curved limbs without higher order parts. This effect was tested further by creating sequences of shapes, in which the part structure of the same object was continuously changed from animal- to plant-like, by making symmetrical part pairs continuously more asymmetric, limbs less curvy and adding second order limbs. Showing steps of many such sequences in a random order (so that individual sequences were not apparent) and letting participants rate each shape on a continuous scale from animal- to plant-like gave us insight into how fine-grained effects of part structure can be. Indeed, participants' ratings were highly correlated to the variations in part structure, with ratings continuously moving from the animal- to the plant-end of the rating-spectrum with associated changes in part structure. This is especially striking given the quite subtle visual changes from step to step in the sequences, underlining just how sensitive the human visual system seems to be to such characteristics.

Together, these findings show how features of part structure guide our superordinate categorization. Spatial relations among parts, such as their symmetry, as well as hierarchical order among parts, like the existence of higher order parts, influence whether we categorize a novel object as plant or animal. These expectations are in line with typical biological growth patterns, suggesting that humans have evolved to expect these patterns in unfamiliar objects. Since it is possible that characteristics of part structure are analyzed in tandem or even as a by-product of part segmentation, it is conceivable that such features are extracted equally early in the visual pipe-line. Typical part structures also fit nicely into the idea of non-random features and generative models: Since growth processes are not random, they tend to create objects with distinctive statistical features. This is also true for other types of object-transformations, such as when paper is crumpled it will be changed in distinctive ways (Fleming & Schmidt, 2019). Therefore, visual features resulting from transformations are prime candidates for categorization, especially 'unavoidable' transformations such as growth patterns, which organisms have to undergo to exist in the first place.

Discussion

An object's shape is its most defining characteristic. Accordingly, humans make ample use of this visual feature to categorize objects into classes both known and novel. But how do we choose which characteristics are relevant for categorization and which are not? In theory, any number of features could be used, yet humans agree on what objects belong together. This holds true even for novel objects, which might not belong into any previously encountered category. This dissertation aimed to investigate the problem of categorization, specifically which features of an object's shape humans use to distinguish categories of novel objects. Study 1 investigated the problem of one-shot learning via a generative task, showing that part structure as well as distinctive parts drove one-shot learning. Study 2 further probed how different part structures, as they tend to result from real-world biological growth processes, influence our superordinate categorization. Here it was found that part-structure features such as symmetry of parts and the hierarchical order of parts determined category membership between animals and plants.

One-shot learning

Humans are able to learn novel categories from a small number of examples. Even a single novel object usually gives us a solid grasp on what other members of that category might look like. Study 1 investigated this ability through a generative task. Participants were presented with a single novel object—an Exemplar—and were asked to draw new objects that belonged to the same category. The resulting categories proved to be visually varied yet were seen as cohesive categories by independent observers. This shows that humans are indeed able to mentally create a whole category from one piece of data and, crucially, that other humans agree that these newly created objects and the Exemplar belong together. Given the abstract nature of the Exmplars, it was particularly interesting which visual features were preserved in the newly created objects, something that will be discussed in the next section.

Non-random features driving categorization

One of the main questions of human categorization is how we decide what visual features are relevant in any given scenario, of which there are infinitely many to choose from. Theories posit that non-random features – features that are the result of regularities of our environment – make fast and robust categorization possible (e.g., Feldman, 1997). Both studies of this dissertation aimed to uncover some of these special features, both for abstract (Study 1) and biologically inspired (Study 2) objects.

Both Study 1 and 2 showed that part structure – the spatial and hierarchical relations among an object's parts – are crucial for categorization. While many perceptual theories posit that humans represent objects in a part-based fashion (e.g., Biederman, 1987; Hoffman & Richards, 1984), less attention has been put on how the hierarchical organization of these parts themselves might be an important perceptual feature. Study 2 showed that two objects made up of the same set of parts can be seen as belonging to two different superordinate categories simply by changing the hierarchical order of those same parts. While the stimuli in this study were inspired by typical biological growth patterns, it is conceivable that this effect is not limited to such realistic examples. Artificial objects we encounter in our life are not subject to the same limitations as biological objects are, meaning they often possess visual organizations quite unlike any animals or plants, such as a bicycle or satellite. Yet, even if we only see a silhouette of these objects, we usually have no problem associating them with the correct category. This notion that part structures are seen as important regardless of their 'naturalness' is confirmed in Study 1, in which the structure of highly abstract shapes was recreated in many drawings. Therefore, since part structures tend to be the result of non-random processes – even artificial objects are subject to constraints such as stability – it is conceivable that humans are biased to interpret a novel objects part structure as a relevant feature, since real-world objects parts are rarely assembled at random. Accordingly, newly generated objects inspired by the Exemplars in Study 1 often maintained the original part structure, even though participants were free to alter the shape in any way they wanted. That alterations did occur, e.g., by removing or adding parts from the Exemplar, suggests that some details of the Exemplars part structure were considered malleable and legal for some alteration. If participants were asked to draw a new object of a known class, such as humans, it is much less likely that they would have added new parts not present in a normal human body.

Another crucial finding of Study 1 is the importance of distinctive parts. While previous research has shown that certain individual parts drive categorization (Tversky, 1989), they tended to be parts that were semantically loaded, such as wings, which are not only a visual cue towards the bird class, but also in their function, i.e., by enabling flying. In Study 1, different groups of observers working on different kinds of tasks, agreed on what parts were the most distinctive in highly abstract novel shapes. Since these parts did not have a clear semantic function, other factors might have made them stand out from the rest. Importantly, these parts were not only seen to stand out, they also determined category membership to a large degree: If the most distinctive part of an object was swapped out by a distinctive part of another category, in many cases the category changed with it, even if the parts made up only a small amount of the whole object. What exactly makes a part distinctive is as of yet unclear. One potential explanation is that certain distinctive parts deviated from global features, making the distinctive part a statistical outlier within the whole shape. One example of this is the "twirl"-part in the Exemplar shown in **Fig 4a** (center, in yellow), which was deemed highly distinctive and deviates from the other portions of the shape in features such as curvedness and perimeter length. This particular example creates an interesting question: Since this object is made up of two perceptual parts, why was the twirl chosen as distinctive and not the other one? Both could be seen as the outlier compared to the other, yet the twirl was consistently chosen. Potentially, humans have an internal taxonomy of typical parts, to which newly encountered parts are compared. If these internal parts are weighted in terms of how likely they are to be encountered, then parts deemed rare might be seen as more distinctive, compared to more basic and common parts. In the "twirl" example the non-distinctive part is a simple convex shape that could be found in a myriad of real-world objects, such as rocks. The twirl, on the other hand, might be encountered much more rarely, e.g., in ferns, thus winning out as the distinctive part in this scenario.

The distinctive parts from Study 1 give us another tentative insight on effects of part structure: Almost all distinctive parts are parts that look like they sprout from or are attached to the 'main body' of the object. This 'main body' – the central part of an object from which other parts seem to originate – could in theory serve as a candidate for most distinctive part, since it is usually the biggest part and is connected to most other parts. Yet in Study 1 distinctive parts were never the main body (except for the Exemplar which only consisted of a main body). Thus, it seems that the main body is not considered for the role of a category-defining part, adding another potential explanation why the aforementioned 'twirl' was chosen over the other part. This might possibly be because real-world main-bodies visually vary less than other parts. Think of a human torso and a horse's central part. If we compare them purely with shape characteristics, they are far more similar to each other than a human's hand and a horse's hoofs. Since main bodies are a necessary 'starting point' of any complex organism, usually housing the organs, other parts tend to take over more specialized and thus category-specific functions, such as claws, hands and wings, also explaining their more diverse appearance. This might be the reason that main bodies were not chosen as a distinctive part in Study 1 and extremities were chosen instead. Therefore, what makes a part a main body and thus narrows down potential category-defining parts is an important question. Unpublished work by the author of this dissertation suggests that humans are biased towards seeing the most central part as the main body. Other factors, such as the amount of neighboring parts also seems to play a role. Study 2 implicitly assumed the central part to be the main body, which in turn determined the whole object's part hierarchy. In this case, real-world growth processes were used as inspiration and it is likely that a novel objects part hierarchy is likewise influenced by visual analyses of potential growth histories and their visual signatures.

Conclusion

The ability to categorize objects both novel and known is a fascinating facet of the human mind. This dissertation aimed to further our understanding of what visual features we use to categorize novel objects. Study 1 investigated one-shot scenarios, where we found that the part structure, as well as other shape-features, such as curvature, are identified as important for that object's category. Most importantly, however, were individual distinctive parts, which decided category-membership to a large extent, even though they often comprised only a small part of the whole object. This effect was so strong that exchanging the distinctive part of an object with that of another object often changed category-membership along with it. An interesting future question is what exactly makes these and other parts so distinctive.

In Study 2 we investigated how typical biological growth patterns influence human categorization. We found that different part structures, as they tend to result from different biological growth processes, impact our perception in the domain of superordinate categories, such as plants and animals. Specifically, the importance of symmetrical/asymmetrical parts as well as of higher order parts heavily influences whether we judge an object to be an animal or plant.

References

Ashby, F. G., & Gott, R. E. (1988). Decision rules in the perception and categorization of multidimensional stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(1), 33.

Ashby, F. G., & Maddox, W. T. (2005). Human category learning. *Annu. Rev. Psychol.*, *56*, 149–178.

Bai, X., & Latecki, L. J. (2008). Path similarity skeleton graph matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *30*(7), 1282–1292.

Baylis, G. C., & Driver, J. (1995). Obligatory edge assignment in vision: The role of figure and part segmentation in symmetry detection. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(6), 1323.

Behrmann, M., Peterson, M. A., Moscovitch, M., & Suzuki, S. (2006). Independent representation of parts and the relations between them: Evidence from integrative agnosia. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(5), 1169.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*(2), 115.

Biederman, I., & Shiffrar, M. M. (1987). Sexing day-old chicks: A case study and expert systems analysis of a difficult perceptual-learning task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(4), 640.

Bruce, V., Doyle, T., Dench, N., & Burton, M. (1991). Remembering facial configurations. *Cognition*, *38*(2), 109–144.

Cave, C. B., & Kosslyn, S. M. (1993). The role of parts and spatial relations in object identification. *Perception*, *22*(2), 229–248.

Cerella, J. (1979). Visual classes and natural categories in the pigeon. *Journal of Experimental Psychology: Human Perception and Performance*, *5*(1), 68.

Cholewiak, S. A., Fleming, R. W., & Singh, M. (2013). Visual perception of the physical stability of asymmetric three-dimensional objects. *Journal of Vision*, *13*(4), 12–12.

Cohen, E. H., & Singh, M. (2006). Perceived orientation of complex shape reflects graded part decomposition. *Journal of Vision*, *6*(8), 4–4.

Cuijpers, R. H., Smeets, J. B., & Brenner, E. (2004). On the relation between object shape and grasping kinematics. *Journal of Neurophysiology*, *91*(6), 2598–2606.

Estes, W. K. (1986). Array models for category learning. *Cognitive Psychology*, *18*(4), 500–

549.

Feldman, J. (1992). Constructing perceptual categories. *Proceedings 1992 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 244–250.

Feldman, J. (1997). The structure of perceptual categories. *Journal of Mathematical Psychology*, *41*(2), 145–170.

Feldman, J., & Singh, M. (2006). Bayesian estimation of the shape skeleton. *Proceedings of the National Academy of Sciences*, *103*(47), 18014–18019.

Fleming, R. W., & Schmidt, F. (2019). Getting "fumpered": Classifying objects by what has been done to them. *Journal of Vision*, *19*(4), 15–15.

Geirhos, R., Rubisch, P., Rauber, J., Temme, C. R. M., Michaelis, C., Brendel, W., Bethge, M., & Wichmann, F. A. (2019). Inducing a human-like shape bias leads to emergent human-level distortion robustness in CNNs. *Journal of Vision*, *19*(10), 209c–209c.

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, *62*(1), 451–482.

Goodman, N. D., Tenenbaum, J. B., Feldman, J., & Griffiths, T. L. (2008). A rational analysis of rule-based concept learning. *Cognitive Science*, *32*(1), 108–154.

Heinke, D., Wachman, P., van Zoest, W., & Leek, E. C. (2021). A failure to learn object shape geometry: Implications for convolutional neural networks as plausible models of biological vision. *Vision Research*, *189*, 81–92.

Hoffman, D. D., & Richards, W. A. (1984). Parts of recognition. *Cognition*, *18*(1–3), 65–96.

Hoffman, D. D., & Singh, M. (1997). Salience of visual parts. *Cognition*, *63*(1), 29–78.

Hulleman, J., Winkel, W. te, & Boselie, F. (2000). Concavities as basic features in visual search: Evidence from search asymmetries. *Perception & Psychophysics*, *62*(1), 162–174.

Jadon, S. (2020). An overview of deep learning architectures in few-shot learning domain. *ArXiv Preprint ArXiv:2008.06365*.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84–90.

Lake, B. M., Salakhutdinov, R. R., & Tenenbaum, J. (2013). One-shot learning by inverting a compositional causal process. *Advances in Neural Information Processing Systems*, 2526–2534.

Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development*, *3*(3), 299–321.

Latecki, L. J., & Lakamper, R. (2000). Shape similarity measure based on correspondence of visual parts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*(10), 1185–1190.

Long, B., Störmer, V. S., & Alvarez, G. A. (2017). Mid-level perceptual features contain early cues to animacy. *Journal of Vision*, *17*(6), 20–20.

Lowet, A. S., Firestone, C., & Scholl, B. J. (2018). Seeing structure: Shape skeletons modulate perceived similarity. *Attention, Perception, & Psychophysics*, *80*(5), 1278–1289.

Machilsen, B., Pauwels, M., & Wagemans, J. (2009). The role of vertical mirror symmetry in visual shape detection. *Journal of Vision*, *9*(12), 11–11.

Mandler, J. M., & McDonough, L. (1998). On developing a knowledge base in infancy. *Developmental Psychology*, *34*(6), 1274.

Marr, D., & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, *200*(1140), 269–294.

Morgenstern, Y., Schmidt, F., & Fleming, R. W. (2019). One-shot categorization of novel object classes in humans. *Vision Research*, *165*, 98–108.

Palmer, S. E. (1977). Hierarchical structure in perceptual representation. *Cognitive Psychology*, *9*(4), 441–474.

Quinn, P. C., Eimas, P. D., & Tarr, M. J. (2001). Perceptual categorization of cat and dog silhouettes by 3-to 4-month-old infants. *Journal of Experimental Child Psychology*, *79*(1), 78–94.

Reece, J. B., Urry, L. A., Cain, M. L., Wasserman, S. A., Minorsky, P. V., & Jackson, R. B. (2014). *Campbell biology*. Pearson Boston; pp. 633.

Rosch, E. H. (1973). Natural categories. *Cognitive Psychology*, *4*(3), 328–350.

Samuel, F., & Kerzel, D. (2011). Is this object balanced or unbalanced? Judgments are on the safe side. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(2), 529.

Samuelson, L. K., & Smith, L. B. (2005). They call it like they see it: Spontaneous naming and attention to shape. *Developmental Science*, *8*(2), 182–198.

Schmidt, F., Hegele, M., & Fleming, R. W. (2017). Perceiving animacy from shape. *Journal of Vision*, *17*(11), 10–10.

Schmidt, F., Kleis, J., Morgenstern, Y., & Fleming, R. W. (2020). The role of semantics in the perceptual organization of shape. *Scientific Reports*, *10*(1), 1–19.

Shepard, R. N. (1987). Toward a universal law of generalization for psychological science.

Science, *237*(4820), 1317–1323.

Siddiqi, K., Tresness, K. J., & Kimia, B. B. (1996). Parts of visual form: Psychophysical aspects. *Perception*, *25*(4), 399–424.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *ArXiv Preprint ArXiv:1409.1556*.

Singer, J., Seeliger, K., & Hebart, M. N. (2020). The representation of object drawings and sketches in deep convolutional neural networks. *NeurIPS 2020 Workshop SVRHM*.

Singh, M., Seyranian, G. D., & Hoffman, D. D. (1999). Parsing silhouettes: The short-cut rule. *Perception & Psychophysics*, *61*(4), 636–660.

Spröte, P., Schmidt, F., & Fleming, R. W. (2016). Visual perception of shape altered by inferred causal history. *Scientific Reports*, *6*, 36245.

Stevens, P. S. (1974). *Patterns in Nature* (1st ed.). Little Brown & Co.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1–9.

Tenenbaum, J. B., & Griffiths, T. L. (2001). Generalization, similarity, and Bayesian inference. *Behavioral and Brain Sciences*, *24*(4), 629–640.

Tiedemann, H., Morgenstern, Y., Schmidt, F., & Fleming, R. W. (2022). One-shot generalization in humans revealed through a drawing task. *Elife*, *11*, e75485.

Tiedemann, H., Schmidt, F., & Fleming, R. W. (2022). Superordinate Categorization Based on the Perceptual Organization of Parts. *Brain Sciences*, *12*(5), 667.

Todd, J. T., & Petrov, A. A. (2022). The many facets of shape. *Journal of Vision*, *22*(1), 1–1.

Tversky, B. (1989). Parts, partonomies, and taxonomies. *Developmental Psychology*, *25*(6), 983.

Tversky, B., & Hemenway, K. (1984). Objects, parts, and categories. *Journal of Experimental Psychology: General*, *113*(2), 169.

Wolfe, J. M., & Bennett, S. C. (1997). Preattentive object files: Shapeless bundles of basic features. *Vision Research*, *37*(1), 25–43.

Xu, Y., & Singh, M. (2002). Early computation of part structure: Evidence from visual search. *Perception & Psychophysics*, *64*(7), 1039–1054.

List of publications

Morgenstern, Y., Hartmann, F., Schmidt, F., Tiedemann, H., Prokott, E., Maiello, G., & Fleming, R. W. (2021). An image-computable model of human visual shape similarity. *PLoS Computational Biology*, *17*(6), e1008981.

Morgenstern, Y., Storrs, K. R., Schmidt, F., Hartmann, F., Tiedemann, H., Wagemans, J., & Fleming, R. (2023). The statistics of natural shapes predict high-level aftereffects in human vision. *BioRxiv*, 2023–01.

Tiedemann, H., Morgenstern, Y., Schmidt, F., & Fleming, R. W. (2022). One-shot generalization in humans revealed through a drawing task. *Elife*, *11*, e75485.

Tiedemann, H., Schmidt, F., & Fleming, R. W. (2022). Superordinate Categorization Based on the Perceptual Organization of Parts. *Brain Sciences*, *12*(5), 667.

I hereby declare that I have prepared the thesis at hand independently and without undue aid or the use of any resources other than indicated within the thesis. All parts of my thesis taken either verbatim or analogously from the published or unpublished works of or based on oral communications with others are indicated as such. Regarding all aspects of my scientific enquiries as they appear in my thesis, I have upheld the tenets of good scientific practice as laid out in the "Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis" and complied with the precept of ethics, data protection and animal welfare. I declare that I have neither directly nor indirectly given monetary or any other valuable considerations to others in connection with the thesis at hand. I declare that I have not presented the thesis at hand, either in an identical or similar form, to an examination office or agency in Germany or any other country as part of any examination or degree. All materials from other sources as well as all works performed by others used or directly referenced within the thesis at hand have been indicated as such. In particular, all persons involved directly or indirectly in the development of the thesis at hand have been named. I agree with the screening of my thesis for plagiarism via offline or online detection-software.

One-shot generalization in humans revealed through a drawing task

Henning Tiedemann[1]*, Yaniv Morgenstern[1,2], Filipp Schmidt[1,3], Roland W Fleming[1,3]

[1]Department of Experimental Psychology, Justus Liebig University Giessen, Giessen, Germany; [2]Laboratory of Experimental Psychology, University of Leuven (KU Leuven), Leuven, Belgium; [3]Center for Mind, Brain and Behavior (CMBB), University of Marburg and Justus Liebig University Giessen, Giessen, Germany

Abstract Humans have the amazing ability to learn new visual concepts from just a single exemplar. How we achieve this remains mysterious. State-of-the-art theories suggest observers rely on internal 'generative models', which not only describe observed objects, but can also synthesize novel variations. However, compelling evidence for generative models in human one-shot learning remains sparse. In most studies, participants merely compare candidate objects created by the experimenters, rather than generating their own ideas. Here, we overcame this key limitation by presenting participants with 2D 'Exemplar' shapes and asking them to draw their own 'Variations' belonging to the same class. The drawings reveal that participants inferred—and synthesized—genuine novel categories that were far more varied than mere copies. Yet, there was striking agreement between participants about which shape features were most distinctive, and these tended to be preserved in the drawn Variations. Indeed, swapping distinctive parts caused objects to swap apparent category. Our findings suggest that internal generative models are key to how humans generalize from single exemplars. When observers see a novel object for the first time, they identify its most distinctive features and infer a generative model of its shape, allowing them to mentally synthesize plausible variants.

***For correspondence:** henning-tiedemann@hotmail.de

Competing interest: The authors declare that no competing interests exist.

Funding: See page 19

Preprinted: 31 May 2021
Received: 11 November 2021
Accepted: 01 May 2022
Published: 10 May 2022

Reviewing Editor: Morgan Barense, University of Toronto, Canada

Editor's evaluation

This paper employs innovative approaches to elegantly tackle the question of how we are able to learn an object category with just a single example, and what features we use to distinguish that category. Through a collection of rigorous experiments and analytical methods, the paper demonstrates people's impressive abilities at rapid category learning and highlights the important role of distinctive features for determining category membership. This paper and its approach will be of interest to those who study learning, memory, and perception, while also contributing to a growing field which uses naturalistic drawing as a window into high-level cognition.

Introduction

Visual recognition and categorization of objects are vital for practically every visual task, from identifying food to locating potential predators. Humans can rapidly classify objects (***Thorpe et al., 1996***; ***Mack and Palmeri, 2015***; ***Serre et al., 2007***), complex scenes (***Wilder et al., 2018***; ***Fei-Fei et al., 2007***), and materials (***Sharan et al., 2010***) into familiar classes, as well as build new classes when presented with sufficient examples (***Op de Beeck and Baker, 2010***; ***Gauthier et al., 1998***). When large numbers of examples are provided, the mathematical basis of human pattern recognition and categorization is well described, and computational models trained with large training sets can emulate many of the human abilities in these areas (***Huang et al., 2017***; ***Krizhevsky et al., 2012***; ***Szegedy et al., 2017***; ***Szegedy et al., 2015***; ***Szegedy et al., 2016***; ***He et al., 2016***; ***Radford, 2021***; ***Kubilius et al., 2016***; ***Jozwik et al., 2018***; ***Jozwik et al., 2016***; ***Jozwik et al., 2019***). However, humans also have the amazing ability to infer new classes when presented with only a small number or even just one single example, as typically occurs when first encountering new concepts ('one-shot categorization'; ***Feldman, 1992***; ***Feldman, 1997***; ***Ons and Wagemans, 2012***; ***Richards et al., 1992***; ***Morgenstern et al., 2019***; ***Ayzenberg and Lourenco, 2021***). Such one-shot learning—which we here use interchangeably with one-shot categorization—is crucial to the formation of human perceptual categories, particularly at early stages of visual development (***Gelman and Markman, 1986***; ***Gelman and Meyer, 2011***; ***Gopnik and Sobel, 2000***; ***Smith and Slone, 2017***; ***Yuan et al., 2020***; ***Gershkoff-Stowe et al., 1997***; ***Landau et al., 1988***; ***Landau et al., 1998***; ***Pereira and Smith, 2009***). Yet how we achieve this remains mysterious and is a significant challenge for artificial learning systems (***Geirhos et al., 2018***; ***Zhang et al., 2019***; ***Zhang et al., 2021***; ***Baker et al., 2018***; ***Michaelis et al., 2020***).

From a computational perspective, the ability to infer a new category from just a single exemplar seems virtually impossible: Given only a single exemplar of a category, there is an infinite number of sets containing that exemplar, any of which could in principle be the true category from which the exemplar was drawn. How can we predict the scope of a category without having witnessed any variability?

The state-of-the-art in psychology and computer science for understanding how humans generalize from few samples suggest that they infer a *generative model*, which considers the observed exemplar as a single sample from a statistical generative process (***Feldman, 1992***; ***Feldman, 1997***; ***Fei-Fei et al., 2006***; ***Goodman et al., 2008a***; ***Goodman et al., 2008b***; ***Lake et al., 2015***; ***Stuhlmuller et al., 2010***). In intuitive terms, it is assumed that observers have a 'deep' understanding of objects, such that they infer the causes or processes that generated the object (***Fleming and Schmidt, 2019***; ***Schmidt et al., 2019***). Importantly, generative models can not only be used to identify behaviourally significant features to judge new samples but can also be used to synthesize or imagine new (i.e., never observed) objects from the same class.

A key step in the inference of a generative model for a given exemplar is the identification of *diagnostic features* that are informative about the underlying generative processes and therefore category membership. Some features—such as the bilaterally symmetric arrangement of animal limbs—are evidence of lawful processes that structure and describe valid members of the category. At the same time, other features of the observed exemplar—such as the specific pose of the limbs—are free to vary across class members. Differentiating between these generic and non-generic (or 'non-transverse'), diagnostic features, such as particular relationships between elements in objects (e.g., symmetry) is one important cue to underlying generative processes (***Feldman, 1997***). Other cue features might include statistical outliers in the shape (e.g., a sharp protrusion, or an angle), that makes a local feature stand out within the shape (***Feldman and Singh, 2005***; ***Feldman and Singh, 2006***; ***Feldman, 2013***; ***Op de Beeck et al., 2008***; ***Kayaert et al., 2005***), or compared to others seen previously. However, how we identify and interpret such category-defining features and use them for generalization remains a matter of debate (***Serre, 2016***).

A stumbling block in testing whether humans infer generative models of objects lies in current experimental methods in category learning. Typically, tasks exploring categorization and generalization ask observers to *discriminate* between multiple presented objects (***Ashby and Maddox, 2005***) (with rare exceptions, e.g., ***Stuhlmuller et al., 2010***) often varying along binary dimensions (e.g., thin vs. thick, square vs. circle; although see ***Hegdé et al., 2008***; ***Kromrey et al., 2010***). Yet, the very process of presenting multiple objects potentially interferes with how observers perform the task. Rather than using a generative model, observers can solve these types of discrimination tasks simply by comparing how similar objects are to one another, without establishing preferences for particular features, and without synthesizing new variants. Another key limitation is that the experimenter—rather than the observer—determines the range of possible options the observers can choose from, thereby crucially constraining the range of responses. Thus, a question of central theoretical importance about whether humans use internal generative models— and if so, how consistent they are across observers—remains unanswered.

To overcome these shortcomings by tapping into generative rather than discriminative processes, we used a task in which observers were presented with single Exemplar objects and were asked to explicitly generate (draw) new objects ('Variations') from the same category, on a tablet computer. With only one Exemplar present, participants could not derive category-defining features by looking for commonalities between category members. Indeed, as their task did not involve comparing objects at all, they could not rely solely on internal discriminative models. Instead, to create new objects, participants were forced to utilize a generative model extracted from the Exemplar, to derive new category members, unless they chose to simply copy the Exemplar with minor deviations, which was discouraged in the instructions. By analyzing the drawings relative to copies of Exemplars, and by asking other participants to (1) categorize drawings, (2) identify their distinctive features, and (3) compare them with Exemplar shapes, we test whether the participants that drew the Variations truly generalized from single exemplars, and determine which features they relied on to do so, giving us insight into what these generative models look like and how similar they are between observers. The strength of unconstrained drawing tasks has been shown in areas such as memory (***Bainbridge et al., 2019***), recognition (***Fan et al., 2020***), efficient representation of both scenes (***Sheng et al., 2021***) and part structure of objects (***Mukherjee et al., 2019***), and developmental changes in children (***Long et al., 2019***), making it an ideal tool to investigate categorization and one-shot learning.

Results and discussion

Systematic generalization in a generative one-shot categorization task (Experiments 1 and 2)

Our first major finding is that participants can synthesize categories of complex objects by generalizing from single exemplars. On each trial, one of eight Exemplar shapes (***Figure 1a***) was presented in the upper half of a tablet computer's screen. Participants were instructed to 'draw a novel object that is not a copy of the Exemplar, yet which belongs to the same category' using a digital pencil (**Experiment 1**, ***Figure 1b***; see Materials and methods). Exemplars were created by the experimenters by combining a main body with varying numbers of parts (ranging from 0 to 5) and displaying a wide range of visual features, such as positive and negative parts (e.g., the indentation in Exemplar 2), curvatures (polygonal and curved shapes), and complexity of parts (e.g., the 'spike' in Exemplar 4 vs. the 'twirl' in Exemplar 6). The aim was to create Exemplars that are diverse without resembling real-world objects, to maximize generalizations as well as limit the impact of semantic knowledge. While hand-crafting stimuli runs the risk of introducing biases, the intended variety of Exemplars made an algorithmic creation approach unfeasible. For each of the 8 Exemplars, 17 participants ('drawers') each drew 12 Variations (yielding a total of 204 Variations per Exemplar, i.e., 1632 drawings overall). To minimize potential carry-over effects of previously seen Exemplars, shapes were shown in randomized order for each participant. As a baseline, another group of participants (n = 15) was asked to copy the Exemplars three times as accurately as possible (**Experiment 2b**).

To assess the range and variety of the generated shapes, a new group of 12 participants rated the similarity of each Variation to the corresponding Exemplar (**Experiment 2**, note that number of data points per participant vary by experiment, for details see Materials and methods). ***Figure 1c*** shows examples of Variations for one Exemplar, with perceptually more similar drawings plotted closer to the Exemplar (see ***Figure 1—figure supplement 1*** for Variations of other Exemplars). The drawings illustrate the range of responses, with some shapes being very similar to the Exemplar while others differ considerably. ***Figure 1d*** shows the average similarity of produced Variations for each participant, demonstrating that participants varied substantially in how much their drawings deviated from the Exemplar, that is, in their creativity when producing Variations. Linear regression showed that age of drawing (i.e., whether a shape was drawn earlier or later within the sequence of 12 shapes for each Exemplar), did not predict similarity ($R^2 < 0.01$, $F(1,10) = 0.08$, p = 0.77). Individual participants tended to stay within a constrained similarity spectrum with only small trends in the direction of less or more similar shapes over time with slopes ranging from −0.13 to 0.12 (mean = 0.07).

To test whether the generated shapes were genuinely new objects or just slightly altered copies of the Exemplars, the copies were compared to a subset of Variations using a 2-AFC task (**Experiment 2b**, see Materials and methods for details), in which a new group of participants (n = 15) was asked to pick the shape that looked more like a copy of the Exemplar. In 95% of trials (chance = 50%) the copies were chosen, showing that the vast majority of a representative subset of Variations were perceived

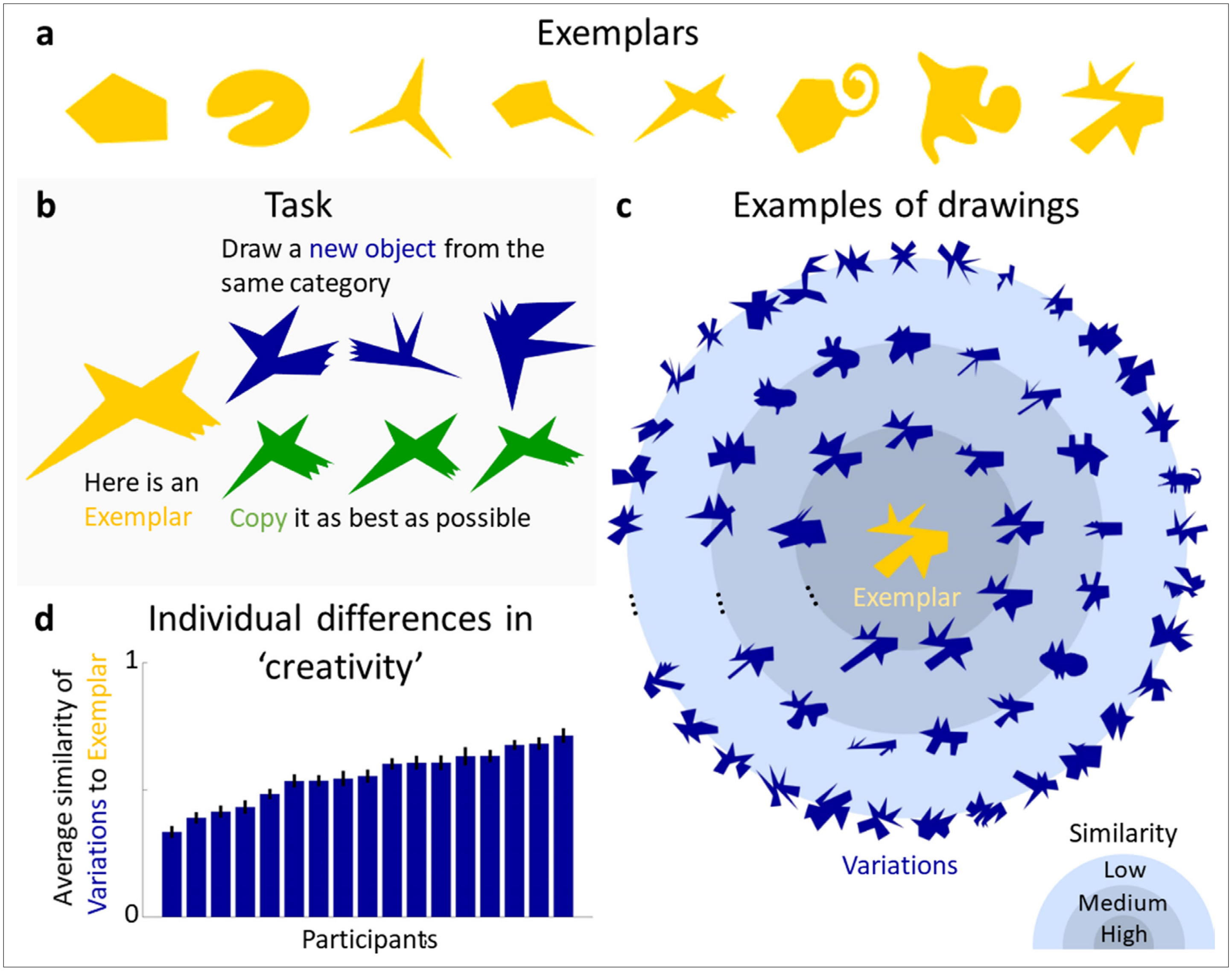

Figure 1. Generative one-shot categorization task and results. (**a**) Exemplars presented in the Experiments. (**b**) In the task, a group of participants was presented with an Exemplar and asked to draw new objects from the same category (blue). As a measure of baseline performance, another group of participants was asked to copy the Exemplar as accurately as possible (green). (**c**) Examples of drawn Variations (blue) generated in response to one of the Exemplars (yellow). Variations are plotted according to their perceptual similarity to the Exemplar, with more similar Variations closer to the centre. (**d**) Individual differences in 'creativity', defined by the average perceived similarity of participant's drawings to the respective Exemplar. Error bars indicate standard errors.

The online version of this article includes the following figure supplement(s) for figure 1:

Figure supplement 1. A subset of Variations (dark blue) created for each Exemplar (yellow, centre).

to be more different from the Exemplars than a mere copy (one-sided binomial test: 50%; $N = 16{,}200$, i.e., number of judgements; K (correct responses) = 15,388, $p < 0.001$).

Responses represent distinct perceptual categories (Experiment 3)

Next, we tested whether the drawings represented distinct perceptual object categories or were mere random variations that did not form coherent groups. The similarity ratings for each Exemplar's Variations from **Experiment 2** were split into subsets (40 bins) that spanned the full similarity continuum (see Materials and methods). A new group of participants ($n = 15$) classified one Variation from each bin for each Exemplar (320 stimuli in total), by sorting in each trial a randomly selected Variation into one of the eight Exemplars' categories. The average percentage of correct

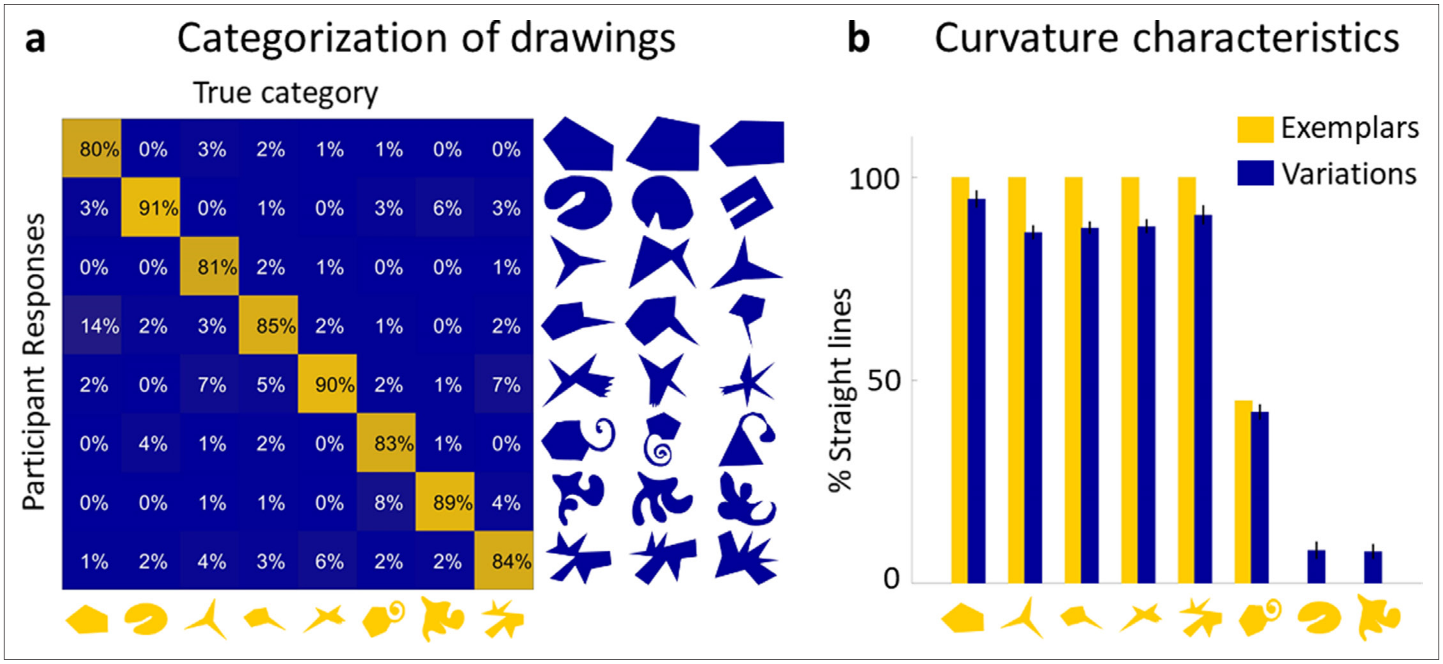

Figure 2. Drawings constitute a real perceptual category. (**a**) The confusion matrix for true classes and participant responses shows that the vast majority of Variations were classified correctly. A subset of Variations that had to be categorized is shown on the right of the matrix with the Exemplars on the bottom. (**b**) Variations largely reproduce the global curvature characteristics of the Exemplar. Curvature similarity across Variations and their Exemplars: Exemplars (yellow) ordered by percentage of perimeter comprised of straight lines, together with average percentage of straight lines in all Variations of that category (blue). Error bars indicate standard error.

classifications was high (86%) and well above chance (one-sided binomial test: 12.5%; $N = 4800$, i.e., number of judgements; K (correct responses) = 4111, $p < 0.001$). ***Figure 2a*** shows the confusion matrix for true classes and participant responses. In almost all instances, observers sorted the Variations correctly into the originating class. Except for a single cell (row 4, column 1), all cells are significantly different from chance with all diagonal cells above—and all others below—chance (one-sided binomial tests with Bonferroni-adjusted p value of 0.015, according to the eight possible outcomes in each row). To investigate whether performance in this task was influenced by the similarity between Variations and Exemplars, the Variations were divided into four similarity bins ranging from very similar to very dissimilar. Performance in the three bins most similar to the Exemplars was virtually identical (overall 89%, 87%, and 88% correct responses), with a maximum of 14% misclassifications in single cells. For the least similar bin, performance dropped to 79%, suggesting that highly dissimilar shapes were significantly more often misclassified, while overall classification accuracy was still way above chance. The maximum of misclassification here was 28% in a single cell (row 4, column 1). This shows that there are almost no systematic miscategorizations between Exemplar categories and demonstrates that the Variations produced in the generative task are samples of robust perceptual categories. Specifically, our results tend to suggest that drawers identified diagnostic features in the Exemplars and reproduced them in the Variations, allowing other observers to identify them as belonging to a common class. Thus, drawers effectively derived distinct novel categories from just a single object.

Identifying category-defining features

So far, we have shown that drawers generated genuinely novel objects, which other observers could nevertheless assign to their corresponding category. This suggests that drawers identified and reproduced those features of the object that were most diagnostic of the class. Thus, we next investigated which features were preserved, the extent participants agreed about the most significant features, and the importance of these features for category assignment.

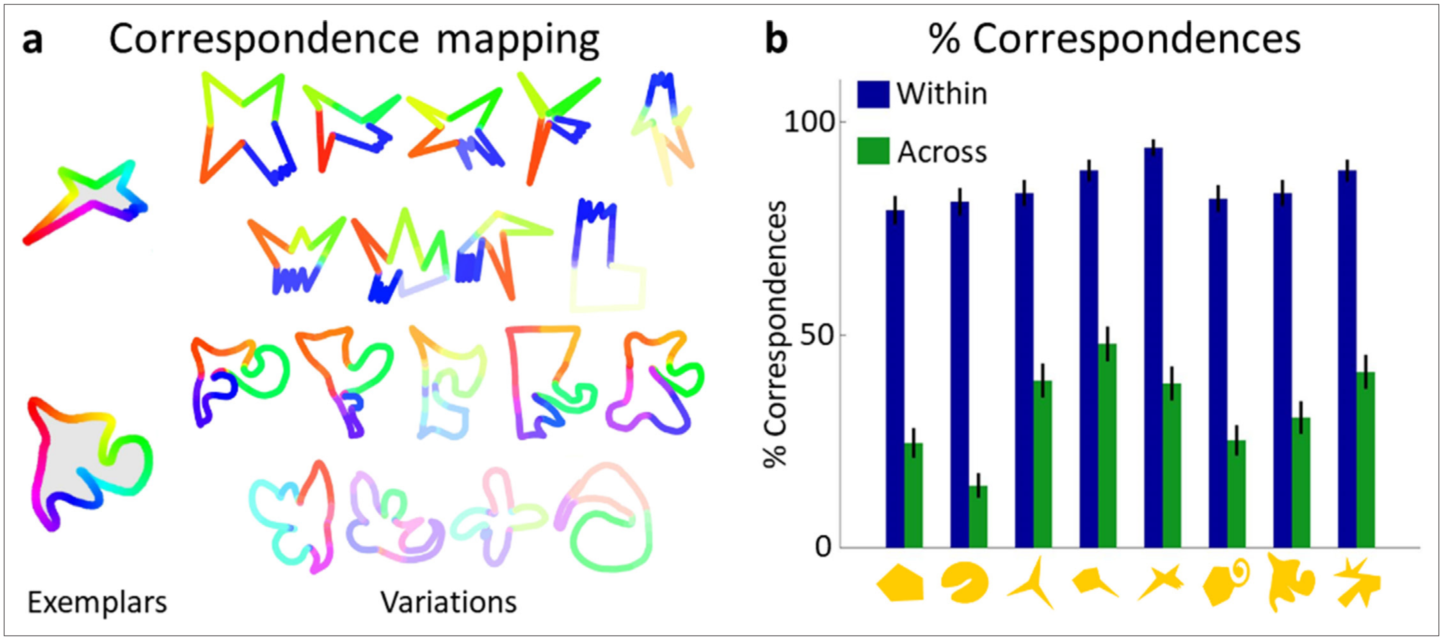

Figure 3. Variations and Exemplars of the same category share most of their parts. (**a**) Aggregated mapping of correspondences for two categories, showing high agreement between participants about corresponding parts. Colours are explained in main text. (**b**) Comparison of percentage of correspondences seen within (blue) and across (green) categories, showing that significantly more correspondence was perceived within categories. Error bars indicate standard error.

The online version of this article includes the following figure supplement(s) for figure 3:

Figure supplement 1. Correspondence data for whole subset.

Figure supplement 2. Analysing part order and part changes in the correspondence task (Experiment 4).

Global curvature features

One significant feature preserved in almost all Variations was the global curvature, namely whether the Exemplar consisted of straight or curved lines. We find that polygonal Exemplars tended to lead to polygonal Variations while curvaceous Exemplars led to curvaceous Variations ($r = 0.99$, $p < 0.001$; ***Figure 2b***). This finding is broadly in line with the concept of non-random features indicating generative processes (***Feldman, 1992***; ***Feldman, 1997***): A pencil tracing a random walk is unlikely to draw a straight line, so straight lines are considered evidence of a significant (i.e., non-random) generative process, which are therefore preserved in Variations.

Part-related features (Experiment 4)

Considering the Variations in ***Figure 1c***, another significant feature that seems to be approximately preserved in many drawings is the part structure—that is, arrangement and number of parts—of the Exemplars (this is especially salient when looking at the most similar Variations; see also **Ideas and Speculations**). At the same time, the parts themselves were often modified in size, orientation, or elongation. This signifies a potential strategy that participants might have used in the drawing task, highlighting their generative approach: Starting from perceptual segmentation of the Exemplar into parts, they may have modified these parts (to varying degrees) and 'put them back together', either in the original or a different order.

To test whether participants used this strategy, another group of participants ($n = 15$) was shown one Exemplar at a time, paired with a Variation of either the same or a different category, and asked to identify any corresponding parts between the two shapes. This allowed us to test whether part correspondence is stronger within categories than across categories—and how much of the Exemplar's part structure was retained in its Variations. When participants perceived parts as corresponding, they delineated each part by drawing a line, thereby 'cutting' the shape into two parts on either side of the line, and then choosing one of the parts to indicate the correspondence to a part in the Exemplar.

Figure 3a shows aggregated mappings of correspondence for two of the categories, with the Exemplar on the left and rows of corresponding Variations in descending similarity to the Exemplar (see ***Figure 3—figure supplement 1b*** for complete set). Each point of the Variations' contour is coloured the same as the Exemplar's contour to which it was perceived to correspond the most. If a point was perceived to correspond to a whole section of the Exemplar (e.g., the green 'nose' of the Exemplar in the bottom row of ***Figure 3a***), then it was coloured the same as the circular median point of that section. Colour saturation indicates how often correspondence was seen for each point, with higher saturation indicating stronger correspondence. This visualization shows the generally high level of agreement about corresponding parts between participants. ***Figure 3b*** compares the correspondence in same-category pairs to that in different-category pairs. On average, 71% of an Exemplar's area corresponded to some part of its Variations, with only 18% correspondence to Variations of other categories (one-sided *t*-test: $t(1199) = 38.84$, $p < 0.001$, ***Figure 3—figure supplement 1a***). To test whether less similar shapes share fewer parts, we analyzed the corresponding area as a function of the similarity of Variations to Exemplars by linear regression. We find that the extent to which the Variation was seen to correspond was well predicted by similarity ($R^2 = 0.92$, $F(1,8) = 96.69$, $p < 0.001$, see ***Figure 3—figure supplement 1a***). This also suggests that the extent to which parts or area corresponded might have been used by observers in **Experiment 2** to assess similarity.

A significant correlation between the percentage of correct categorizations from **Experiment 3** and the percentage of corresponding area ($r = 0.53$, $p < 0.01$) indicates that Variations that share more parts with the Exemplar are grouped more often with their Exemplar compared to those sharing fewer parts. Notably, out of the 80 Variations used in **Experiment 4**, 49 were always classified correctly, with a range in corresponding area between 30% and 100%, and a mean of 81%.

Visual inspection of Variations suggested that Exemplars' part orderings were often retained in Variations, that is, visually different corresponding parts were arranged in the same sequence. However, in some Variations this order was changed, as if the parts were 'shuffled'. To test the prevalence of different part orderings in the Variations, we created a compressed representation of part order for each shape pair's correspondences (***Figure 3—figure supplement 2***): Moving clockwise around the shape's silhouette, we listed each part indicated by the participant, as well as gaps with no correspondences, resulting in a 'part circle' describing the order of labelled parts for each shape (see ***Figure 3—figure supplement 2a***). This representation was chosen because Exemplars were created to comprise a main body with parts arranged roughly circularly around it, making the resulting part circles approximately analogous to perceived part ordering. To analyze part order, part circles were then compared to quantify the extent to which orders were identical, reversed, or shuffled (i.e., non-identical). There is an important limitation to this analysis, however. By this definition, part order could only be changed with more than two corresponding parts, amounting to 30% of all trials. For the remaining 70% of trials, this analysis is not possible as with two or fewer corresponding parts, part order is automatically conserved or cannot be defined. This relatively low number is explained by participants often aggregating corresponding parts in their responses, for example, by considering the three 'non-bitten' spikes in the top Exemplar in ***Figure 3a*** as a single corresponding part—resulting in only two correspondences even though the shape may have featured more perceptual parts. Of those trials with more than two correspondences, part order was identical in 77%, reversed in 7%, and shuffled in 17% of cases. Analyzing gaps in part circles also provided a tentative measure of how often parts of the Exemplar were omitted (4%), substituted with another non-corresponding part (25%), or new parts were added that did not exist in the Exemplar (8%). Notably, more part substitutions occurred for less similar Variations to the Exemplar (***Figure 3—figure supplement 2d***). For further details of this analysis, see Materials and methods.

While we have to be careful to generalize to the full set of Variations from this limited number of data points, this analysis nevertheless provides some insight into potential strategies used in the creation of new shapes: Many Variations retained the Exemplar's part order, sometimes in reversed form, suggesting that the shape was changed from its original form (as opposed to being created completely from scratch) either globally or part by part to create an appreciably different object displaying the same part order. In contrast, Variations with a shuffled part order, or omitted and added parts, point to a 'building block' approach where parts are treated as independent, re-combinable elements.

Together, these findings demonstrate that Exemplars and associated Variations share a considerable portion of their parts, indicating that drawers preserve identifiable parts in their Variations even though they varied specific geometrical properties of these parts. Furthermore, Variations also tended to preserve the order of corresponding parts—with some exceptions where part order was changed—overall pointing to a highly part-based creation approach.

Identification and preservation of distinctive parts (Experiment 5)

Another notable feature of the generated shape categories is that certain parts of the Exemplars seem to be more distinctive than others, and that these often also appear in the Variations. We suggest that these distinctive parts are a major driving force for correct categorization (**Experiment 3**). To address whether participants agreed about which shape features are most distinctive, we showed a new group of participants (n = 10) all 8 Exemplars together with 39 Variations from each of the categories, in random order. They were asked to mark up to three parts of each object's silhouette through a painting interface, starting with the most distinctive part, followed by the second and third most distinctive parts. ***Figure 4a*** shows all responses for one shape, together with the aggregated response, which indicates a high level of agreement across observers (see ***Figure 4—figure supplement 1*** for the complete set). ***Figure 4b*** shows aggregated responses for a subset of Variations per category. Contrasting these data with randomized responses mimicking the number of areas painted and the lengths of consecutive areas painted, but with randomized placement of those areas (see Methods for details), shows significantly higher agreement between humans compared to chance: The mean consecutive area of a shapes' silhouette with a high distinctiveness score (above 75% of the highest score) for the human data made up 19% of a shape's perimeter compared to <1% for the randomized data (two-sample Kolmogorov–Smirnov test, $p < 0.001$). This high agreement among observers is especially noteworthy given that the quite vague concept of 'distinctiveness' could be interpreted differently by different observers, pointing to just how important and characteristic these parts are considered compared to the rest of the shape.

Visual inspection suggests that for most categories, participants tend to consistently indicate specific parts as being the most distinctive across different Variations (e.g., indentation for category 2, spike for category 4, jagged feature for category 5, and twirl for category 6). Cross-referencing these data with the part correspondences from **Experiment 4** allowed us to test whether distinctive parts within a category are indeed conserved (i.e., whether the most distinctive parts in each shape are those that are judged to correspond to each other). For each corresponding part pair, average distinctiveness scores of both parts were calculated. Comparing the scores of all Exemplar parts with the scores of all Variation parts of the same category shows a high correlation ($r = 0.63$; $p < 0.001$), suggesting that distinctive parts remain distinctive even when modified or shuffled within the overall part structure. Equivalently, indistinctive parts in Exemplars tend to remain indistinctive in Variations, lending further support to the finding that participants modify individual parts to create new shapes. This raises the possibility that parts that are perceived to be distinctive are particularly important in determining category membership.

Causal role of distinctive parts in categorization (Experiment 6)

We next sought to test the impact of distinctive parts on category membership more directly. To do this, we created new stimuli from Variations by replacing the most distinctive part—as determined from **Experiment 5**—with the most distinctive part of the Exemplar of another category. For comparison, we created stimuli where we replaced the same Variation's least distinctive part with the least distinctive part of the Exemplar of another category (***Figure 5a***; controlling for perimeter length—see Materials and methods). A new group of participants (n = 15) grouped each of these newly generated stimuli (280 with the distinctive, 280 with the indistinctive part swapped) with one of the Exemplars, similar to **Experiment 2**.

Figure 5b compares the results with the percentage of correct categorizations in **Experiment 2**. Binomial tests indicate that the proportion of correct responses from **Experiment 2** (86%) was significantly higher than both swap conditions (69% for indistinctive swap, $p < 0.001$, and 32% for distinctive swap, $p < 0.001$). Further, we find a large difference in effect sizes: while swapping an indistinctive part does have a small effect (Cohen's $h = 0.29$), swapping a distinctive part has an immense effect on correct categorizations (Cohen's $h = 1.05$). This shows that distinctive parts were a driving force

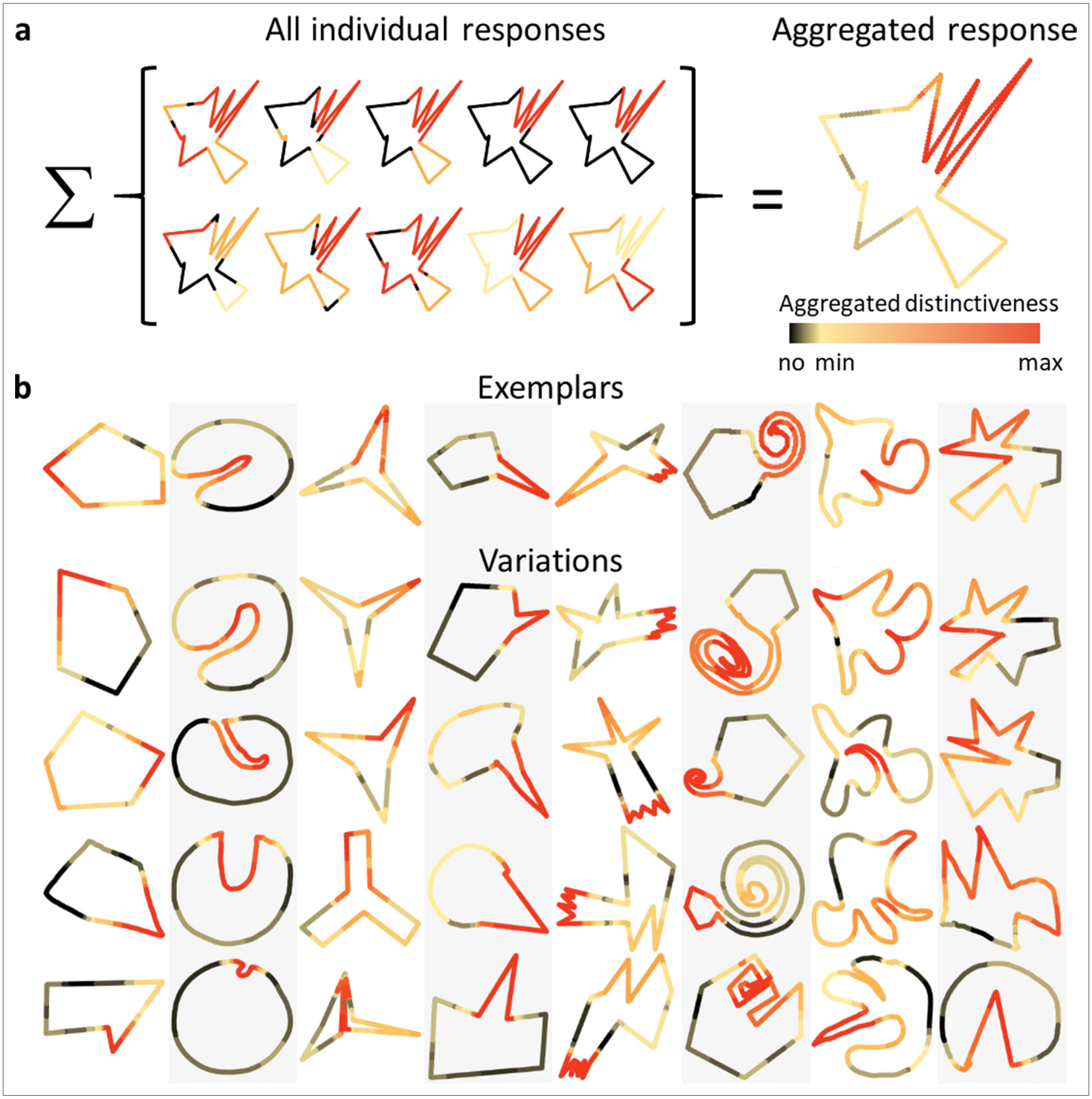

Figure 4. Observers agree on the most distinctive parts. (**a**) Individual responses of all participants for an example Variation, showing which parts were marked as most distinctive (red), second most distinctive (orange), and third most distinctive (yellow). Aggregating these responses results in the shape on the right. The redder each point in the contour, the more distinctive it was evaluated across all participants. (**b**) Comparing aggregated responses between Exemplars and corresponding Variations (a subset shown here) suggests that in most categories similar parts (e.g., the indentation, spike or twirl) were identified as distinctive across most Variations.

The online version of this article includes the following figure supplement(s) for figure 4:

Figure supplement 1. Aggregated distinctiveness scores for all shapes.

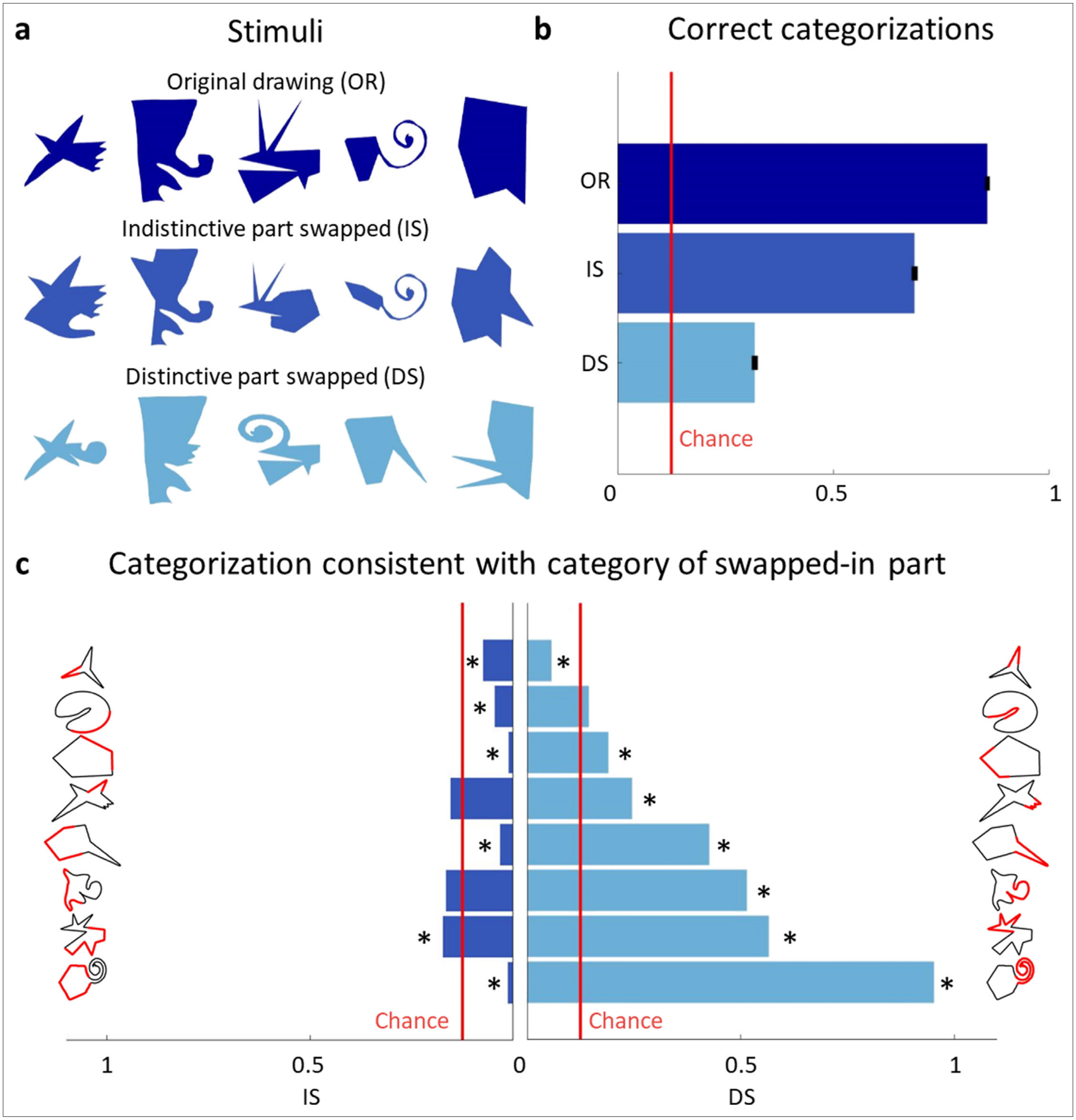

Figure 5. Distinctive parts are the main categorization cue. (**a**) Original drawings (top row) were altered so that either the least distinctive part was swapped with the least distinctive part of a different category (second row) or the most distinctive part was swapped with the most distinctive part of another category (third row). (**b**) Comparison of the percentage of correct categorizations for the original shapes (from Experiment 2), swapped indistinctive parts and swapped distinctive parts (error bars indicate standard errors). (**c**) Bar plots showing how often the category of the swapped-in part determined the categorization choice. The indistinctive (left) and distinctive (right) parts are shown in red on the shapes' silhouette. Bars significantly different from chance ($p < 0.01$) are marked with an asterisk.

in categorization decisions, even though they often comprised only a small percentage of shapes' contours. ***Figure 5c*** summarizes how often the category of the new, swapped-in part determined the categorization decision, separately for indistinctive and distinctive parts. Performing one-sided binomial tests on each of swapped-part conditions (with a Bonferroni-adjusted p level for 16 tests) shows that most distinctive parts strongly biased participant's responses toward their categories. In contrast, only one indistinctive part had a significantly positive impact on its category choice, while most indistinctive parts resulted in choices of their category even significantly below chance.

Previous studies showed that small-to-intermediate fragments of an image are sufficient for correct categorization (***Hegdé et al., 2008***; ***Ullman et al., 2002***). These informative fragments were defined by implicit analysis of the statistical distribution of features across the complete object category or categories (i.e., the fragments were learned during a training phase). In contrast, in our experiments, parts were identified as distinctive from just a single piece of data (i.e., one object), consistently across many observers. We suggest that observers may use statistical outliers within shapes (***McCarthy and Warrington, 1986***), or outliers relative to previously seen objects to identify distinctive parts.

In summary, observers agreed on the most distinctive parts of shapes and used this information as one of the main cues for category membership. In line with this, when creating new shapes, these distinctive parts were reproduced with modifications that retained their specific characteristics making them both distinctive and signifiers of their category.

General discussion

Our ability to organize objects into categories at a glance is of fundamental importance to everyday activities. It allows us to access a wealth of knowledge about objects from previous experiences, rather than having to discover each newly encountered item's properties de novo. Visual object categorization is typically so effortless that we easily take it for granted. Yet disorders of object perception—such as visual agnosia (***McCarthy and Warrington, 1986***; ***Riddoch and Humphreys, 1987***; ***Goodale et al., 1991***; ***Behrmann and Nishimura, 2010***) —lead to profound deficits. A particularly striking characteristic of healthy human object perception is how rapidly observers learn new categories. Whereas machine-learning systems typically require thousands of examples per category to rival human performance at recognizing objects in photographs (***Huang et al., 2017***; ***Krizhevsky et al., 2012***; ***Szegedy et al., 2017***; ***Szegedy et al., 2015***; ***Szegedy et al., 2016***; ***He et al., 2016***; ***Radford, 2021***; ***Kubilius et al., 2016***; ***Jozwik et al., 2018***; ***Jozwik et al., 2016***; ***Jozwik et al., 2019***), human infants and adults appear to be able to generalize successfully from just a single example of a new category (***Gelman and Markman, 1986***; ***Gelman and Meyer, 2011***; ***Gopnik and Sobel, 2000***; ***Smith and Slone, 2017***; ***Yuan et al., 2020***; ***Gershkoff-Stowe et al., 1997***; ***Landau et al., 1988***; ***Landau et al., 1998***; ***Pereira and Smith, 2009***)—so-called 'one-shot learning' (***Feldman, 1992***; ***Feldman, 1997***; ***Ons and Wagemans, 2012***; ***Richards et al., 1992***; ***Morgenstern et al., 2019***; ***Ayzenberg and Lourenco, 2021***).

One-shot categorization is a formally under-constrained inference problem (***Feldman, 1992***). There are infinitely many sets containing any given exemplar, any of which could be the true category. It is thus remarkable that humans seem to be able to draw consistent conclusions about category membership from such sparse data (***Feldman, 1992***; ***Feldman, 1997***). Their judgements presumably reflect assumptions about how—and by how much—objects within a category tend to differ from one another, which constrains the space of variants that are deemed likely. Yet how this occurs remains elusive.

We suggest that generative statistical models of shape play a key role in one-shot categorization. However, to date, direct evidence for generative models has been scant. Most experimental studies on categorization use sets of pre-selected stimuli to define categories and *discriminative* tasks to test their hypotheses. The main drawback of this approach is that the assumed category-defining features are determined by the experimenter and might therefore differ from the features used in unrestricted categorization decisions. Moreover, discriminative tasks allow for simple strategies based on comparisons between objects, rather than probing the visual system's internal generative models of objects which are thought to be central to how humans and machines can learn categories from sparse data (e.g., ***Fleming and Schmidt, 2019***).

Here, by contrast, we used a generative one-shot categorization task to tap directly into generative models and human creativity—and to identify category-defining features—by allowing participants to generate their own new category samples, rather than merely selecting between experimenter-generated alternatives. The resulting Variations are thus shaped by the features that participants consider important in the context of their previous experiences and by their internal visual imagery processes. In principle, different observers might consider different features as significant so that across observers no coherent object categories would emerge. However, in our findings this was clearly not the case. We found a high degree of agreement between observers, suggesting general principles of how humans analyze single objects and extrapolate new category members from its features. Participants created Variations for different Exemplars varying in curvature characteristics and number of parts, resulting in a large dataset of over 1600 drawings, which were significantly more variable than mere copies of the Exemplars. Yet, despite the wide variety of Variations, the overwhelming majority of a representative subset of these shapes was correctly grouped with the originating Exemplar by independent observers, showing that our task yielded genuine distinct perceptual categories.

There are, of course, some limitations to our approach. By asking participants to draw variants, the responses are necessarily limited by participant's skills with the digital drawing interface. Some people have lesser confidence or ability in drawing, and thus it is possible that some of drawings might not accurately reflect their mental image of a given variant. However, this is somewhat mitigated by the relatively straightforward 2D forms they were asked to draw. No ability to render perspective, shading or other more advanced artistic skills were required. It is also worth noting that there was a substantial difference between copies (which were quite accurate) and variants (which were highly diverse). This suggests that artistic ability did not prevent participants from depicting at least a subset of the varied and novel instances brought to mind by the exemplar.

Another potential bias introduced by our approach is possible carry-over effects of Exemplars seen in previous trials influencing the drawings of later ones. While we aimed to minimize these effects by randomizing the order of Exemplars, future studies could reduce this further by presenting just a single Exemplar per participant, for example in an online experiment with larger numbers of participants.

A third potential issue with the method was our use of hand-created experimental stimuli. The presence of distinctive features in some of the stimuli (e.g., the 'twirl' in Exemplar 6) was therefore not accidental but the result of a decision by the experimenter, and it could be that such features occur less frequently in natural objects. This limitation could be mitigated in future experiments by creating stimuli through a stochastic process that is less under experimenter control. For example, by training a generative algorithm (e.g., Generative Adversarial Networks ***Goodfellow et al., 2014***) on a dataset of natural shapes, it is possible to create stimuli that share statistical properties with natural shapes

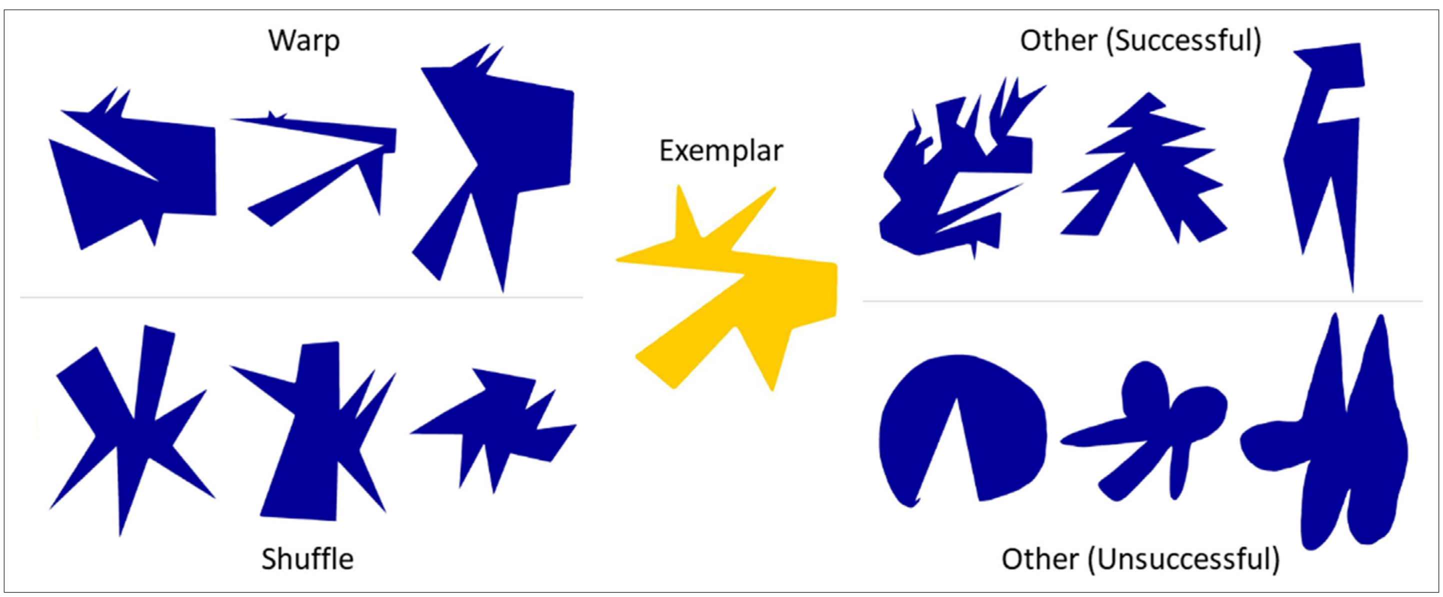

Figure 6. Examples of proposed strategies used by participants. Selected examples representing part-based strategies like warp, shuffle, and other strategies, both successful (i.e., shapes were correctly classified almost all the time) and unsuccessful strategies (i.e., shapes were correctly classified far below average).

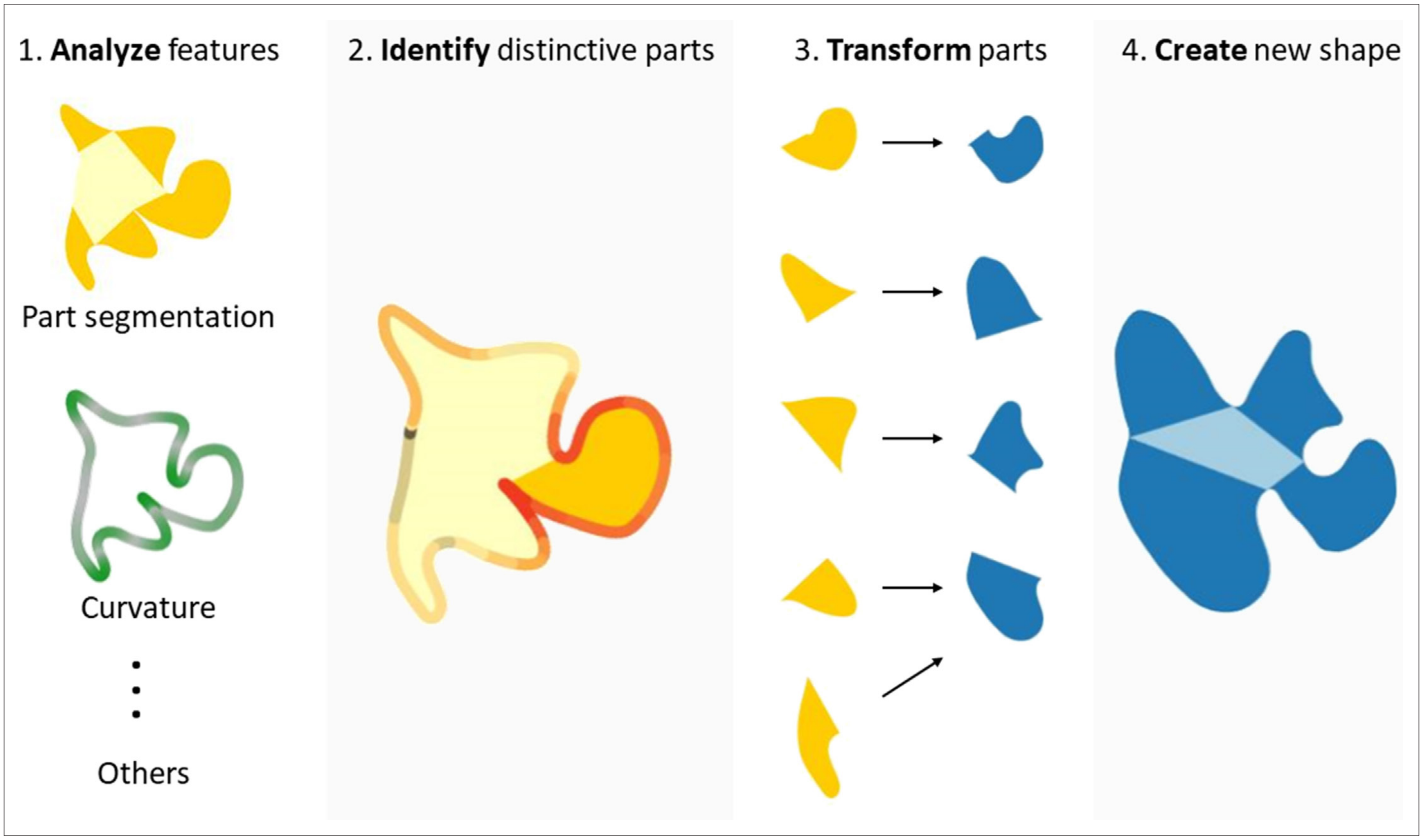

Figure 7. Proposed steps of shape creation using a predominantly part-based approach. After analyzing shape features and identifying distinctive parts, the individual parts get transformed and re-assembled to form a new shape. In the transformation process, distinctive parts are changed to a lesser extent than other, non-distinctive parts. The creation of a new shape from the individual parts might involve adding, merging, or removing parts, as well as changing their order. Note that this is an illustration of only one strategy (albeit an important one); however, Variations might also be created using non-part-based strategies, where other features such as curvature are varied.

(***Morgenstern et al., 2021***), although care would need to be taken to rule out stimuli that resemble recognizable items (which often result from such methods). However, even algorithmic stimulus generation does not fully solve the problem. Through their decisions about the algorithm and training set, experimenters necessarily exert some degree of control over the kinds of stimuli presented to participants, even if they do not hand-create or hand-select the individual instances, as we did. Despite our use of hand-drawn Exemplars, the finding that participants could consistently identify distinctive features without any explicit training—and treated these as defining class membership—suggests that both experimenter and participants shared assumptions about which features are important. More broadly, the finding that participants created Variations that other participants recognized as belonging to the same class provides an existence proof for systematic generative processes, irrespective of how frequently such distinctive features occur in natural objects.

Ideas and speculation

How might a generative model operate in practice to allow a participant to synthesize new objects? While it is difficult to probe the exact strategies used to create each Variation, visual inspection of the drawings, as well as the part correspondence task in **Experiment 4,** provide some tentative insights. We summarize our speculations about some of the possible strategies in ***Figure 6***. For example, we call one strategy 'Warping', as many variants appear to be related to the original Exemplar by a relatively straightforward (non-linear) spatial distortion operation that preserves the ordering of points, but alters their relative distances. Through such warping, a wide variety of shapes can be synthesized while largely retaining one-to-one correspondence with features of the prototype from which they are derived.

Some strategies seem to involve segmenting the Exemplar into parts, which are then modified to various degrees and recombined to create a Variation (***Figure 7***). There is considerable evidence for perceptual organization processes that segment shapes into distinct subcomponents based on their geometrical properties (***Hoffman and Richards, 1984***; ***Hoffman and Singh, 1997***; ***Siddiqi and Kimia, 1995***; ***Biederman, 1987***; ***Singh et al., 1999***; ***Marr and Nishihara, 1978***) and causal origin (***Spröte et al., 2016***). It thus seems intuitively plausible that parts might play a significant role in the generative process too. This idea is further supported by the following observations: (1) Variations' parts often corresponded to parts of the Exemplar (**Experiment 4**); (2) part structure was retained or merely shuffled in a significant number of shapes (**Experiment 4**); and (3) distinctive parts were crucial for categorization (**Experiment 6**). As a result, even highly dissimilar shapes still shared 34% of their area with the originating Exemplar (see ***Figure 3—figure supplement 1a*** right panel). Importantly, some parts were identified as being highly characteristic of the Exemplar's category and accordingly were varied in the Variations in ways to remain category defining (**Experiment 6**).

In addition to reusing and changing individual parts, the order of parts also seems to have been retained in many Variations—showing that drawers relied on both individual parts as well as their relationships to create a new object. Indeed, even shapes that appear to be created by a straightforward 'Warp' strategy are not inconsistent with a part-based analysis and recombination.

In other Variations, however, part order was shuffled, illustrating a local strategy in which parts were considered independent and recombined in different order. This part-based approach is also reflected by the omission or addition of parts to the Exemplar, as revealed by our 'part circle' analysis. One limitation of the analyses supporting these suggestions is that in **Experiment 4**, observers often (70% of trials) indicated only two or fewer corresponding parts. For these shapes, it is not possible to identify part-order relationships using the 'part circle' analysis. This was in part because participants often appear to have bundled together multiple elements—which a part-detection algorithm would treat as discrete—into composite shape elements, effectively processing the parts in concert. Moreover, although the drawings themselves were entirely unconstrained, it should be noted that **Experiment 4** was explicitly focussed on corresponding parts by design, which only allows us to assess part-based strategies.

Visual inspection of the dataset also suggests additional strategies that are harder to quantify and less decidedly part focussed. Some of these are shown in ***Figure 6*** (top right), which seem to change the Exemplar's part structure substantially while nevertheless retaining and exaggerating the (statistical) motif of 'spikiness'. The exact nature of the features that define abstract concepts such as 'spikiness'—and thus allow observers to relate the Variations to the Exemplar—remain elusive. However, presumably it is necessary to retain certain statistical properties of the shape, such as the prevalence of sharp angles and straight edges. The bottom right panel shows Variations which were less often linked to the original category in **Experiment 3** than the average Variation. Very few if any of the parts in the Exemplar are retained, and curved elements abound despite the original shape being entirely angular. Evidently, eliminating core features—such as parts or straightness—was not a successful strategy for generating legal Variations. Consequently, we suggest that observers analyze the Exemplar with respect to a number of features, ranging from basic properties like curvature to more high-level properties like 'spikiness', and then vary one or multiple of these features to create a new object. This is consistent with recent work showing that perceived similarity between shapes can be predicted by considering a large number of shape features (***Morgenstern et al., 2021***). Yet, in general, observers seemed to share assumptions about the extent to which features could be varied while still retaining the categories' identity, as indicated by the small number of misclassifications in **Experiment 3**. It is interesting to speculate that these assumptions may therefore be derived from the statistics of variations between items in the natural world.

Many of the generated shapes not only shared most of their parts, but, strikingly, specific parts were also reliably perceived as more distinctive than others and were the main catalyst driving categorization decisions, even if they comprised only a minority of the shapes' area. This suggests that a crucial stage in one-shot categorization is the identification of those features within an object that are most likely to be 'distinctive' for defining the category.

What is the basis of such inferences? Given only a single exemplar, how are the 'most distinctive' features determined? We suggest that observers' decisions are related to processes that identify signatures of the underlying *generative processes* responsible for creating the observed shape and which thus define the category. This is closely related to Feldman's (***Feldman, 1992***; ***Feldman, 1997***) theory of non-accidental features. Statistical relations that are unlikely to occur under a random model (e.g., collinearity of features) are evidence of the operation of non-random generative processes. Similarly, parts that have statistically distinct properties from the rest of the shape (***Schmidt et al., 2019***)—or which are statistically distinct from parts of objects seen previously (***Lake et al., 2015***)—are likely evidence of a category-defining process. Consistent with this idea, we find that parts that were deemed distinctive were also those that had the most significant effect of categorization (e.g., the ‘twirl’ feature in category 6; see ***Figure 1a***). On a more global scale, curvature characteristics (straight or curved lines) of Exemplars were preserved almost perfectly in most drawings, showing that these features were also deemed non-random. Conversely, random features were altered more freely: Since the presented shapes were abstract, the part arrangement held no meaning and was therefore free to be modified (in contrast to familiar objects like a chair or a human), resulting in some Variations with changed part-ordering from the Exemplar. Yet an important open question is which properties of parts or features are used to determine their status as outliers, and how these properties are determined, based on familiar objects and processing constraints of the visual system. For example, even though distinctiveness is an intuitive concept—as demonstrated by the high agreement by observers in **Experiment 5**—it is challenging to define it formally. A large number of factors can affect the distinctiveness of a part or feature, ranging from purely geometric qualities such as frequency of curvature changes (***Baker et al., 2021***; ***Attneave, 1957***; e.g., a tentacle vs. a stub) to comparative analyses within the object (difference between part and other object parts) or across other objects seen before (difference between part and other objects’ parts). More generally, parts might be compared to an internal taxonomy of parts, each weighted with its probability of appearance, potentially informed by frequencies of parts in real-world objects or by their function (***Mukherjee et al., 2019***; ***Tversky, 1989***). Another important open question is how parts are parameterized so that identity-preserving variations can be generated. Skeletal representations (e.g., ***Feldman and Singh, 2006***) offer a promising avenue for potential representations (***Destler et al., 2019***; ***Wilder et al., 2011***; ***Wilder et al., 2018***; ***Wilder et al., 2019***).

Conclusions and future directions

Taken together, the results of our experiments suggest that humans are not merely passive observers, who assign objects to categories based on their relative similarities in a fixed feature space. Instead, a key aspect of human generalization is our ability to identify important signatures of generative processes and then to run internal routines that actively synthesize new objects that we have never actually observed. Thus, although drawing tasks present experimenters with challenges—particularly in terms of analyzing the resulting drawings—they also provide a promising avenue into research not only about object categorization but also human creativity. As seen in ***Figure 1d***, there is a notable difference between participants as to how much they tended to deviate from the Exemplars, raising the question of what drives these individual differences. Another particularly fascinating open question is the extent to which the synthesis of novel objects by observers recruits physical simulation processes (***Battaglia et al., 2013***). It seems plausible that experience with the natural ways that objects, materials, animals, and plants move and change shape over time (***Schmidt and Fleming, 2016a***; ***Spröte et al., 2016***; ***Schmidt et al., 2016b***) might influence the types of variations that we tend to imagine (e.g., articulating limbs into different poses). The fact that observers can derive diverse but constrained variations from a single exemplar suggests a deep perceptual understanding about the ways things in the natural world tend to vary.

In future work, it would also be interesting to analyze the similarities between human drawings (and their similarities to the exemplars) using artificial neural networks trained on datasets of line drawings, such as the ‘Quick Draw!’ dataset (***Jongejan et al., 2016***; ***Ha and Eck, 2017***; ***Xu et al., 2021***; ***Kabakus, 2020***). In particular, each of the human drawings (or exemplars) could be fed into such a network, and a feature vector describing the shape derived, allowing an automated quantitative similarity analysis to be performed. Here, we used human-based ratings to perform such analyses, but this approach does not scale well to larger datasets. Even more intriguing is the possibility of training a neural network model to reproduce the types of generalizations that humans produce when presented with a single exemplar. This would answer questions, such as whether it is necessary to be exposed to human-created line drawings in order to generalize like humans, or whether a visual diet consisting entirely of natural images is sufficient. It would also allow a more rigorous test of whether part-based representations are necessary to capture the full range of human generalization. However, to date, one-shot generalization as exhibited by our observers remains a significant challenge for artificial vision systems.

Materials and methods

Participants were recruited through university mailing lists. All participants gave informed consent before the experiments in accordance with the Declaration of Helsinki. The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of the Department of Psychology and Sports Sciences of the Justus-Liebig University Giessen (LEK-FB06; protocol code 2016-0007, approved 18 April 2016). **Experiment 1** was conducted on a touchpad computer, all others on a computer.

Sample sizes

In the only studies known by the authors that employed a generative one-shot categorization task (***Feldman, 1997***), the number of objects drawn per Exemplar ranged from 96 to 168. Given that these Exemplars were comparatively simple (e.g., a dot on a line, two connected lines), we aimed for a substantially larger number of drawings per Exemplar (204) to allow all potentially relevant features of our Exemplars to be expressed in the generated objects. Given the novelty of this paradigm, estimating effect sizes for follow-up studies using the drawn shapes proved difficult. For this reason, all follow-up experiments in this study aimed for very high sample sizes to allow the detection of potentially very small effects. For example, in **Experiment 3**, in which chance level is at 12.5%, to detect a significant difference from that level of 2.5% using a one-sided binomial test (with Bonferroni-adjusted alpha = 0.00625 and power = 0.95) we would need a sample size of 3197 trials, a number we have exceeded (4800 trials) for additional sensitivity. Concerning *t*-tests, to detect a small Cohen's *d* of 0.2 (with alpha = 0.05 and power = 0.95), a minimum sample size of 1084 is necessary (for a one-sided *t*-test), which is exceeded, for example, in **Experiment 4** with 1200 data points per condition.

Experiment 1: generative one-shot categorization task

Seventeen participants drew 12 new shapes for each of the 8 Exemplars (204 shapes per participant, overall 1632 shapes). They were instructed to draw a new object, belonging to the Exemplar's class, that was not just a mere copy of the Exemplar. The Exemplars were shown in randomized order in the upper part of an iPad Pro (12.9") oriented in portrait mode and observers' drawings were on the lower part of the screen. While drawing, participants could switch between drawing freely and drawing straight lines. Participants could clear the current drawing or undo the last drawn segment. After putting down the pen they could only continue drawing from the end-point of the last segment to prevent gaps between lines. After finishing a shape, the drawing area was cleared and a new object could be drawn. A shape could only be finished if the contour was closed, that is, the first and last point drawn were on the same spot. Participants were instructed not to draw any overlapping or crossing lines.

Experiment 2: perceived similarity to Exemplars

For each category, 12 new participants judged the perceived similarity of all drawings created in Experiment 1 to their originating Exemplar. Before making any judgements, all Variations of a category were shown in grids of 3 by 4 shapes consecutively to give the observers an idea of the range of shapes to be judged. In total, participants viewed 17 grids of 12 shapes. They were then presented with the Exemplar on the far left of the screen. On the screen bottom, participants used the mouse to drag and drop 3 randomly chosen Variations from that category into the upper part of the screen and arrange them based on similarity to the Exemplar: the closer the shape was placed on the *x*-axis towards the Exemplar, the more similar it was judged to be. Participants were instructed to try and keep a consistent scale of similarity, meaning that equally similar shapes relative to the Exemplar should be placed in the same area of the screen, across trials and irrespective of how similar the shapes were to each other. After placing three shapes relative to the Exemplar, a button press revealed the next three shapes on the screen's bottom portion. Already placed shapes were greyed out but could still be adjusted in position if so desired. To prevent the screen from getting cluttered, old shapes disappeared after 4 trials so that only 12 shapes were shown at one time. For each of the participants responses, the similarity values were normalized between 0 (the Exemplar) and 1 (the least similar drawing). The final similarity value of each drawing was averaged across observers over these normalized responses.

Picking stimuli from similarity space

In **Experiments 2b**, **3**, **4**, **5**, and **6**, we selected a subset of shapes from each category ensuring that the subset spanned the whole range of similarities. As an example, a subset of 20 shapes was created by taking all shapes of a category (minus the Exemplar) and dividing them into 20 equally sized bins spanning the perceived similarity space derived in **Experiment 2**. For each bin, we selected the shape with the lowest between-participant variance in similarity judgements. If a bin was empty, neighbouring bins were searched for a shape not yet used.

Experiment 2b: comparing copies and new drawings

Fifteen participants were instructed to copy each Exemplar three times as best as possible with the drawing interface used in **Experiment 1**, resulting in 45 copies of each Exemplar. Then, 15 new participants were shown one of these copies, along with the Exemplar and one Variation of that category per trial. The task was to pick the shape that was a copy of the Exemplar. The 45 Variations per category were selected to span the range of similarities (see **Picking stimuli from similarity space**) and randomly paired with one of the copies. Each category comprised one block of trials with blocks being randomized in order, resulting in 45 × 8 = 360 trials. These trials were repeated three times resulting in 1080 trials overall. In each repeat, the order of blocks and pairings of copies and Variations was randomized.

Experiment 3: perceptual category membership

Fifteen participants sequentially judged which Exemplar category 40 drawings of each category (chosen as described in **Picking stimuli from similarity space**) belonged to, resulting in 320 data points per participant. If participants were unsure about the category membership, they were instructed to pick the category they thought the shape belonged to the most.

Experiment 4: corresponding parts experiment

Fifteen participants participated in this experiment. For each Exemplar, a subset of 10 drawings spanning the range of similarities was chosen (see **Picking stimuli from similarity space**). Participants were shown an Exemplar on the left and a drawing on the right, either belonging to the same or a different category. They were then asked if they saw any corresponding parts between these shapes. If not, the next shape pair was shown. Otherwise they were asked to pick the corresponding parts by first picking the part in the Exemplar and then the corresponding part in the drawing. To pick a part, the two delineating points of the part were to be clicked in succession, creating a line between these points within the shape. After that, the polygon on either side of the line could be chosen as the final part. A part could only be picked if the resulting line between the points did not intersect with other parts of the shape. Instead of picking a part, the rest of the shape (comprised of anything not yet picked) could also be picked, indicated by a red dot at the centroid of the remaining shape. Any section already picked could not be used for another part. Any number of part pairs could be picked, unless the remaining shape was picked, after which the trial was ended, since no more unpicked parts remained.

Each Exemplar was paired with 10 corresponding and 10 Variations of another randomly chosen category, resulting in 8 × 20 = 160 trials per participant.

Part order analysis

To compare part ordering in Exemplar/Variation pairs we created a simplified representation of parts (see ***Figure 3—figure supplement 2a***): Starting with the left-most point of a shape and moving in clockwise order, each point was tested whether it was within a designated part, or not. Accordingly, a value or an empty value (or 'gap') was added to an array. The array values only changed when encountering a new part (or non-part; i.e., if the current part was not the same as for the previous point). While this is a simplified representation of part order, we believe it to be a sensible approach for the shapes of this study, given that they mainly consisted of one central main body with a differing number of parts attached. This analysis results in a circular array showing the clockwise ordering of parts for each shape ('part circle'). To compare part order, non-corresponding parts (i.e., gap segments) were ignored, since we were only interested in the order of corresponding parts. By comparing these arrays, we could see whether part ordering was identical, reversed, or shuffled. Since order between two circular arrays could only be different when they consisted of more than two values (i.e., the array [A,B] is considered the same as [B,A]), only trials with more than two correspondences were considered, amounting to 30% of trials. Part changes, for example, omitted or substituted parts of an Exemplar or parts added to the Variation that did not exist in the Exemplar, were defined by gaps in the part order array: A gap in the Exemplar with no counterpart in the Variation at that same spot was an omission. A gap in the Variation but not the Exemplar was considered an addition. If there was a gap at the same spot in both shapes it was defined as a substitution. For this part-change analysis all stimuli where any correspondence was seen was used (85% of trials).

Experiment 5: marking distinctive parts

Ten participants were sequentially shown a subset of 39 shapes (chosen as described in **Picking stimuli from similarity space**) for each category, plus the Exemplar, resulting in 320 trials overall. They were asked to mark the 'most distinctive areas' of each shape. They could freely paint on the shape's silhouette and were not constrained to paint only consecutive parts. After painting the most distinctive part or parts in red they could switch to the next lower distinctiveness tier (orange) and paint the second most distinctive areas. After that they could switch to yellow for the third most distinctive area. At least one part had to be painted red, the lower distinctiveness tiers were optional.

To aggregate these responses (as in ***Figure 4*** and ***Figure 3—figure supplement 1***), each point of the contour was given a score, with each red response adding 3, each orange 2 and each yellow adding 1 to the score of that point. After summing the individual scores, the resultant scores were normalized between 0 (score = 0) and 100 (highest possible score).

Creation of randomized responses for comparison with data from Experiment 5

For each categories' responses, we computed how often each distinctiveness tier was used (except the first tier, which was always used). In addition, we computed the distribution of lengths of each tier's consecutive painted areas and the distribution of number of non-consecutive areas per tier and category. With these distributions 10 (the number of participants in **Experiment 5**) randomized responses were created for each of the shapes from **Experiment 5**, meaning the average number of areas and consecutive lengths mimicked the human distribution, while the placement on the shape was random.

Experiment 6: swapping distinctive parts experiment. Stimuli

First, we segmented each shape of **Experiment 5** into one distinctive and one indistinctive part. The distinctive part was the largest consecutive part of the shape with an aggregated distinctiveness score higher than 75 in each point. The indistinctive part was the rest of the shape.

In 'distinctive part swapped' shapes, we swapped the distinctive part with a distinctive part of another categories' Exemplar. In 'indistinctive part swapped' shapes, we swapped an indistinctive part of the shape with the same number of points as the distinctive part that had the lowest average distinctiveness score. Since points were equally spaced along the contour both parts had the same perimeter length. In either case the swapped-in part was rotated and either uniformly compressed or stretched to fit the gap as best as possible.

In each case a drawing from the subset from **Experiment 5** (excluding the Exemplars) was used as the base shape. In this way for each of the 8 categories, 5 shapes were created with parts swapped from one of each of the other 7 categories both once with the distinctive part and once with the indistinctive part swapped, resulting in $8 \times 7 \times 2 \times 5 = 560$ shapes.

Because this process sometimes created shapes with large self-intersections, the final five shapes for each condition were hand-picked to create well-formed, artefact-free shapes.

Procedure

Fifteen participants then conducted a repeat of the 'Perceptual category-membership' experiment with the 560 stimuli (see **Experiment 3**).

Perceived Exemplar complexity

Twenty participants arranged the eight Exemplars in order of perceived complexity from simple to complex. The average rank of these responses was used to order the Exemplars and their categories in most plots of this paper.

Acknowledgements

Research funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation—project number 222641018—SFB/TRR 135 TP C1), by the European Research Council (ERC) Consolidator Award 'SHAPE'—project number ERC-2015-CoG-682859 and by 'The Adaptive Mind', funded by the Excellence Program of the Hessian Ministry of Higher Education, Science, Research and Art.

Additional information

Funding

Funder	Grant reference number	Author
Deutsche Forschungsgemeinschaft	222641018-SFB/TRR 135 TP C1	Roland W Fleming Filipp Schmidt
European Research Council	ERC-2015-CoG-682859	Roland W Fleming
Hessian Ministry of Higher Education, Science, Research and Art	The Adaptive Mind	Roland W Fleming

The funders had no role in study design, data collection, and interpretation, or the decision to submit the work for publication.

Author contributions

Henning Tiedemann, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Visualization, Writing - original draft; Yaniv Morgenstern, Filipp Schmidt, Conceptualization, Formal analysis, Methodology, Resources, Software, Supervision, Writing – review and editing; Roland W Fleming, Conceptualization, Formal analysis, Funding acquisition, Project administration, Resources, Software, Supervision, Writing – review and editing

Author ORCIDs

Henning Tiedemann http://orcid.org/0000-0003-2637-7888
Yaniv Morgenstern http://orcid.org/0000-0001-6902-667X
Filipp Schmidt http://orcid.org/0000-0001-8023-7304
Roland W Fleming http://orcid.org/0000-0001-5033-5069

Ethics

The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of the Department of Psychology and Sports Sciences of the Justus-Liebig University Giessen (LEK-FB06; protocol code 2016-0007, approved 18 April 2016).

Decision letter and Author response

Decision letter https://doi.org/10.7554/eLife.75485.sa1
Author response https://doi.org/10.7554/eLife.75485.sa2

Additional files

Supplementary files

- Transparent reporting form

Data availability

All data generated and analyzed are available at zenodo at: https://zenodo.org/record/5230306.

The following dataset was generated:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Tiedemann H, Morgenstern Y, Schmidt F, Fleming RW	2021	One shot generalization in humans revealed through a drawing task - Dataset	https://zenodo.org/record/5230306	Zenodo, 10.5281/zenodo.5230306

References

Ashby FG, Maddox WT. 2005. Human category learning. *Annual Review of Psychology* **56**:149–178. DOI: https://doi.org/10.1146/annurev.psych.56.091103.070217, PMID: 15709932

Attneave F. 1957. Physical determinants of the judged complexity of shapes. *Journal of Experimental Psychology* **53**:221–227. DOI: https://doi.org/10.1037/h0043921, PMID: 13416488

Ayzenberg V, Lourenco S. 2021. The shape skeleton supports one-shot categorization in human infants. *Journal of Vision* **21**:2316. DOI: https://doi.org/10.1167/jov.21.9.2316

Bainbridge WA, Hall EH, Baker CI. 2019. Drawings of real-world scenes during free recall reveal detailed object and spatial information in memory. *Nature Communications* **10**:1–13. DOI: https://doi.org/10.1038/s41467-018-07830-6, PMID: 30602785

Baker N, Lu H, Erlikhman G, Kellman PJ. 2018. Deep convolutional networks do not classify based on global object shape. *PLOS Computational Biology* **14**:e1006613. DOI: https://doi.org/10.1371/journal.pcbi.1006613, PMID: 30532273

Baker N, Garrigan P, Kellman PJ. 2021. Constant curvature segments as building blocks of 2D shape representation. *Journal of Experimental Psychology. General* **150**:1556–1580. DOI: https://doi.org/10.1037/xge0001007, PMID: 33332142

Battaglia PW, Hamrick JB, Tenenbaum JB. 2013. Simulation as an engine of physical scene understanding. *PNAS* **110**:18327–18332. DOI: https://doi.org/10.1073/pnas.1306572110, PMID: 24145417

Behrmann M, Nishimura M. 2010. Agnosias. *Wiley Interdisciplinary Reviews. Cognitive Science* **1**:203–213. DOI: https://doi.org/10.1002/wcs.42, PMID: 26271235

Biederman I. 1987. Recognition-by-components: a theory of human image understanding. *Psychological Review* **94**:115–147. DOI: https://doi.org/10.1037/0033-295X.94.2.115, PMID: 3575582

Destler N, Singh M, Feldman J. 2019. Shape discrimination along morph-spaces. *Vision Research* **158**:189–199. DOI: https://doi.org/10.1016/j.visres.2019.03.002, PMID: 30878276

Fan JE, Wammes JD, Gunn JB, Yamins DLK, Norman KA, Turk-Browne NB. 2020. Relating Visual Production and Recognition of Objects in Human Visual Cortex. *The Journal of Neuroscience* **40**:1710–1721. DOI: https://doi.org/10.1523/JNEUROSCI.1843-19.2019, PMID: 31871278

Fei-Fei L, Fergus R, Perona P. 2006. One-shot learning of object categories. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **28**:594–611. DOI: https://doi.org/10.1109/TPAMI.2006.79, PMID: 16566508

Fei-Fei L, Iyer A, Koch C, Perona P. 2007. What do we perceive in a glance of a real-world scene? *Journal of Vision* **7**:10. DOI: https://doi.org/10.1167/7.1.10, PMID: 17461678

Feldman J. 1992. Constructing perceptual categories. IEEE Computer Society Conference on Computer Vision and Pattern Recognition. . DOI: https://doi.org/10.1109/CVPR.1992.223268

Feldman J. 1997. The Structure of Perceptual Categories. *Journal of Mathematical Psychology* **41**:145–170. DOI: https://doi.org/10.1006/jmps.1997.1154, PMID: 9237918

Feldman J, Singh M. 2005. Information along contours and object boundaries. *Psychological Review* **112**:243–252. DOI: https://doi.org/10.1037/0033-295X.112.1.243, PMID: 15631595

Feldman J, Singh M. 2006. Bayesian estimation of the shape skeleton. *PNAS* **103**:18014–18019. DOI: https://doi.org/10.1073/pnas.0608811103, PMID: 17101989

Feldman J. 2013. An integrated Bayesian approach to shape representation and perceptual organization. Feldman J (Ed). *In Shape Perception in Human and Computer Vision*. Springer. p. 55–70. DOI: https://doi.org/10.1007/978-1-4471-5195-1_4

Fleming RW, Schmidt F. 2019. Getting “fumpered”: Classifying objects by what has been done to them. *Journal of Vision* **19**:15. DOI: https://doi.org/10.1167/19.4.15

Gauthier I, Williams P, Tarr MJ, Tanaka J. 1998. Training “greeble” experts: a framework for studying expert object recognition processes. *Vision Research* **38**:2401–2428. DOI: https://doi.org/10.1016/s0042-6989(97)00442-2, PMID: 9798007

Geirhos R, Temme CRM, Rauber J, Schütt HH, Bethge M, Wichmann FA. 2018. Generalisation in Humans and Deep Neural Networks. *arXiv*. https://arxiv.org/abs/1808.08750DOI: https://doi.org/https://arxiv.org/abs/1808.08750

Gelman SA, Markman EM. 1986. Categories and induction in young children. *Cognition* **23**:183–209. DOI: https://doi.org/10.1016/0010-0277(86)90034-x, PMID: 3791915

Gelman SA, Meyer M. 2011. Child categorization. *Wiley Interdisciplinary Reviews. Cognitive Science* **2**:95–105. DOI: https://doi.org/10.1002/wcs.96, PMID: 23440312

Gershkoff-Stowe L, Thal DJ, Smith LB, Namy LL. 1997. Categorization and Its Developmental Relation to Early Language. *Child Development* **68**:843–859. DOI: https://doi.org/10.1111/j.1467-8624.1997.tb01966.x, PMID: 29106722

Goodale MA, Milner AD, Jakobson LS, Carey DP. 1991. A neurological dissociation between perceiving objects and grasping them. *Nature* **349**:154–156. DOI: https://doi.org/10.1038/349154a0, PMID: 1986306

Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Bengio Y. 2014. Generative adversarial nets. Advances in neural information processing systems. .

Goodman ND, Tenenbaum JB, Feldman J, Griffiths TL. 2008a. A rational analysis of rule-based concept learning. *Cognitive Science* **32**:108–154. DOI: https://doi.org/10.1080/03640210701802071, PMID: 21635333

Goodman ND, Tenenbaum JB, Griffiths TL, Feldman J. 2008b. Compositionality in rational analysis: Grammar-based induction for concept learning. The Probabilistic Mind: Prospects for Bayesian Cognitive Science. . DOI: https://doi.org/10.1093/acprof:oso/9780199216093.001.0001

Gopnik A, Sobel DM. 2000. Detecting blickets: how young children use information about novel causal powers in categorization and induction. *Child Development* **71**:1205–1222. DOI: https://doi.org/10.1111/1467-8624.00224, PMID: 11108092

Ha D, Eck D. 2017. A Neural Representation of Sketch Drawings. *arXiv*. https://arxiv.org/pdf/1704.03477.pdf

He K, Zhang X, Ren S, Sun J. 2016. Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR. . DOI: https://doi.org/10.1109/CVPR.2016.90

Hegdé J, Bart E, Kersten D. 2008. Fragment-based learning of visual object categories. *Current Biology* **18**:597–601. DOI: https://doi.org/10.1016/j.cub.2008.03.058, PMID: 18424145

Hoffman DD, Richards WA. 1984. Parts of recognition. *Cognition* **18**:65–96. DOI: https://doi.org/10.1016/0010-0277(84)90022-2, PMID: 6543164

Hoffman DD, Singh M. 1997. Salience of visual parts. *Cognition* **63**:29–78. DOI: https://doi.org/10.1016/s0010-0277(96)00791-3, PMID: 9187064

Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. 2017. Densely Connected Convolutional Networks. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR. 4700–4708. DOI: https://doi.org/10.1109/CVPR.2017.243

Jongejan J, Rowley H, Kawashima T, Kim J, Fox-Gieg N. 2016. The Quick, 953 Draw! [Experiment]. https://quickdraw.withgoogle.com [Accessed May 6, 2022].

Jozwik K, Kriegeskorte N, Cichy R, Mur M. 2016. Visual features versus categories: Explaining object representations in primate IT and deep neural networks with weighted representational modeling. *Journal of Vision* **16**:511. DOI: https://doi.org/10.1167/16.12.511

Jozwik KM, Kriegeskorte N, Cichy RM, Mur M. 2018. Deep convolutional neural networks, features, and categories perform similarly at explaining primate high-level visual representations. 2018 Conference on Cognitive Computational Neuroscience. . DOI: https://doi.org/10.32470/CCN.2018.1232-0

Jozwik KM, Lee M, Marques T, Schrimpf M, Bashivan P. 2019. Large-Scale Hyperparameter Search for Predicting Human Brain Responses in the Algonauts Challenge. *bioRxiv*. DOI: https://doi.org/10.1101/689844

Kabakus AT. 2020. A Novel Sketch Recognition Model based on Convolutional Neural Networks. 2020 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA). , 1–6. DOI: https://doi.org/10.1109/HORA49412.2020.9152911

Kayaert G, Biederman I, Op de Beeck HP, Vogels R. 2005. Tuning for shape dimensions in macaque inferior temporal cortex. *The European Journal of Neuroscience* **22**:212–224. DOI: https://doi.org/10.1111/j.1460-9568.2005.04202.x, PMID: 16029211

Krizhevsky A, Sutskever I, Hinton GE. 2012. Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems. 1097–1105.

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: https://doi.org/10.1371/journal.pone.0015444, PMID: 21124837

Kubilius J, Bracci S, Op de Beeck HP. 2016. Deep Neural Networks as a Computational Model for Human Shape Sensitivity. *PLOS Computational Biology* **12**:e1004896. DOI: https://doi.org/10.1371/journal.pcbi.1004896, PMID: 27124699

Lake BM, Salakhutdinov R, Tenenbaum JB. 2015. Human-level concept learning through probabilistic program induction. *Science (New York, N.Y.)* **350**:1332–1338. DOI: https://doi.org/10.1126/science.aab3050, PMID: 26659050

Landau B, Smith LB, Jones SS. 1988. The importance of shape in early lexical learning. *Cognitive Development* **3**:299–321. DOI: https://doi.org/10.1016/0885-2014(88)90014-7

Landau B, Smith L, Jones S. 1998. Object perception and object naming in early development. *Trends in Cognitive Sciences* **2**:19–24. DOI: https://doi.org/10.1016/s1364-6613(97)01111-x, PMID: 21244958

Long B, Fan J, Chai Z, Frank MC. 2019. Developmental Changes in the Ability to Draw Distinctive Features of Object Categories. *PsyArXiv*. DOI: https://doi.org/10.31234/osf.io/8rzku

Mack ML, Palmeri TJ. 2015. The dynamics of categorization: Unraveling rapid categorization. *Journal of Experimental Psychology. General* **144**:551–569. DOI: https://doi.org/10.1037/a0039184, PMID: 25938178

Marr D, Nishihara HK. 1978. Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London. Series B, Biological Sciences* **200**:269–294. DOI: https://doi.org/10.1098/rspb.1978.0020, PMID: 24223

McCarthy RA, Warrington EK. 1986. Visual associative agnosia: a clinico-anatomical study of a single case. *Journal of Neurology, Neurosurgery, and Psychiatry* **49**:1233–1240. DOI: https://doi.org/10.1136/jnnp.49.11.1233, PMID: 3794729

Michaelis C, Bethge M, Ecker AS. 2020. Closing the Generalization Gap in One-Shot Object Detection. *arXiv*. https://arxiv.org/abs/2011.04267

Morgenstern Y, Schmidt F, Fleming RW. 2019. One-shot categorization of novel object classes in humans. *Vision Research* **165**:98–108. DOI: https://doi.org/10.1016/j.visres.2019.09.005, PMID: 31707254

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 Computational Biology* **17**:e1008981. DOI: https://doi.org/10.1371/journal.pcbi.1008981, PMID: 34061825

Mukherjee K, Hawkins RX, Fan JW. 2019. Communicating semantic part information in drawings. CogSci... Annual Conference of the Cognitive Science Society. Cognitive Science Society (U.S.). Conference. 2413–2419.

Ons B, Wagemans J. 2012. Generalization of visual shapes by flexible and simple rules. *Seeing and Perceiving* **25**:237–261. DOI: https://doi.org/10.1163/187847511X571519, PMID: 21771394

Op de Beeck HP, Torfs K, Wagemans J. 2008. Perceived shape similarity among unfamiliar objects and the organization of the human object vision pathway. *The Journal of Neuroscience* **28**:10111–10123. DOI: https://doi.org/10.1523/JNEUROSCI.2511-08.2008, PMID: 18829969

Op de Beeck HP, Baker CI. 2010. The neural basis of visual object learning. *Trends in Cognitive Sciences* **14**:22–30. DOI: https://doi.org/10.1016/j.tics.2009.11.002, PMID: 19945336

Pereira AF, Smith LB. 2009. Developmental changes in visual object recognition between 18 and 24 months of age. *Developmental Science* **12**:67–80. DOI: https://doi.org/10.1111/j.1467-7687.2008.00747.x, PMID: 19120414

Radford A. 2021. Learning Transferable Visual Models from Natural Language Supervision. *arXiv*. https://arxiv.org/abs/2103.00020

Richards W, Feldman J, Jepson A. 1992. From Features to Perceptual Categories. British Machine Vision Conference 1992. 99–108. DOI: https://doi.org/10.1007/978-1-4471-3201-1

Riddoch MJ, Humphreys GW. 1987. A case of integrative visual agnosia. *Brain* **110 (Pt 6)**:1431–1462. DOI: https://doi.org/10.1093/brain/110.6.1431, PMID: 3427396

Schmidt F, Fleming RW. 2016a. Visual perception of complex shape-transforming processes. *Cognitive Psychology* **90**:48–70. DOI: https://doi.org/10.1016/j.cogpsych.2016.08.002, PMID: 27631704

Schmidt F, Spröte P, Fleming RW. 2016b. Perception of shape and space across rigid transformations. *Vision Research* **126**:318–329. DOI: https://doi.org/10.1016/j.visres.2015.04.011, PMID: 25937375

Schmidt F, Phillips F, Fleming RW. 2019. Visual perception of shape-transforming processes: "Shape Scission." *Cognition* **189**:167–180. DOI: https://doi.org/10.1016/j.cognition.2019.04.006, PMID: 30986590

Serre T, Oliva A, Poggio T. 2007. A feedforward architecture accounts for rapid categorization. *PNAS* **104**:6424–6429. DOI: https://doi.org/10.1073/pnas.0700622104

Serre T. 2016. Models of visual categorization. *Wiley Interdisciplinary Reviews. Cognitive Science* **7**:197–213. DOI: https://doi.org/10.1002/wcs.1385, PMID: 26997155

Sharan L, Rosenholtz R, Adelson E. 2010. Material perception: What can you see in a brief glance? *Journal of Vision* **9**:784. DOI: https://doi.org/10.1167/9.8.784

Sheng H, Wilder J, Walther DB. 2021. Where to draw the line? *PLOS ONE* **16**:e0258376. DOI: https://doi.org/10.1371/journal.pone.0258376, PMID: 34748556

Siddiqi K, Kimia BB. 1995. Parts of visual form: computational aspects. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **17**:239–251. DOI: https://doi.org/10.1109/34.368189

Singh M, Seyranian GD, Hoffman DD. 1999. Parsing silhouettes: the short-cut rule. *Perception & Psychophysics* **61**:636–660. DOI: https://doi.org/10.3758/bf03205536, PMID: 10370334

Smith LB, Slone LK. 2017. A Developmental Approach to Machine Learning? *Frontiers in Psychology* **8**:2124. DOI: https://doi.org/10.3389/fpsyg.2017.02124, PMID: 29259573

Spröte P, Schmidt F, Fleming RW. 2016. Visual perception of shape altered by inferred causal history. *Scientific Reports* **6**:1–11. DOI: https://doi.org/10.1038/srep36245, PMID: 27824094

Stuhlmuller A, Tenenbaum JB, Goodman ND. 2010. Learning structured generative concepts. Cognitive Science Society. .

Szegedy C, Wei L, Yangqing J, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. 2015. Going deeper with convolutions. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR. 1–9. DOI: https://doi.org/10.1109/CVPR.2015.7298594

Szegedy C., Vanhoucke V, Ioffe S, Shlens J, Wojna Z. 2016. Rethinking the Inception Architecture for Computer Vision. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA, 2818–2826. DOI: https://doi.org/10.1109/CVPR.2016.308

Szegedy C, Ioffe S, Vanhoucke V, Alemi AA. 2017. Inception-v4, inception-resnet and the impact of residual connections on learning. in Thirty-first AAAI conference on artificial intelligence. .

Thorpe S, Fize D, Marlot C. 1996. Speed of processing in the human visual system. *Nature* **381**:520–522. DOI: https://doi.org/10.1038/381520a0, PMID: 8632824

Tversky B. 1989. Parts, partonomies, and taxonomies. *Developmental Psychology* **25**:983–995. DOI: https://doi.org/10.1037/0012-1649.25.6.983

Ullman S, Vidal-Naquet M, Sali E. 2002. Visual features of intermediate complexity and their use in classification. *Nature Neuroscience* **5**:682–687. DOI: https://doi.org/10.1038/nn870, PMID: 12055634

Wilder J, Feldman J, Singh M. 2011. Superordinate shape classification using natural shape statistics. *Cognition* **119**:325–340. DOI: https://doi.org/10.1016/j.cognition.2011.01.009, PMID: 21440250

Wilder J, Dickinson S, Jepson A, Walther DB. 2018. Spatial relationships between contours impact rapid scene classification. *Journal of Vision* **18**:1. DOI: https://doi.org/10.1167/18.8.1, PMID: 30073270

Wilder J, Rezanejad M, Dickinson S, Siddiqi K, Jepson A, Walther DB. 2019. Local contour symmetry facilitates scene categorization. *Cognition* **182**:307–317. DOI: https://doi.org/10.1016/j.cognition.2018.09.014, PMID: 30415132

Xu P, Song Z, Yin Q, Song YZ, Wang L. 2021. Deep Self-Supervised Representation Learning for Free-Hand Sketch. *IEEE Transactions on Circuits and Systems for Video Technology* **31**:1503–1513. DOI: https://doi.org/10.1109/TCSVT.2020.3003048

Yuan L, Xiang V, Crandall D, Smith L. 2020. Learning the generative principles of a symbol system from limited examples. *Cognition* **200**:104243. DOI: https://doi.org/10.1016/j.cognition.2020.104243, PMID: 32151856

Zhang H, Yu Y, Jiao J, Xing E, El Ghaoui L, Jordan M. 2019. Theoretically principled trade-off between robustness and accuracy. In International Conference on Machine Learning. 7472–7482.

Zhang C, Bengio S, Hardt M, Recht B, Vinyals O. 2021. Understanding deep learning (still) requires rethinking generalization. *Communications of the ACM* **64**:107–115. DOI: https://doi.org/10.1145/3446776

brain sciences

MDPI

Article

Superordinate Categorization Based on the Perceptual Organization of Parts

Henning Tiedemann [1,*], Filipp Schmidt [1,2] and Roland W. Fleming [1,2]

[1] Department of Experimental Psychology, Justus Liebig University Giessen, 35394 Giessen, Germany; filipp.schmidt@psychol.uni-giessen.de (F.S.); roland.w.fleming@psychol.uni-giessen.de (R.W.F.)
[2] Center for Mind, Brain and Behavior (CMBB), University of Marburg and Justus Liebig University Giessen, 35037 Marburg, Germany
* Correspondence: henning.tiedemann@psychol.uni-giessen.de

Abstract: Plants and animals are among the most behaviorally significant superordinate categories for humans. Visually assigning objects to such high-level classes is challenging because highly distinct items must be grouped together (e.g., chimpanzees and geckos) while more similar items must sometimes be separated (e.g., stick insects and twigs). As both animals and plants typically possess complex multi-limbed shapes, the perceptual organization of shape into parts likely plays a crucial rule in identifying them. Here, we identify a number of distinctive growth characteristics that affect the spatial arrangement and properties of limbs, yielding useful cues for differentiating plants from animals. We developed a novel algorithm based on shape skeletons to create many novel object pairs that differ in their part structure but are otherwise very similar. We found that particular part organizations cause stimuli to look systematically more like plants or animals. We then generated other 110 sequences of shapes morphing from animal- to plant-like appearance by modifying three aspects of part structure: sprouting parts, curvedness of parts, and symmetry of part pairs. We found that all three parameters correlated strongly with human animal/plant judgments. Together our findings suggest that subtle changes in the properties and organization of parts can provide powerful cues in superordinate categorization.

Keywords: visual perception; objects; shape; features; visual cognition

check for updates

Citation: Tiedemann, H.; Schmidt, F.; Fleming, R.W. Superordinate Categorization Based on the Perceptual Organization of Parts. *Brain Sci.* **2022**, *12*, 667. https://doi.org/10.3390/brainsci12050667

Academic Editors: Baingio Pinna and Amedeo D'Angiulli

Received: 10 April 2022
Accepted: 17 May 2022
Published: 20 May 2022

1. Introduction

Visually differentiating between superordinate classes, e.g., plants and animals, is a vital skill with direct consequences for our behavior. For example, many animals can move at speeds that demand immediate action—such as fleeing or attacking—while plants do not. However, the visual distinction between these classes is not always easy. Even though objects within a superordinate class typically share particular characteristics (e.g., visible motion), they might still look very different. For example, elephants and insects both belong to the animal class, but are very different in their visual appearance. In contrast, insects and twigs can share many visual features, such as thin and spindly limbs, but belong to different superordinate classes. Consequently, we need to be able to quickly analyze and interpret sometimes subtle shape differences to achieve correct categorizations and adjust our behavior accordingly.

Previous research identified some of the visual aspects distinguishing superordinate classes like animate and inanimate objects [1–3] as well as aspects of the neural processing that drives categorization [4–6]. Specifically, these studies focused on diagnostic local features such as eyes and mouth (e.g., [1]), individual parts such as tails [7], global features such as canonical animal postures [8], or overall curvature [9–11]. However, all of these studies are virtually agnostic with respect to the role of perceptual organization of object shape, that is, the spatial arrangement and relations between the different parts of an object.

Hereinafter, we use the term "part structure" to refer to the visual segmentation of an object into its parts (e.g., [12,13]) as well as to the analysis of relational information—for example, to which other parts a given part is connected (e.g., [14]), whether parts form a symmetrical pair (e.g., [15]), or other non-accidental configurations like collinearity of parts [16,17].

There are several reasons why part structure might contribute to superordinate object categorization. First, natural growth processes tend to follow systematic rules [18], resulting in highly regular growth patterns. This constrains the perceptual organization of object parts in otherwise visually highly varied categories (Figure 1). For example, the presence or absence of symmetrical parts is a distinctive quality of part structure. Animals such as insects, elephants and vultures could hardly be more visually distinct, yet all of them are comprised of a main body with prominent pairs of symmetrical limbs, most often arms and legs. Indeed, biologically, the vast majority of animals follow a bilaterally symmetric body plan [19] supporting the notion that humans have evolved (or learned) to expect this part structure in animals. In line with this, previous studies have found that symmetry is indeed a strong animal-cue [3]. In contrast, plants such as oak trees or ferns typically grow parts radially in a Fibonacci sequence [20], resulting in a roughly alternating left/right placement of parts when viewed from the side. Experiments using abstract shapes have shown that regularities in part structure can, in certain situations, drive categorization [21], supporting the notion that these ubiquitous natural regularities play a role as well.

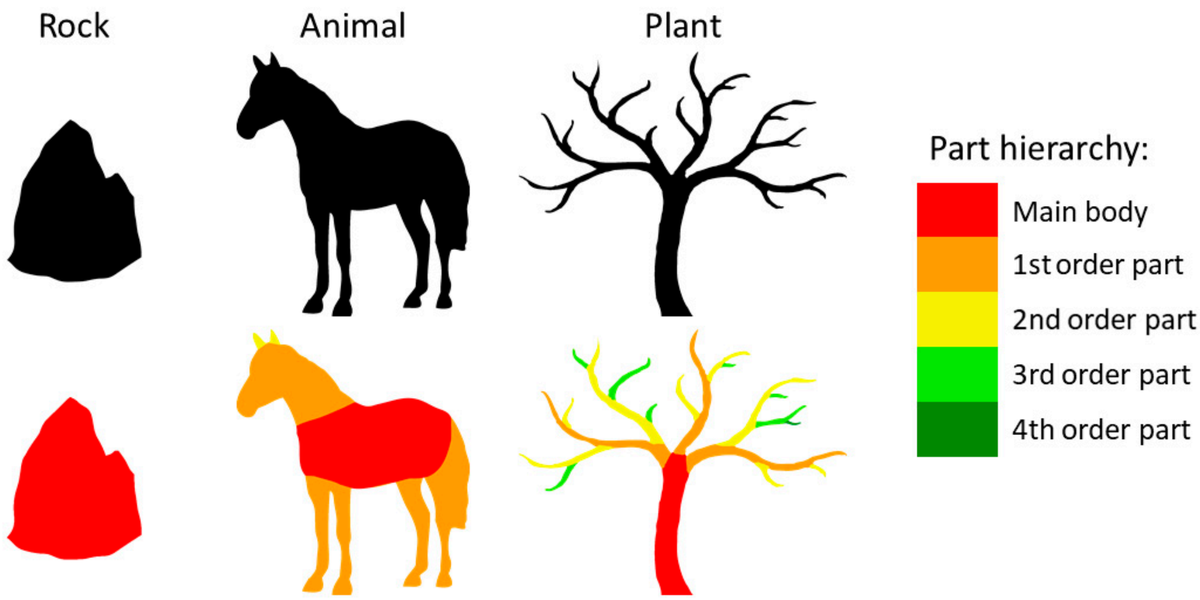

Figure 1. Scheme illustrating that different superordinate categories often show striking differences in their perceptual organization: rocks are usually made up of only one central body, without any limbs; animals tend to have prominent first-order parts growing from the main body with few higher-order parts; plants often have additional layers of parts further removed from the main body. Part organizations were created by the authors for illustrative purposes, not as an algorithm.

Second, part structure is a useful visual feature because it is independent of temporary limb articulation. For example, an ape rarely positions its limbs in perfect mirror symmetry formation. Instead, the limbs are configured according to the task currently being performed. However, by attending to the part structure, that is, to where the limbs are attached to the ape's main body, an observer can abstract away from their current shape and positioning. Although limb pose can be highly flexible, an animal's part structure rarely changes with motion, so it is an immutable feature of its growth history and class.

Third, animals and plants can be differentiated perceptually by the presence of higher-order parts. Animals mainly consist of a main body with prominent first-order parts (limbs that grow directly from the main body). In contrast, plants are often composed of a hierarchy including many higher-order parts, i.e., parts that do not connect directly to the main body (red in Figure 1) but rather derive from a first-order or a higher-order part (orange to green parts in Figure 1). These higher-order parts of plants often sprout off their parents, resulting in sharp angles that make them more salient and visually distinct. In contrast, if animals possess higher-order parts at all—for example, fingers or claws—they often grow off the parent part's end, and the resulting obtuse angles make them look more like a continuation of the parent than distinct parts on their own. As a result of these typical differences, symmetry of part pairs and sprouting higher-order parts are potentially useful cues for distinguishing plants from animals.

Previous findings suggest that part structure is an important feature for the speed of object identification [22] and for discriminating between abstract shapes [23,24]. Other findings indicate that part segmentation is available early in visual processing [25,26]. Together, this suggests that part structure might be a powerful cue for superordinate categorization of objects. Here, we ask to what extent humans make use of such cues.

We can also use differences in part structure to map out the superordinate category boundary along the spectrum of shapes spanning plants and animals. Previous studies did this by generating shapes along a continuum of superordinate classes (e.g., animals and leaves [27]) using a global shape-morphing algorithm to move between contours of different classes (e.g., [28]). This method of global shape-morphing, however, can often result in shapes that do not belong clearly to either of the superordinate classes and, moreover, does not permit testing the role of specific individual cues in participants' judgments easily. Here, we used our observations about systematic differences in part structures between animals and plants to morph continuously between classes by parametrically varying these cues. We focused on novel, generated objects to provide greater control over the parameters than is possible through photographs or silhouettes of natural shapes and to help reduce the influence of semantic knowledge and ground-truth category membership. We reasoned that while objects are generated to possess more or fewer features of a particular class, their novelty and the continuous variation of parameters ensures that the superordinate categorization task is nontrivial.

2. Experiment 1: Differences in Part Symmetry and Curvedness

To investigate the role of part symmetry as a superordinate categorization cue, we sought to generate pairs of shapes that differed only in that aspect—with one having symmetrical pairs of parts and the other having the same parts slightly displaced with respect to each other. Previous studies also found curvature to be a distinguishing factor between animals and plants (e.g., [9]); however, they analyzed and manipulated curvature at the image level. In contrast, we decided to investigate the role of curvature by creating pairs of shapes that only differed in the articulation of parts—namely, in straighter or curved limbs. To achieve these aims, we developed a generative algorithm based on shape skeletons [14,29,30] that allowed us to vary shape parameters such as the number of skeletal limbs (parts) and the width and length of each part (Figure 2). This algorithm was used to create all stimuli in this study.

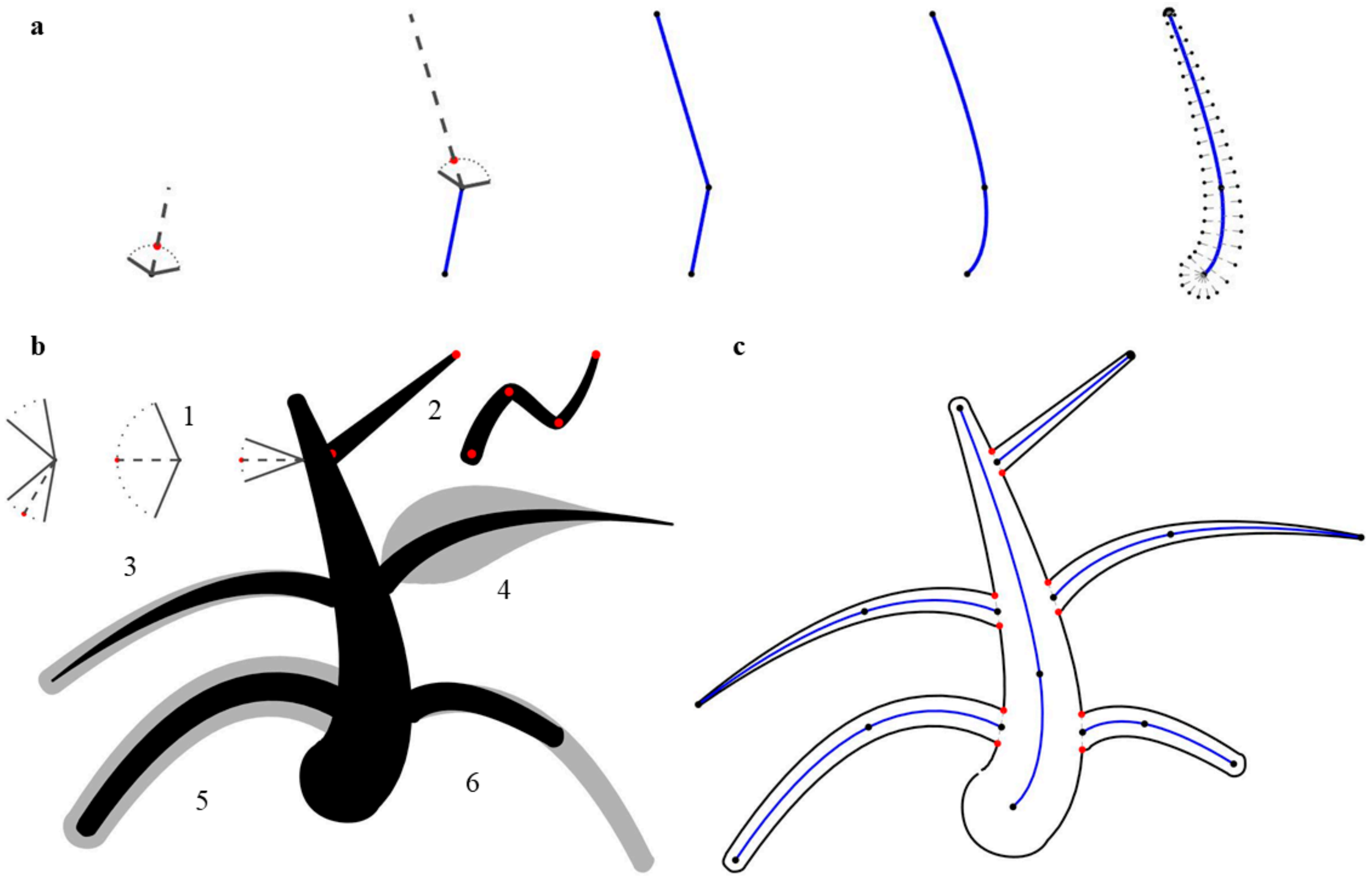

Figure 2. Overview of the shape creation algorithm. (**a**) Growth process of a part. From a starting point, we choose a random angle within a specified range (here, from −60° to +60° relative to a starting angle, which was pointed upwards for the first part, or orthogonal to the parent part in later parts). Then, we create a new point at a random distance deviating from the previous angle by a value chosen randomly from the same range as before. This process is repeated until the desired number of joints is created (the limbs of stimuli in this study ranged from three to five joints). The points are then connected by a line with each point defining a skeletal joint (panels 1–3). After smoothing the line (panel 4), we create silhouette points along the line with specified width values (panel 5). All parts are created analogously. (**b**) Examples of growth parameters. 1. The range of angles between joints can be varied from narrow to wide. 2. Number of joints. 3. Tapering. 4. For each joint, a width factor can be specified. Values between joints are interpolated. 5. Overall width relative to the parent. 6. Overall length relative to the parent. (**c**) Resulting skeletal representation. Joints (black dots) are shown together with skeletons of the main body and the parts (blue lines). Intersection points (red dots) between them are used for part segmentation.

2.1. Participants

The participants (n = 13) were recruited through the university mailing list. All the participants gave informed consent before the experiments in accordance with the Declaration of Helsinki. The procedures were approved by the Local Ethics Committee of the Department of Psychology and Sports Sciences of Justus Liebig University Giessen.

2.2. Materials and Methods

For Experiment 1, we created pairs of shapes with the simultaneous aims of maximizing visual similarity while varying symmetry of part pairs or curvedness of parts. Therefore, each shape pair consisted of the same main body, which was a simple elongated shape in upright orientation, with the same number of limbs growing from the main body of both shapes.

For the symmetry condition, one limb of each part pair in the second shape was displaced (top row of Figure 3)—that is, moved randomly up or down the main body by either 2.5%, 5.0%, 7.5%, 10.0%, or 12.5% of the main body's perimeter—creating pairs with subtle to very large asymmetries.

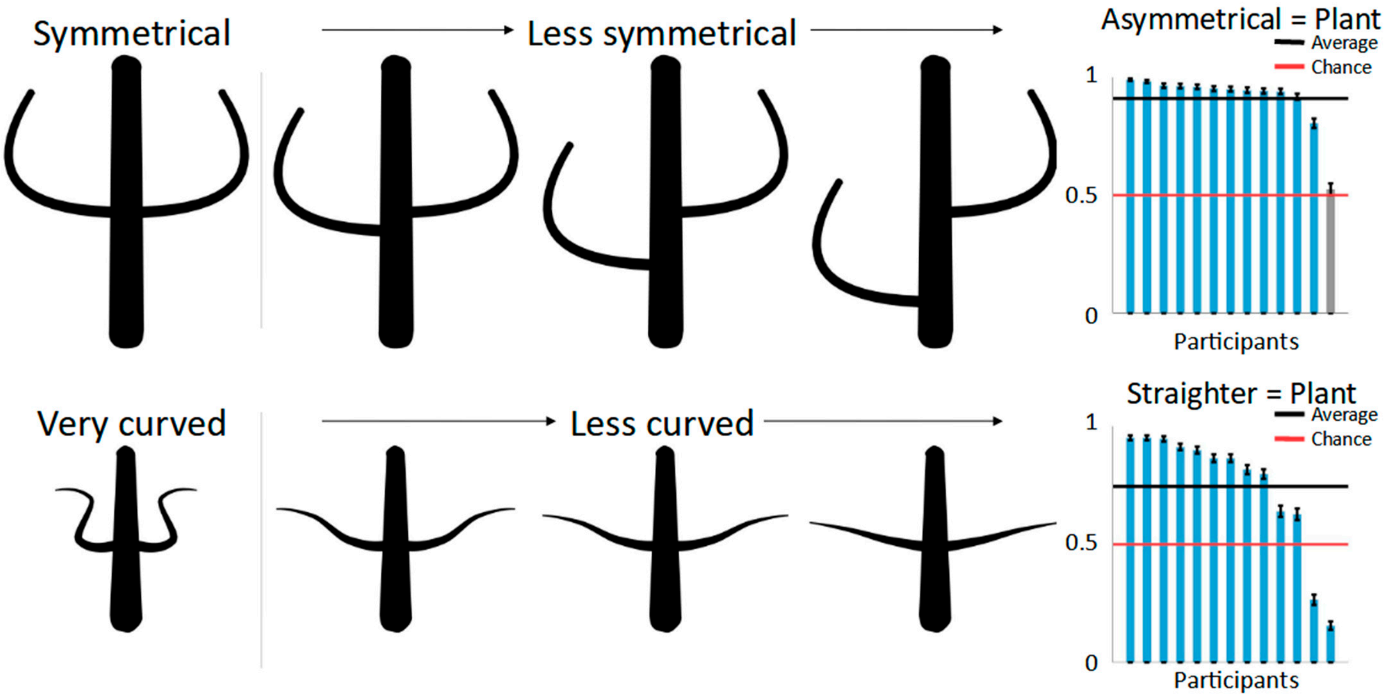

Figure 3. Stimuli and results of the 2-AFC task. The top row shows an object with symmetrical part pairs (left) and shapes becoming increasingly asymmetrical. Each shape pair in the 2-AFC task consisted of one symmetrical and one asymmetrical shape, and the participants were to choose the more plant-like shape. The results on a per-participant basis are plotted in the bar graph on the right with grey bars being not significantly different from chance (red line; evaluated by one-sided binomial test). The high average performance (black line) shows that asymmetry serves as a strong plant cue. The bottom row shows stimuli and findings for curvedness.

For the curvedness condition, we firstly created the limbs of the first shape with high curvedness (the bottom row of Figure 3), with angles between joints ranging from 50 to 100 degrees. For the second shape, these limbs were "straightened out" by moving the individual joints so that the resulting angle between all the joints was reduced by one of five factors (ranging from 0.05 to 0.25 in steps of 0.05). For example, multiplication using a factor of 0.25 with a limb's joints with an angle of 100° resulted in a limb with joints with an angle of 25° while keeping the distances between them the same. As a result, the parts of the second shape varied between somewhat curved to being almost completely straight.

For both conditions, the number of limbs attached to the main body varied from two to five. Twenty shape pairs were created for each number of limbs (4) and parameter values (5) resulting in $20 \times 4 \times 5 = 400$ stimuli per condition. In the 2-AFC task, the participants were presented with each shape pair and instructed to identify which of the two shapes belongs to the category of plants—and implicitly that the other belongs to the category of animals—with a press of a button. Thus, the same shape could never be plant-like and animal-like at the same time. They were further instructed that there were no right or wrong answers and, if unsure, to pick the response that fit the respective categories best. No further elaborations as to the research question were made. In spite of the shapes' relatively abstract visual appearance, the participants reported no problem in performing the task. The shape pairs were shown side by side in randomized order and with randomized left/right placement.

2.3. Results

Bar plots in Figure 3 show the results for both symmetry of part pairs (top) and curvedness (bottom) on a per-participant basis. The shapes with asymmetrical part pairs were judged to be more plant-like than their counterparts by all but one participant, resulting in a strong overall preference in 91% of the trials (one-sided binomial test: 50%; N = 5200, i.e., number of judgements; K (correct responses) = 4735, $p < 0.001$), showing that this aspect of part structure is indeed a strong differentiating cue between plants and animals. Curvedness also had a significant, although smaller, effect on categorization with less curved shapes seen as more plant-like in 75% of the trials (one-sided binomial test: 50%; N = 5200, i.e., number of judgements; K (correct responses) = 3879, $p < 0.001$)—with the exception of two participants who displayed the opposite pattern, with more curved shapes appearing more plant-like. This suggests that part structure is an effective cue for superordinate categorization, just like part articulation.

As described above, asymmetry and straightness varied from subtle to strong in five steps to test if larger differences in these two aspects would in turn lead to stronger categorization effects. A one-sided ANOVA for the different levels of asymmetry was significant ($F(4,395) = 3.49$, $p < 0.01$). Performing Tukey's post-hoc tests showed that the smallest asymmetry (displacement of 2.5% of the main body perimeter) was significantly different from both the second-largest and the largest displacements at $p < 0.05$ (10% and 12.5% displacement, respectively). Note that these differences are, however, relatively small in absolute numbers, with the smallest asymmetry seen as plant-like 88% of the time, compared to 93% and 92% for the two largest asymmetries. The ANOVA for the different levels of curvedness was not significant ($F4,395 = 0.76$, $p = 0.55$).

3. Experiment 2: Differences in Part Hierarchy between Animals and Plants

To investigate the role of part hierarchy in superordinate categorization, we generated shape pairs that differed in whether they exhibited higher-order limbs (i.e., sprouting) but were otherwise similar in appearance.

3.1. Participants

The participants ($n = 15$) were recruited through the university mailing list. All the participants gave informed consent before the experiments in accordance with the Declaration of Helsinki. Procedures were approved by the Local Ethics Committee of the Department of Psychology and Sports Sciences of Justus Liebig University Giessen.

3.2. Materials and Methods

Equivalently to Experiment 1, we created shape pairs only differing in the presence of second-order limbs. In each pair, the first shape consisted of the main body and two sets of symmetrical first-order limbs, emulating the look of an insect-like animal. The second shape consisted of the same main body and first-order limbs rearranged to imitate the typical part structure of a tree-like plant (by "transplanting" one set of first-order limbs onto another set, Figure 4a). As Experiment 1 showed that symmetry of part pairs and curvedness of limbs are strong categorization cues which might modulate potential effects of higher-order limbs, we created four groups of stimulus pairs from all the possible combinations of symmetrical/asymmetrical and curved/straight limbs (Figure 4a). Thus, we were able to test the effect of sprouting limbs on shapes that we expected to appear more plant-like (straight and asymmetrical limbs), animal-like (curved and symmetrical limbs), or in-between (straight and symmetrical as well as curved and asymmetrical limbs). For each of these four stimulus types, we created 125 shape pairs, resulting in 500 trials with no repetitions.

The participants were instructed to identify one shape that was more plant-like versus another shape that was more animal-like, meaning one shape could not be both more animal- and plant-like, as in Experiment 1. Apart from the stimuli, the experiment was identical to Experiment 1: the shape pairs were shown side by side in randomized order and with randomized left/right placement. The participants would then choose the more plant-like object with a press of a button (which also defined the other shape as more animal-like). This way we could test to what extent the presence of second-order limbs served as a cue to superordinate categorization of plants and animals.

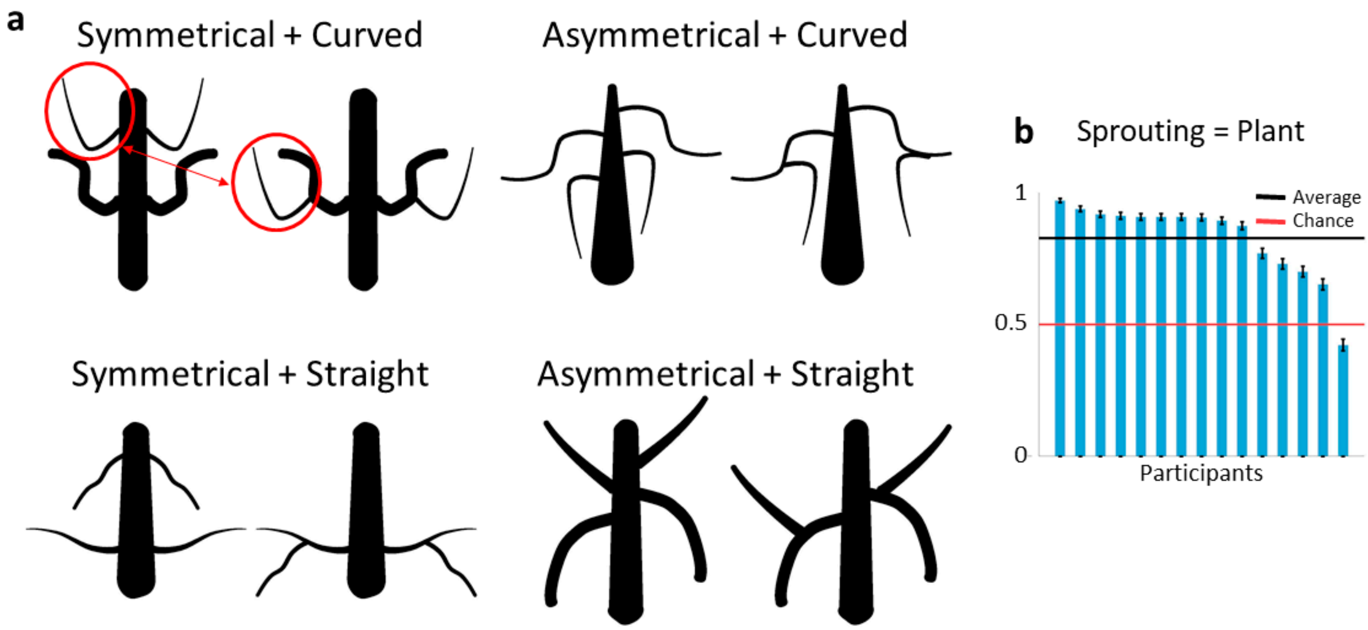

Figure 4. Sprouting limbs are a strong differentiation cue between plants and animals. (**a**) Types of stimuli created for the 2-AFC task combining symmetrical/asymmetrical and curved/straight limbs to make them appear more or less animal-/plant-like. (**b**) Results on a per-participant basis. All the participants' responses were significantly different from chance (red line; evaluated by one-sided binomial test). The participants were instructed to pick the more plant-like shape versus the more animal-like shape.

3.3. Results

Figure 4b shows the results for Experiment 2 on a per-participant basis. The shapes with sprouting limbs (i.e., with second-order limbs) were consistently judged to be more plant-like than their counterparts and, conversely, the shapes without sprouting were judged to be more animal-like (one-sided binomial test: 50%; N = 7500; K = 6205, $p < 0.001$). This illustrates the effectiveness of sprouting limbs in making a shape seem distinctively more plant- and less animal-like.

With respect to the additional parameters of symmetry and curvedness, we found no significant difference between the four stimulus types (ANOVA: F3,496 = 1.42, $p = 0.24$). This suggests that sprouting parts are the main driving force in this stimulus set independent of the other available cues.

4. Experiment 3: Morphing from Animals to Plants with Changes in Part Structure

After having established that features of part structure are effective cues for differentiating between animals and plants, we set out to investigate whether continuous changes in these features across the spectrum would result in corresponding changes in categorical decisions. Previous studies morphed between two end-states of different categories (e.g., animals and leaves [27]). However, by using global shape-morphing algorithms based on contours, in-between morphs did not always look like real objects belonging to any specific category. In contrast, we manipulated objects in a part-based fashion so that the number of parts as well as the parts themselves (apart from articulation) stayed the same throughout the morphing sequence. This way we could test whether differences in part structure alone are sufficient to define perceived category membership, without additional cues like texture or part identity (e.g., leaves or wings). Specifically, we asked whether the "animal-ness" of an object decreases continuously with the increase in the corresponding plant-like changes in part structure up to an ambiguous midpoint after which the object appears more like a plant. Alternatively, it would be possible that the chosen parameters were not sufficient to induce gradual changes in category membership. In that case, all shapes along the morphing spectrum might look either animal-like or plant-like (e.g., when changes in part curvedness would be interpreted as the same object with a different articulation of its limbs rather than another object from a different class). The often-subtle differences in part structure between morphing steps (in contrast to more blatant differences like the number of parts) allowed us to test the type of fine-grained categorical decisions that are often necessary in the real world, for example, when differentiating twigs from insects. Given that single parameters tend to explain little variance in the distinctions between superordinate classes [2,3], we decided to vary all three parameters from Experiments 1 and 2 together (symmetry of part pairs, curvedness of parts, and sprouting parts).

4.1. Participants

The participants (n = 14) were recruited through the university mailing list. All the participants gave informed consent before the experiments in accordance with the Declaration of Helsinki. The procedures were approved by the Local Ethics Committee of the Department of Psychology and Sports Sciences of Justus Liebig University Giessen.

4.2. Materials and Methods

We morphed along three parameters at the same time: (1) the symmetry of part pairs, varying between symmetrical and asymmetrical; (2) the curvedness of parts, varying between curved and almost straight; (3) the organization of parts, varying in two different ways (Figure 5). The first type of stimuli featured a more plant-like form of sprouting (growing sprouting condition): sprouting limbs grew from random points on their parent limb and were articulated independently, without making a symmetrical pair with the sprout on the corresponding mirrored parent part. Furthermore, we varied the size of the sprouts, with second-order limbs growing from non-existent to relatively large. The second type of stimuli featured sprouting limbs with more animal-like characteristics (symmetrical sprouting condition): second-order limbs formed symmetrical pairs (if the parent limb was also part of a symmetrical pair) and were of the same size throughout the morphing sequence—emulating real-world second-order limbs of animals like claws or thumbs. In both (growing sprouting, symmetrical sprouting) conditions, the sprouting limbs changed in the same way as the first-order limbs did along the morphing sequence—they became straighter and more asymmetrical. Along with these parameters, shapes varied in the number of their first-order limbs (from two to six; held constant along the morphing sequence).

With these parameters, we created 110 sequences of five morph levels from animal- to plant-like, resulting in 550 shapes in total. Each shape consisted of a vertically oriented main body, with a different number of first-order limbs (ranging from two to six) that formed symmetrical pairs. With odd numbers of limbs, the additional limb would not have a symmetrical counterpart.

All the 550 shapes were shown in randomized order one by one so that the morphing sequences were not apparent to the viewer. The participants were asked to rate each shape on a continuous scale from animal- to plant-like by moving a slider, with the midpoint representing a shape that is equally animal- and plant-like.

The different types of stimuli allowed us to investigate two questions. First, can we generalize the previous findings of Wilder et al. (2011) by showing that a larger number of limbs not only makes an object look more like a leaf, but also generally more like a plant? If that were the case, we should find that shapes with more first-order limbs are judged to be more plant-like compared to shapes with fewer first-order limbs. Second, are shapes with an odd number of first-order limbs, and therefore no bilateral symmetry, more biased towards the plant side than shapes with an even number of limbs?

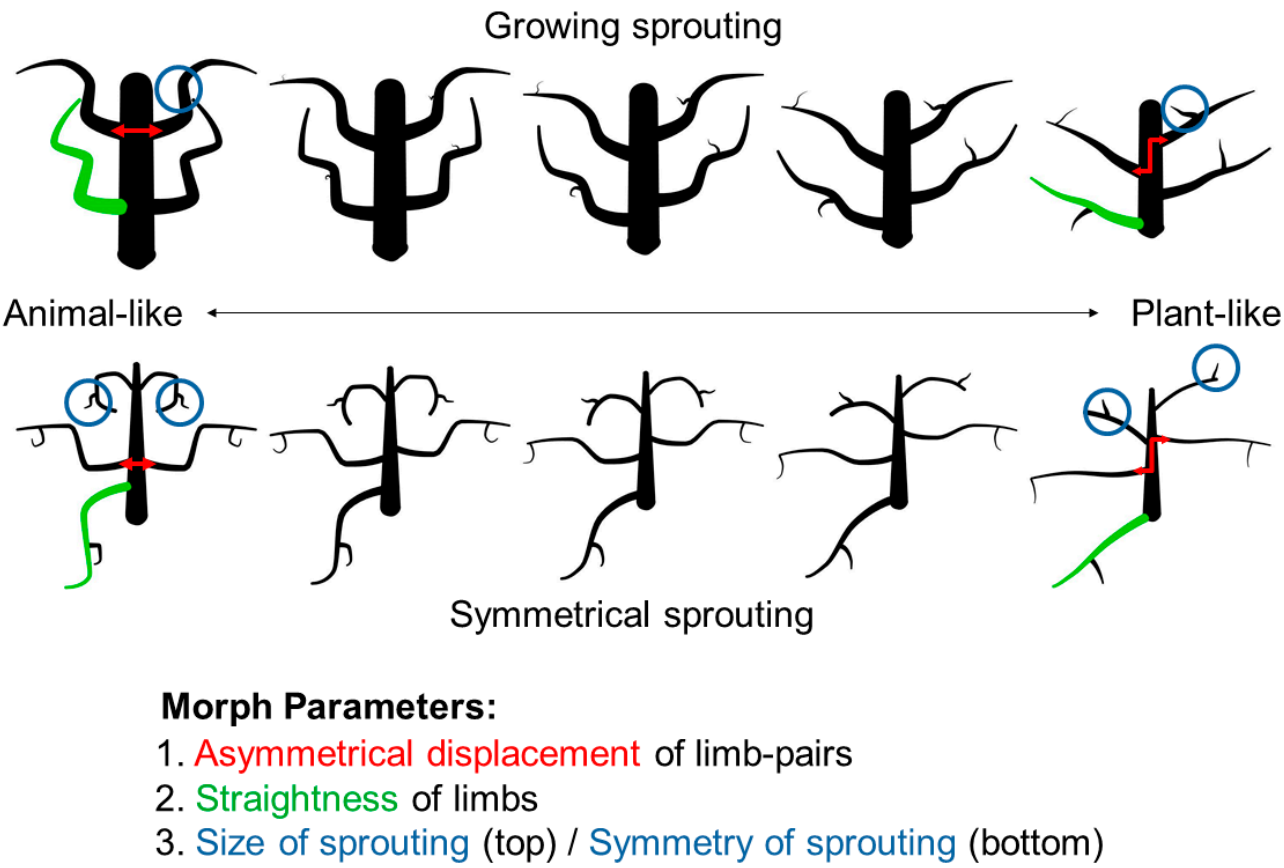

Figure 5. Morphing sequences from animal- to plant-like and the corresponding parameters. The top row shows an example sequence for the "growing sprouting" condition and the bottom row for the "symmetrical sprouting" condition. Changes in three morph parameters are illustrated: symmetrical limb pairs are increasingly displaced (red), parts get continuously straighter (green), and sprouting parts (blue) either grow (from non-existent to largest) from random starting points (top row of shapes) or are always the same size but make up symmetrical pairs where possible (bottom row).

4.3. Results and Discussion

The participants' responses were closely related to variations in our parameter space, in which symmetry, curvedness, and sprouting were manipulated simultaneously. Shapes with more animal-like parts (more curved, symmetrical, and with fewer prominent second-order limbs) were perceived as more animal-like and vice versa for more plant-like parts (Figure 6). To summarize the sigmoidal trends in the data, we fit psychometric functions to all the responses using psignifit [31]. Overall, the responses were biased towards the plant end of the spectrum (first panel of Figure 6) with the average categorization value of 0.59 (most animal-like = 0, most plant-like = 1)—documented by a one-sample *t*-test comparing all the responses with the midpoint of 0.5 ($t(7699) = 28.47$, $p < 0.001$). This bias might be explained by the existence of (albeit small) sprouting limbs in all the conditions, except for the most animal-like stimulus type in the "growing sprouting" condition (meaning that 90% of the stimuli had some form of second-order limbs). Furthermore, only the most animal-like step had complete symmetry among the limb pairs, with the other four steps introducing asymmetry, a potent plant cue, potentially adding to the plant bias. Moreover, it is also worth noting that we did not systematically manipulate the local symmetry of second-order limbs meaning that they may have more closely resembled plant branches than animal limbs.

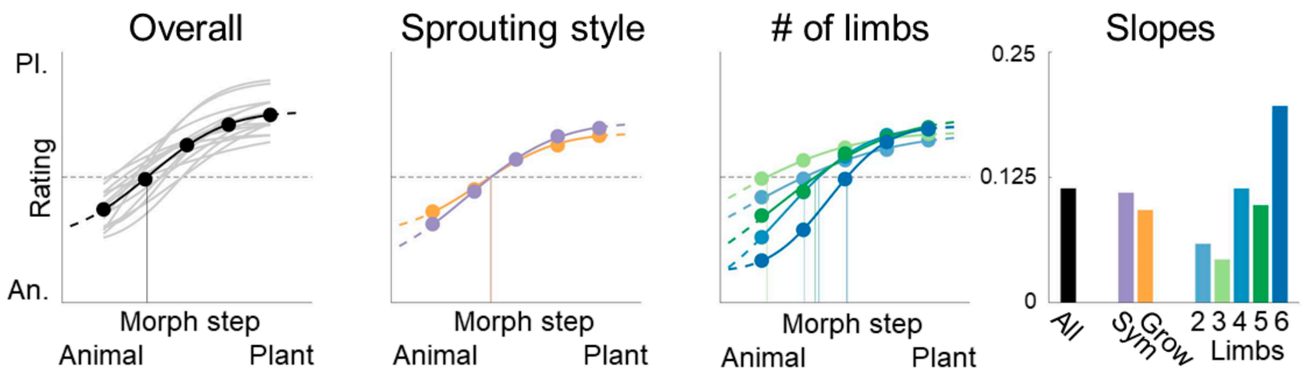

Figure 6. Morph sequences agree with human categorizations. Panel 1, aggregated responses (y-axis, Pl. = plant, An. = Animal) as a function of morphing steps (x-axis), with each participant's average response in faint grey, shows the overall agreement between the morphing steps and the observers' categorizations. The midpoint of the animal/plant spectrum used by observers is indicated by a dashed line. Panel 2, different sprouting styles (growing sprouting in orange and symmetrical sprouting in purple) had very similar effects on categorization. Panel 3, the more parts a shape had, the stronger was the difference between the morphing steps. Panel 4, slopes of curves in Panels 1–3. Noticeably, the parts with an odd number of non-main body limbs have a smaller slope and are more biased towards plant-likeness compared to the shapes with an even number of limbs.

Different sprouting styles (second panel of Figure 6) elicited highly similar responses, showing that they had roughly the same impact on categorization. An ANOVA yielded no significant difference between the two sprouting styles ($F(1,548) = 0.01$, $p = 0.92$). To establish whether we had evidence for no difference between the sprouting styles, we calculated a scaled JZS Bayes factor using a Jeffrey–Zellner–Siow prior (Cauchy distribution on effect size) with a default scale factor of 0.707 [32], resulting in a BF10 of 0.03. This measure describes the probability of the data given H1 relative to H0, where a BF10 < 0.1 can be considered "strong evidence" for H0 [32]. This suggests that, across the complete morphing sequence, it did not matter whether sprouting limbs grew from random points and varied in size (growing sprouting) or whether they formed symmetrical pairs that did not change in size (symmetrical sprouting).

Finally, increasing numbers of limbs attached to the main body (third and fourth panels in Figure 6) increased the slope of the fit function, indicating a stronger distinction between animal- and plant-like shapes. While the number of parts did not affect the plant side of the morphing spectrum much, it strongly affected the judgement of animal-like shapes. This is corroborated by separate ANOVAs for each morphing step, showing significant differences within steps 1 (most animal-like) to 3 (midpoint), but not within steps 4 and 5 (most plant-like; Bonferroni-adjusted *p*-value of 0.01 for five tests). Consequently, the number of limbs only affected judgement of ambiguous and animal-like shapes—with animal-like shapes with fewer limbs judged as much more plant-like compared to animal-like shapes with many limbs. Furthermore, an odd number of limbs created an overall more plant-like impression along the morphing sequence (fourth panel in Figure 6), with a smaller slope compared to shapes with an even number of limbs (one limb either more or less), presumably because animals canonically have an even number of limbs.

5. Discussion

Visually distinguishing between superordinate classes is a vital yet challenging skill. Different members of the same superordinate class may share certain features, but also differ from one another in other highly salient ways. Somehow, the visual system must determine which feature differences are relevant for superordinate classification and which otherwise significant differences in appearance ought to be ignored. How humans achieve this remains unclear. Most previous studies investigated local object features such as eyes and mouths (e.g., [1]), individual parts such as tails [7], global features such as canonical animal posture [8], or overall curvature [9–11] without considering the perceptual organization of objects. However, since organic objects grow in specific ways according to particular laws (e.g., [18,20]), the resulting visual part structure is highly constrained by biological class, therefore specifying a potentially powerful cue for superordinate categorization.

Based on this observation, we conducted two categorization experiments with pairs of shapes that were identical except for the symmetry of part pairs, the curvedness of parts, or the presence of second-order limbs (Experiments 1 and 2). We found that both symmetry and curvedness were effective cues for differentiating between animals and plants. The presence of second-order limbs was also used to categorize novel objects as plants or animals. Together, these findings suggest that the perceptual organization of objects—the spatial arrangement and relations between their parts—helps us to distinguish between superordinate classes of animals and plants, and potentially other superordinate categorizations.

In the final experiment, we used these newly identified categorization cues to create sequences of shapes morphing from animal- to plant-like in a continuous, part-based fashion. Since superordinate categorization tends to be affected by multiple factors at once [2,3], all three features varied simultaneously. We also included two different styles of sprouting, one closer to the biological reality of plants (growing sprouting condition) and the other emulating animal-like second-order limbs (symmetrical sprouting). We found that the three categorization cues predicted category membership very well even if the two sprouting styles did not create appreciable differences in categorization. Additionally, the more parts a shape had, the stronger the perceived difference across the morphing spectrum, especially for shapes with animal characteristics: the shapes with fewer than four parts were seen as more plant-like across the spectrum compared to shapes with four or five limbs. It is interesting to speculate whether this might have been because canonically mammals possess four limbs, whereas plants deviate more often from this number. Finally, shapes with an odd number of first-order limbs were more biased towards the plant end of the spectrum regardless of curvedness or symmetry of part pairs compared to even-numbered counterparts—showing a further effect of structural symmetry on categorization (i.e., the symmetry of the points of attachment along the length of the body). Overall, our findings show that even small differences in part structure can change the superordinate categorization of objects as opposed to more blatant changes like adding new parts [27] or distinctive features [1,7], emphasizing just how subtle decisive category features can be.

Given the speed of visual part segmentation [25,26] and categorization [1] in humans, it is possible that the perceptual organization of objects is analyzed early in visual processing as well. In fact, a lot of information about part structure would be obtained already in the process of segmenting an object into its parts—as, for example, indicated by studies on the early availability of symmetry information (e.g., [15,33,34]) and an extended neural network processing that information [35,36].

Previous studies developed concepts similar to our definition of part structure, all of which put emphasis on the information about relational aspects of object parts (e.g., [37–42]), even though some of these approaches do not discuss part segmentation as a perceptual problem. What many of these different concepts have in common is the reduction of visual information about object parts to a "neighborhood tree", tree-like graph representations that mainly encode which parts are connected to which other parts. While this information alone might be sufficient to match many similar objects (e.g., [40]), thereby showing that even the most basic form of part structure is rich in information, this type of representation is not sufficient to explain our current findings. For example, a graph representation is insufficient to differentiate between the symmetrical and asymmetrical shape pairs of Experiment 1 where the connections between parts (as well as individual parts) are identical. Thus, an accurate account of the human visual representation of part structure has to include not only the connections between parts, but also relational information of unconnected parts (e.g., whether a limb has a mirror-symmetric counterpart on the other side of its parent).

Medial axes representations (e.g., [14]) are a promising concept to encode these types of relational features (e.g., [43,44]).

Our results show that part structure is used to differentiate between superordinate classes, specifically those differing in their patterns of biological growth. It is highly feasible that we also represent categories of artificial objects with respect to their part structure: even though that structure might be less constrained, or constrained by alternative organizing principles (e.g., a helicopter is not the result of a growth process), most objects are composed in systematic ways according to particular laws that produce regularities in their part structure. It is also worth noting that in some cases there may even be similarities with classes of natural objects. For example, the part structure of an airplane is similar to that of a bird, given that both "solve" the problem of flight. This similarity in part structure between biological and artificial objects is further strengthened by the common constraints of structural integrity that only allow for certain viable arrangements. An exciting direction for future work is to identify other constraints, both universal and restricted to specific categories, and analyze how they affect the perceptual organization of objects, and consequently their mental representation and categorization.

Author Contributions: Conceptualization, H.T., F.S. and R.W.F.; methodology, H.T., F.S. and R.W.F.; software, H.T.; validation, F.S. and R.W.F.; formal analysis, H.T.; investigation, H.T.; resources, R.W.F.; data curation, H.T.; writing—original draft preparation, H.T.; writing—review and editing, F.S. and R.W.F.; visualization, H.T., F.S. and R.W.F.; supervision, F.S. and R.W.F.; project administration, H.T.; funding acquisition, R.W.F. All authors have read and agreed to the published version of the manuscript.

Funding: Research funded by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) (project No. 222641018-SFB/TRR 135 TP C1), by the European Research Council (ERC) Consolidator Award "SHAPE" (project No. ERC-2015-CoG-682859), and by the Research Cluster "The Adaptive Mind" funded by the Excellence Program of the Hessian Ministry of Higher Education, Science, Research, and Art.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of the Department of Psychology and Sports Sciences of Justus Liebig University Giessen (LEK-FB06; protocol code 2016-0007, approved on 18 April 2016).

Informed Consent Statement: Informed consent was obtained from all the subjects involved in the studies.

Data Availability Statement: All the stimuli will be made available on Zenodo upon acceptance.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Fabre-Thorpe, M. The characteristics and limits of rapid visual categorization. *Front. Psychol.* **2011**, *2*, 243. [CrossRef] [PubMed]
2. Jozwik, K.M.; Kriegeskorte, N.; Mur, M. Visual features as stepping stones toward semantics: Explaining object similarity in IT and perception with non-negative least squares. *Neuropsychologia* **2016**, *83*, 201–226. [CrossRef] [PubMed]
3. Schmidt, F.; Hegele, M.; Fleming, R.W. Perceiving animacy from shape. *J. Vis.* **2017**, *17*, 10. [CrossRef]
4. de Beeck, H.O.; Béatse, E.; Wagemans, J.; Sunaert, S.; Van Hecke, P. The Representation of Shape in the Context of Visual Object Categorization Tasks. *NeuroImage* **2000**, *12*, 28–40. [CrossRef] [PubMed]
5. Gillebert, C.R.; Op de Beeck, H.P.; Panis, S.; Wagemans, J. Subordinate Categorization Enhances the Neural Selectivity in Human Object-selective Cortex for Fine Shape Differences. *J. Cogn. Neurosci.* **2009**, *21*, 1054–1064. [CrossRef] [PubMed]
6. De Beeck, H.P.O.; Torfs, K.; Wagemans, J. Perceived Shape Similarity among Unfamiliar Objects and the Organization of the Human Object Vision Pathway. *J. Neurosci.* **2008**, *28*, 10111–10123. [CrossRef]
7. Tversky, B. Parts, partonomies, and taxonomies. *Dev. Psychol.* **1989**, *25*, 983. [CrossRef]
8. Delorme, A.; Richard, G.; Fabre-Thorpe, M. Key visual features for rapid categorization of animals in natural scenes. *Front. Psychol.* **2010**, *1*, 21. [CrossRef]
9. Long, B.; Störmer, V.S.; Alvarez, G.A. Mid-level perceptual features contain early cues to animacy. *J. Vis.* **2017**, *17*, 20. [CrossRef]
10. Long, B.; Yu, C.-P.; Konkle, T. Mid-level visual features underlie the high-level categorical organization of the ventral stream. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E9015–E9024. [CrossRef]
11. Perrinet, L.U.; Bednar, J.A. Edge co-occurrences can account for rapid categorization of natural versus animal images. *Sci. Rep.* **2015**, *5*, 11400. [CrossRef] [PubMed]

12. Hoffman, D.D.; Richards, W.A. Parts of recognition. *Cognition* **1984**, *18*, 65–96. [CrossRef]
13. De Winter, J.; Wagemans, J. Segmentation of object outlines into parts: A large-scale integrative study. *Cognition* **2006**, *99*, 275–325. [CrossRef] [PubMed]
14. Feldman, J.; Singh, M. Bayesian estimation of the shape skeleton. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 18014–18019. [CrossRef]
15. Koning, A.; Wagemans, J. Detection of symmetry and repetition in one and two objects: Structures versus strategies. *Exp. Psychol.* **2009**, *56*, 5–17. [CrossRef]
16. Kubilius, J.; Sleurs, C.; Wagemans, J. Sensitivity to Nonaccidental Configurations of Two-Line Stimuli. *i-Perception* **2017**, *8*, 2041669517699628. [CrossRef]
17. Kubilius, J.; Wagemans, J.; De Beeck, H.P.O. Encoding of configural regularity in the human visual system. *J. Vis.* **2014**, *14*, 11. [CrossRef]
18. Stevens, P.S. *Patterns in Nature*, 1st ed.; Little Brown & Co.: Boston, MA, USA, 1974.
19. Reece, J.B.; Urry, L.A.; Cain, M.L.; Wasserman, S.A.; Minorsky, P.V.; Jackson, R.B. *Campbell Biology*; Pearson Boston: Boston, MA, USA, 2014.
20. Newell, A.C.; Shipman, P.D. Plants and fibonacci. *J. Stat. Phys.* **2005**, *121*, 937–968. [CrossRef]
21. Ons, B.; Wagemans, J. Generalization of Visual Shapes by Flexible and Simple Rules. *Seeing Perceiving* **2012**, *25*, 237–261. [CrossRef]
22. Cave, C.B.; Kosslyn, S.M. The Role of Parts and Spatial Relations in Object Identification. *Perception* **1993**, *22*, 229–248. [CrossRef]
23. Mash, C. Multidimensional shape similarity in the development of visual object classification. *J. Exp. Child Psychol.* **2006**, *95*, 128–152. [CrossRef] [PubMed]
24. Lowet, A.S.; Firestone, C.; Scholl, B.J. Seeing structure: Shape skeletons modulate perceived similarity. *Atten. Percept. Psychophys.* **2018**, *80*, 1278–1289. [CrossRef] [PubMed]
25. Wolfe, J.M.; Bennett, S.C. Preattentive Object Files: Shapeless Bundles of Basic Features. *Vis. Res.* **1997**, *37*, 25–43. [CrossRef]
26. Xu, Y.; Singh, M. Early computation of part structure: Evidence from visual search. *Percept. Psychophys.* **2002**, *64*, 1039–1054. [CrossRef]
27. Wilder, J.; Feldman, J.; Singh, M. Superordinate shape classification using natural shape statistics. *Cognition* **2011**, *119*, 325–340. [CrossRef]
28. Destler, N.; Singh, M.; Feldman, J. Shape discrimination along morph-spaces. *Vis. Res.* **2019**, *158*, 189–199. [CrossRef]
29. Blum, H. Biological shape and visual senses (Part I). *J. Theor. Biol.* **1973**, *38*, 205–287. [CrossRef]
30. Feldman, J.; Singh, M.; Briscoe, E.; Froyen, V.; Kim, S.; Wilder, J. An Integrated Bayesian Approach to Shape Representation and Perceptual Organization. In *Shape Perception in Human and Computer Vision*; Springer: London, UK, 2013; pp. 55–70. [CrossRef]
31. Schütt, H.H.; Harmeling, S.; Macke, J.H.; Wichmann, F.A. Painfree and accurate Bayesian estimation of psychometric functions for (potentially) overdispersed data. *Vis. Res.* **2016**, *122*, 105–123. [CrossRef]
32. Rouder, J.N.; Speckman, P.L.; Sun, D.; Morey, R.D.; Iverson, G. Bayesian t tests for accepting and rejecting the null hypothesis. *Psychon. Bull. Rev.* **2009**, *16*, 225–237. [CrossRef]
33. Schmidt, F.; Schmidt, T. Rapid processing of closure and viewpoint-invariant symmetry: Behavioral criteria for feedforward processing. *Psychol. Res.* **2014**, *78*, 37–54. [CrossRef]
34. Machilsen, B.; Pauwels, M.; Wagemans, J. The role of vertical mirror symmetry in visual shape detection. *J. Vis.* **2009**, *9*, 11. [CrossRef] [PubMed]
35. Alp, N.; Kohler, P.J.; Kogo, N.; Wagemans, J.; Norcia, A.M. Measuring Integration Processes in Visual Symmetry with Frequency-Tagged EEG. *Sci. Rep.* **2018**, *8*, 6969. [CrossRef] [PubMed]
36. Bertamini, M.; Silvanto, J.; Norcia, A.M.; Makin, A.D.; Wagemans, J. The neural basis of visual symmetry and its role in mid- and high-level visual processing. *Ann. N. Y. Acad. Sci.* **2018**, *1426*, 111–126. [CrossRef] [PubMed]
37. Siddiqi, K.; Shokoufandeh, A.; Dickinson, S.J.; Zucker, S.W. Shock graphs and shape matching. *Int. J. Comput. Vis.* **1999**, *35*, 13–32. [CrossRef]
38. Zhu, S.C.; Yuille, A.L. FORMS: A flexible object recognition and modelling system. *Int. J. Comput. Vis.* **1996**, *20*, 187–212. [CrossRef]
39. Marr, D.; Nishihara, H.K. Representation and recognition of the spatial organization of three-dimensional shapes. *Proc. R. Soc. Lond. Ser. B Biol. Sci.* **1978**, *200*, 269–294.
40. Latecki, L.J.; Lakamper, R. Shape similarity measure based on correspondence of visual parts. *IEEE Trans. Pattern Anal. Mach. Intell.* **2000**, *22*, 1185–1190. [CrossRef]
41. Kupeev, K.Y.; Wolfson, H.J. On shape similarity. In Proceedings of the 12th International Conference on Pattern Recognition, Jerusalem, Israel, 9–13 October 1994; Volume 1, pp. 227–231.
42. Connell, J.H.; Brady, M. Generating and generalizing models of visual objects. *Artif. Intell.* **1987**, *31*, 159–183. [CrossRef]
43. Ayzenberg, V.; Kamps, F.S.; Dilks, D.D.; Lourenco, S.F. Skeletal representations of shape in the human visual cortex. *Neuropsychologia* **2021**, *164*, 108092. [CrossRef]
44. Ayzenberg, V.; Lourenco, S.F. Skeletal descriptions of shape provide unique perceptual information for object recognition. *Sci. Rep.* **2019**, *9*, 9359. [CrossRef]