

## Reviewed Preprint

v1 • May 7, 2026

Not revised

## ✉ For correspondence:

sparun@iisc.ac.in

## Competing interests: No

competing interests declared

## Funding: See page 23

Reviewing editor: Iris Groen,  
University of Amsterdam,  
Netherlands

© 2026, Cherian et al. This article is distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use and redistribution provided that the original author and source are credited.

# Do monkeys see the way we do? Qualitative similarities and differences between monkey and human perception

Thomas Cherian<sup>1</sup>, Georgin Jacob<sup>2,1</sup>, SP Arun<sup>1,2</sup> ✉

<sup>1</sup>Centre for Neuroscience, Indian Institute of Science, Bengaluru, India • <sup>2</sup>Department of Electrical Communication Engineering, Indian Institute of Science, Bengaluru, India

## eLife Assessment

This **valuable** study presents a comprehensive comparison of human and macaque monkey behavior across a range of visual perceptual phenomena. The use of a unified oddball visual search paradigm enables direct cross-species comparison while minimizing task-related confounds. It provides **solid** evidence that visual perception is largely similar between these two species, with some interesting exceptions. These insights into qualitative and quantitative differences between species are relevant for evaluating macaques as a model organism for understanding human vision.

<https://doi.org/10.7554/eLife.110743.1.sa3>

## Abstract

Monkeys are widely used as model organisms for human vision and cognition. While their anatomy and physiology strongly correspond to humans, it is unclear to what extent their perception matches with ours. Previous studies have evaluated specific aspects of perception, after extensive training on customized tasks that could have altered their perception. To resolve these issues, we trained monkeys to perform an oddball visual search task on natural images, tested them on carefully controlled images to detect a variety of perceptual phenomena, and compared them with humans. This revealed a number of qualitative similarities and differences. Like humans, monkeys showed similar object relations, Weber's law and amodal completion. However, unlike humans, monkeys did not show mirror confusion or a global advantage. These findings represent a first comprehensive evaluation of visual perception in monkeys and humans, revealing the limitations of monkeys as a model organism for human vision.

## Significance statement

Monkeys are widely used to study for vision and cognition, but is their perception like ours? Most previous studies have tested only single aspects of perception in monkeys, often after extensive training which could have altered their perception. Here, we overcame these limitations by training monkeys on a highly general oddball visual search task, testing them on many perceptual phenomena and comparing them with humans on the same tasks. Our results show that monkeys do indeed see the way we do, but with important differences that highlight their limitations as a model for human vision.

## Introduction

*The Master sees things as they are,  
without trying to control them.*

-- *Tao Te Ching* (Mitchell, 1988 [↗](#))

Macaque monkeys are widely used as model organisms to study the neural basis of cognition (Passingham, 2009 [↗](#); Roelfsema and Treue, 2014 [↗](#); Buffalo et al., 2019 [↗](#)). Specifically in the case of vision, monkeys have similar visual acuity and contrast thresholds (Weinstein and Grether, 1940 [↗](#); De Valois et al., 1974b [↗](#), 1974a [↗](#)), and have extensively characterized visual regions whose homologues continue to be elaborated and studied in humans (Quiroga et al., 2005 [↗](#); Hiramatsu et al., 2011 [↗](#); Tang et al., 2014 [↗](#)). Moreover, humans and monkeys share similar object representations at a global level (Kriegeskorte et al., 2008a [↗](#); Rajalingham et al., 2015 [↗](#), 2018 [↗](#)), and their visual cortexes also show similar segregation into functional sub-units (Bell et al., 2009 [↗](#)). These converging lines of evidence confirm that macaque monkeys are indeed an excellent model for human visual cognition.

However, our understanding of any model is incomplete without knowing its limitations. In this case, it is not clear if there are any fundamental differences in visual perception between monkeys and humans. Do monkeys see the way we do? A comprehensive answer to this question would require comparing monkeys and humans directly on a battery of tests of visual perception. However, this can be extremely challenging in practice. Humans can easily be instructed verbally to perform a variety of visual tasks, but monkeys require extensive behavioural training on each specific visual task. As a result, most studies have tested primates on various specific aspects of perception, as detailed below.

At a coarse level, monkeys trained on categorization tasks show similar performance on natural objects as humans (Rajalingham et al., 2015 [↗](#), 2018 [↗](#)). They also show similar gaze patterns on normal faces and illusory faces (Dahl et al., 2010 [↗](#); Weldon et al., 2013 [↗](#); Beran et al., 2017 [↗](#); Taubert et al., 2017 [↗](#)). But on other measures, humans and monkeys have shown different patterns of performance. While humans are faster to report shape at the global level compared to the local level (Navon, 1977 [↗](#); Kimchi, 1992 [↗](#)), there is conflicting evidence for this effect in monkeys: monkeys trained on shape categorization responded faster to global shape (Tanaka and Fujita, 2000 [↗](#)), but in other cases responded faster to local shape (Deruelle and Fagot, 1997 [↗](#); Fagot and Tomonaga, 1999 [↗](#); Nielsen et al., 2006 [↗](#)). These differences could arise due to task or stimulus differences.

At a more fine-grained level, macaque monkeys have been tested on a variety of visual illusions. Monkeys trained on size matching tasks are susceptible to size illusions like we do (Behar and Samuel, 1982 [↗](#); Fujita, 1996 [↗](#), 1997 [↗](#); Barbet and Fagot, 2002 [↗](#); Parron and Fagot, 2007 [↗](#); Agrillo et al., 2019 [↗](#)), although they are less sensitive to context than humans (Fujita, 1997 [↗](#)). Monkeys trained on length-matching tasks show the Muller-Lyer illusion (Tudusciuc and Nieder, 2010 [↗](#)). Monkeys trained on image matching tasks show mirror confusion on abstract shapes (Riopelle et al., 1964 [↗](#); Washburn, 1993 [↗](#)). Monkeys trained on numerosity tasks show a Weber's law for numerosity (Nieder and Miller, 2003 [↗](#); Jordan and Brannon, 2006 [↗](#)) and biased responses to regular dot configurations (Beran, 2006 [↗](#)). However, they do not show certain grouping illusions (Parrish et al., 2019 [↗](#)).

## Overview of this study

A major limitation of the above studies is that each finding is based on a highly specialized training regimen tested on a different group of individuals. As a result, we lack a comprehensive evaluation of visual perception in monkeys and how this might compare to humans. Here, we overcame this limitation by training monkeys on a generic oddball visual search task, evaluating multiple perceptual phenomena within the same overall task and comparing these patterns of performance to humans.

In the oddball search task, the participant has to touch the oddball item in an array containing one oddball item and multiple identical distractors (Figure 1 [↗](#)). This is an extremely easy task for human participants and required minimal training for monkeys to reach high levels of

performance (see Methods). The response time of an oddball search is a natural index of the dissimilarity between the target and distractors (Arun, 2012 [↗](#); Pramod and Arun, 2014 [↗](#), 2016a [↗](#); Sunder and Arun, 2016 [↗](#)).

Importantly, oddball search can be used to measure a variety of perceptual phenomena depending on the specific images used (Jacob et al., 2021b [↗](#)). For example, it is well known that humans experience confusion between vertical mirror images. In visual search, mirror confusion is evidenced by longer response times for finding a target among distractors when the target is a vertical mirror image of the distractor than when it is a horizontal mirror image. Likewise, according to the global advantage effect, humans can report global shape faster than local shape. In visual search, the global advantage effect manifests as faster response times for targets differing in global shape compared to when they differ in local shape.

We identified several perceptual phenomena for testing on both humans and monkeys, which included those used recently to compare humans with deep neural networks (Jacob et al., 2021b [↗](#)). These included testing coarse object similarities, Weber's law for luminance, amodal completion of occluded objects, mirror-image confusion, global advantage and part processing. These revealed both qualitative similarities as well as differences between human and monkey perception.

## Results

To obtain a comprehensive comparison of visual perception in monkeys and humans, we tested 3 monkeys and 24 humans on an oddball visual search task (Figure 1 [↗](#)) across several experiments. Each experiment involved a different set of images designed to address a specific aspect of visual processing, but importantly, the task remained unchanged for the participants: they had to always locate an oddball item in the search array. This design ensured that task and performance considerations of the participants (humans and monkeys) were completely unrelated to the goals of each experiment.

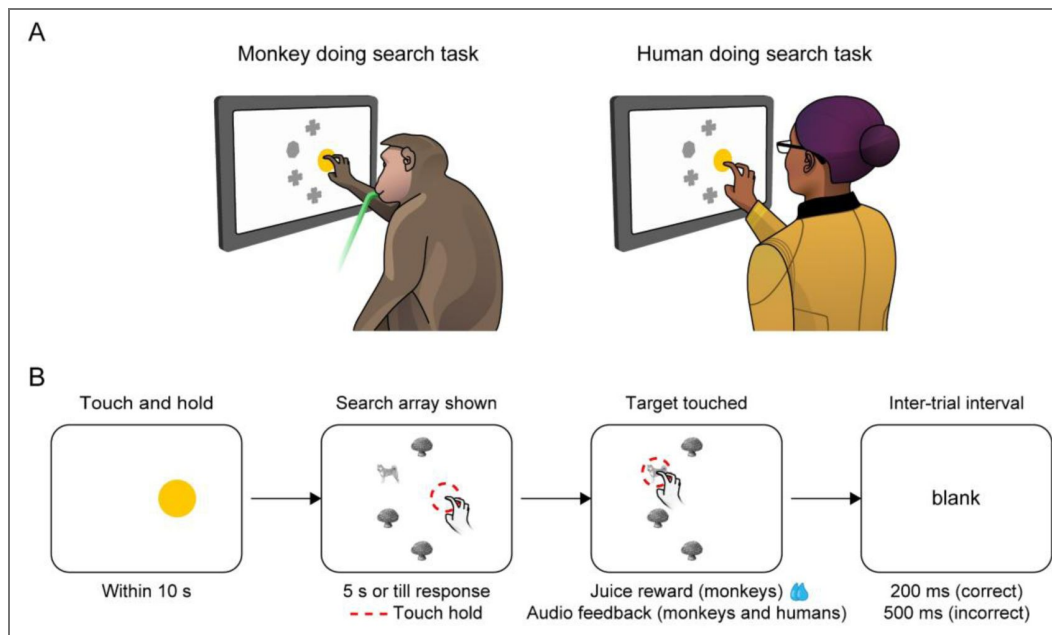
### Experiment 1: Do monkeys have object representations similar to humans?

In Experiment 1, we asked whether monkeys and humans perceive objects similarly at the coarse level. Since visual search is more difficult when the target is similar to the distractors, the reciprocal of the response time can be taken as a measure of perceived dissimilarity between the target and distractor. We therefore asked how well the perceived dissimilarities between objects for monkeys matched those obtained from humans.

To this end, we chose a set of animate and inanimate objects (Figure 2A [↗](#)) and created search arrays involving all possible pairs of objects. We analysed previously published data from 12 human participants on an oddball search task (see Methods). Human reaction times were highly consistent ( $r_{sh} = 0.90$ ,  $p_{sh} < 0.005$ ,  $n = 28$  search pairs; see Methods). To calculate perceived dissimilarity, we took the reciprocal of the average response time (across trials and participants) for each image-pair. The resulting pairwise dissimilarities are depicted in Figure 2B [↗](#). It can be seen that humans perceive animate objects to be similar to each other, and dissimilar to inanimate objects, as evidenced by the blue cluster in Figure 2B [↗](#).

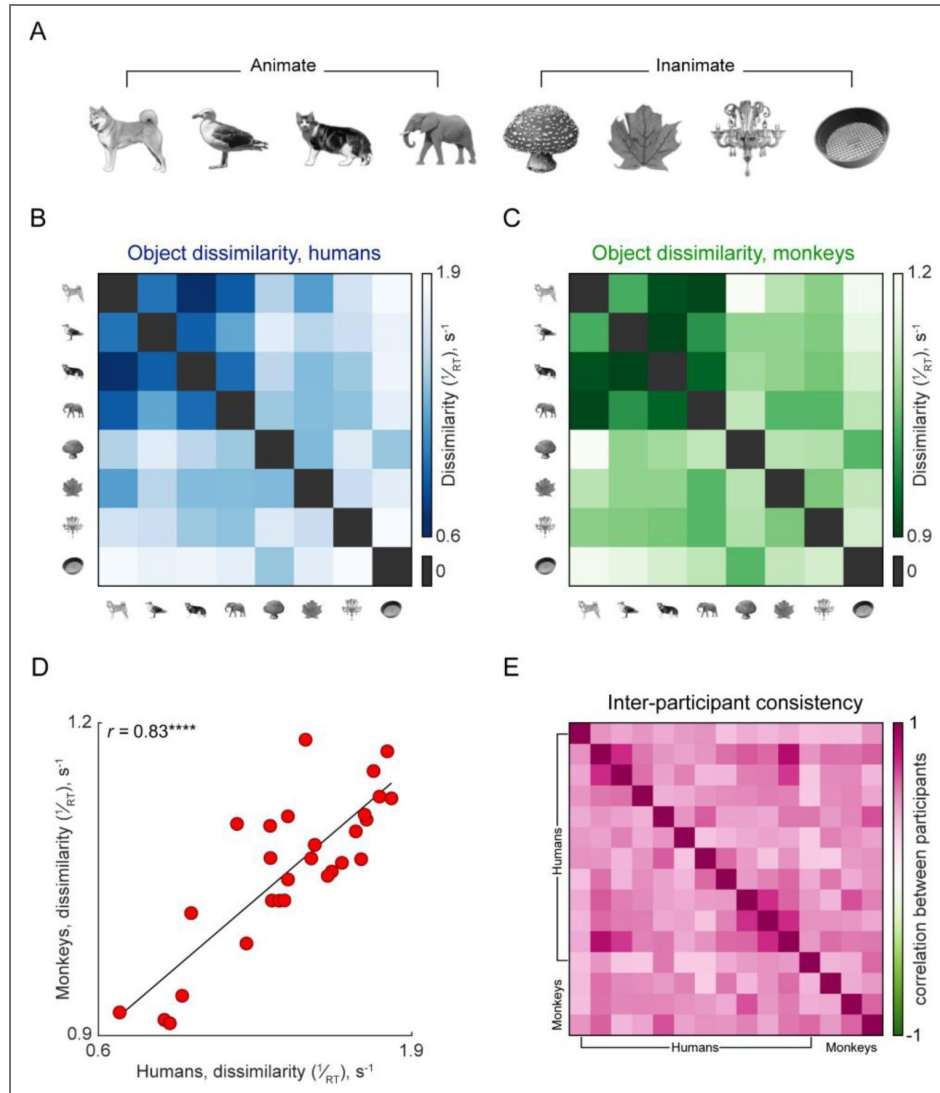
We then proceeded to ask if monkeys show similar object representations as humans. All three monkeys were highly accurate in this task (mean  $\pm$  s.d.: 82%  $\pm$  5%; note that chance accuracy is 25% since there are four items in each search array) and their reaction times were consistent ( $r_{sh} = 0.82$ ,  $p_{sh} < 0.00005$ ,  $n = 28$  search pairs). We then calculated pairwise dissimilarities as before, and the resulting dissimilarity matrix is depicted in Figure 2C [↗](#). Here too, we find that monkeys perceived animate objects to be similar to each other and dissimilar to inanimate objects (Figure 2C [↗](#), green cluster).

To quantify the overall match between perceived object dissimilarities between monkeys and humans, we plotted the dissimilarity of each pair of objects for monkeys against the corresponding value obtained from humans (Figure 2D [↗](#)). This revealed a strong and statistically significant



**Figure 1. Testing visual perception in monkeys and humans**

(A) Schematic of the setup used to test visual perception in monkeys and humans. Monkeys were given a juice reward on making a correct response (indicated by the green juice pipe). An example video of one of the monkeys performing this task is available in [Supplementary Video S1](#). Humans received no such tangible rewards but were instead compensated monetarily for their participation. (B) Schematic of the oddball visual search task used to test visual perception in monkeys and humans. Participants had to touch and hold (red dashed lines are for our reference) a yellow button till a search array appeared. The search array always contained one oddball item among multiple identical distractors. Participants could touch any item to complete the trial; touching the oddball target was taken as a correct response.



**Figure 2. Monkeys have similar coarse object representations as humans**

(A) Animate and inanimate objects used in Experiment 1. (B) Object dissimilarities measured from visual search for humans, who perceive animate objects to be similar to each other (blue cluster), and dissimilar to other objects. (C) Same as in (B), but for monkeys who also perceive animate objects to be similar to each other (green cluster) and dissimilar to all other objects. (D) Object dissimilarities for monkeys plotted against object dissimilarities in humans across all 28 pairs of objects tested.  $r$  is the Pearson's correlation (\*\*\*\* indicates  $p < 0.00005$ ). (E) Inter-participant consistency for all pairs of participants (both human and monkey), calculated as the pairwise correlation between participants' object dissimilarities. Large values indicate high agreement between that pair of subjects.

correlation (Figure 2D [↗](#);  $r = 0.83$ ,  $p < 0.00005$ ,  $n = 28$  search pairs).

The global match in perceived dissimilarities between humans and monkeys could potentially obscure individual differences. To evaluate this possibility, we calculated the pairwise object dissimilarity for each participant (monkey or human) and calculated the correlation between these dissimilarities for every pair of participants. A high correlation implies that searches that were easy for one participant were also easy for the other, which implies high agreement. Conversely, a low correlation could arise if searches that are easy for one participant are hard for the other. If monkeys and humans had qualitatively different object perception, we would expect them to form two groups, where monkeys' responses agree with each other but not with humans and likewise human responses agree with each other more than with monkeys. However, we observed no such pattern (Figure 2E [↗](#)), suggesting that there is a good overall match in object perception between monkeys and humans.

We conclude that monkeys have similar object representations as humans at least at a coarse level.

## Experiment 2: Do monkeys show Weber's law like humans?

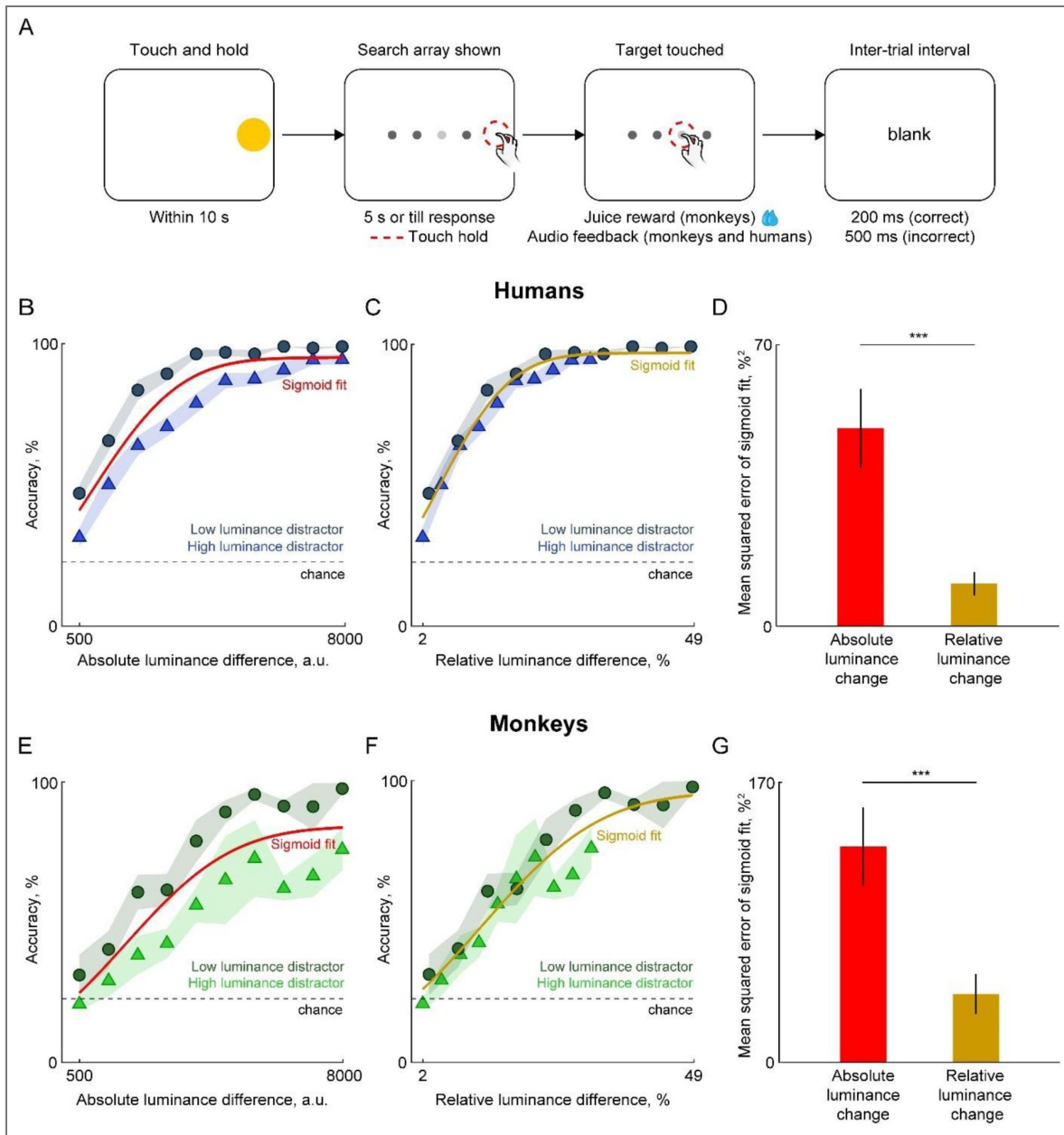
Next, we tested humans and monkeys on a universal law of psychophysics, namely Weber's law. According to Weber's law, the just noticeable difference in a stimulus depends on the baseline magnitude of the stimulus. Weber's law is interesting because it is not a low-level visual property but rather requires higher order processing of relative rather than absolute magnitude. We selected luminance as a simple visual feature and asked whether visual search for oddball targets differing in luminance would follow Weber's law.

To this end, we created search displays in which the oddball target differed only in luminance from the distractors (Figure 3A [↗](#)). According to Weber's law, participants should find it easy to detect the same luminance change when the distractors have low luminance than when they have high luminance. Conversely, the difficulty of search should be similar when the percentage luminance change is the same, regardless of the absolute luminance. We tested these key predictions in human and monkey participants.

Humans were highly accurate on this task (mean  $\pm$  s.d.: 82%  $\pm$  5%) and their reaction times were also highly consistent ( $r_{sh} = 0.97$ ,  $p_{sh} < 0.005$ ,  $n = 20$  search pairs). The performance of 16 human participants is shown in Figure 3B-D [↗](#). When we plotted accuracy as a function of absolute luminance changes, humans showed greater accuracy for detecting larger luminance changes (Figure 3B [↗](#)). But this accuracy was larger for low-luminance distractors (absolute luminance = 16,384 a.u.) compared to high-luminance distractors (absolute luminance = 25,600 a.u.). Importantly, these accuracy differences disappeared when the same data was replotted as a function of percentage difference in luminance (Figure 3C [↗](#)), which is consistent with Weber's law.

To quantify these differences, we fit a sigmoid function to the mean accuracy values for both the low-luminance and high-luminance distractor conditions, and compared the mean squared error between the observed accuracies and the sigmoid fits when the data was plotted as a function of absolute luminance difference and when it was plotted as a function of relative changes in luminance. We found a smaller mean-squared error for relative intensity differences (Figure 3D [↗](#); squared error in accuracy, mean  $\pm$  s.e.m.: 49.18  $\pm$  9.85 %<sup>2</sup> for absolute luminance differences; 10.56  $\pm$  2.85 %<sup>2</sup> for relative luminance differences,  $p < 0.0005$ , rank-sum test,  $n = 20$  search pairs). This confirms that humans showed Weber's law for detecting luminance differences.

Next, we asked whether monkeys also show similar sensitivity to relative luminance compared to absolute luminance. Monkeys also performed well above chance overall (accuracy, mean  $\pm$  s.d.: 64%  $\pm$  4%) and their reaction times were also consistent ( $r_{sh} = 0.69$ ,  $p_{sh} < 0.05$ ,  $n = 20$  search pairs). However, they showed a shallower increase in accuracy with absolute luminance changes compared to humans (Figure 3E [↗](#)). Interestingly, they too showed greater accuracy for detecting luminance changes in low-luminance distractors compared to high-luminance distractors (Figure 3F [↗](#)). This difference in accuracy reduced greatly when the same data was replotted using



**Figure 3. Monkeys experience Weber's law like humans**

(A) Schematic of an oddball visual search trial for testing luminance discrimination. The stimuli were placed in a line centred on the screen and participants had to touch the brightest stimuli, which was always the target, to complete the trial successfully. (B) Accuracy of humans on oddball search for targets differing in luminance as a function of absolute luminance change for high luminance (blue) or low luminance distractors (gray). Shaded regions depict s.e.m. of accuracy across 16 participants. The best-fitting sigmoid line across both conditions is overlaid in red. The dotted line indicates chance accuracy level. (C) Same data as in panel B but replotted as a function of percent change in luminance of the target relative to the distractor. The best fitting sigmoid line across both conditions is overlaid in ochre. The dotted line indicates chance accuracy level. (D) Mean squared error between the sigmoid fit and the observed data for absolute luminance changes (red) and relative luminance changes (ochre). Asterisks indicate statistical significance obtained from a rank-sum test across 20 accuracy values in both conditions (\*\*\*) indicates  $p < 0.0005$ ) (E-G) Same as B-D but for data from 3 monkeys doing the same task.

relative changes in luminance (Figure 3G). This was evident upon comparing the mean-squared error in the sigmoid fits for absolute and relative luminance differences (Figure 3F; squared error, mean  $\pm$  s.e.m.:  $131.09 \pm 23.78 \%^2$  for absolute luminance differences;  $41.46 \pm 12.09 \%^2$  for relative luminance differences,  $p < 0.0005$ , rank-sum test,  $n = 20$  search pairs).

To confirm that these differences are present in individual participants, we repeated these analyses for each participant separately. Nearly all (15/16) human participants and all 3 monkey participants showed a larger error in the sigmoid fit for absolute luminance changes compared to relative luminance changes.

We conclude that monkeys, like humans, experience Weber's law for detecting luminance changes in visual search.

### Experiment 3. Do monkeys perceive occluded shapes like humans?

When we see an occluded display such as the one on the left of Figure 4A, we have no trouble understanding the shape that continues behind the occluder. This kind of amodal completion is thought to be rapid and automatic based on visual search studies (Rensink and Enns, 1998). In these studies, humans find it hard to locate an occluded display among distractors with “likely completion” compared to finding a “mosaic” display in which the two shapes abut without occluding (Figure 4A). Thus, the dissimilarity between occluded and likely displays is smaller than the dissimilarity between likely and mosaic displays. We set out to investigate if monkeys also experience amodal completion as humans do.

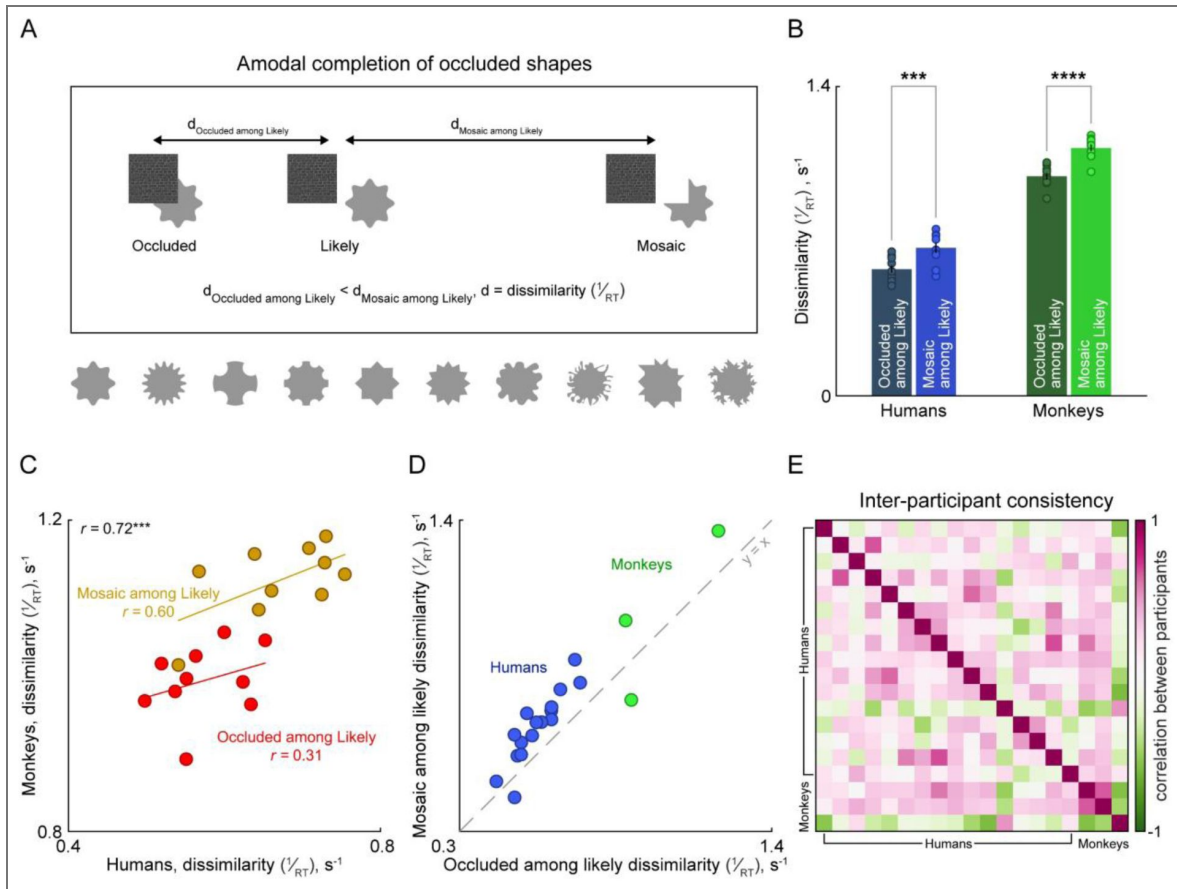
To test amodal completion in monkeys, we took 10 shapes (Figure 4A) and their occluded, likely and mosaic versions from a previous study (Cherian and Arun, 2024). For each shape we used the likely display as the distractor with the occluded or mosaic display as the targets. We then compared the dissimilarities measured in monkeys to humans on the same search pairs.

Humans were highly accurate on this task (accuracy, mean  $\pm$  s.d.:  $96 \pm 4\%$  across 16 participants; chance is 8.3%) and their reaction times were consistent ( $r_{sh} = 0.64$ ,  $p_{sh} = 0.07$ ,  $n = 20$  search pairs). We calculated perceived dissimilarity in visual search as before, by taking the reciprocal of the average response time in each condition. In humans, the dissimilarity between occluded and likely displays were significantly less than the dissimilarity between mosaic and likely displays (Figure 4B; dissimilarity, mean  $\pm$  s.e.m.:  $0.57 \pm 0.16 \text{ s}^{-1}$  for occluded targets among likely displays,  $0.66 \pm 0.23 \text{ s}^{-1}$  for mosaic targets among likely displays;  $p < 0.0005$ , paired t-test across 10 searches). This confirms that humans experience amodal completion.

Monkeys also performed this task accurately (accuracy, mean  $\pm$  s.d.:  $75\% \pm 12\%$  across three monkeys; chance is 25%) and their reaction times were consistent ( $r_{sh} = 0.72$ ,  $p_{sh} = 0.001$ ,  $n = 20$  search pairs). Even for monkeys, the dissimilarity between occluded and likely displays was significantly smaller than the dissimilarity between likely and mosaic displays (Figure 4B; dissimilarities, mean  $\pm$  s.e.m.:  $0.99 \pm 0.14 \text{ s}^{-1}$  for occluded among likely displays;  $1.12 \pm 0.15 \text{ s}^{-1}$  for mosaic among likely displays;  $p < 0.00005$ , paired t-test across 10 searches of each type).

To quantify the overall match between perceived dissimilarities between monkeys and humans, we plotted the dissimilarity for each search condition for monkeys against the corresponding value for humans. This revealed a strong and statistically significant correlation (Figure 4C;  $r = 0.72$ ,  $p < 0.0005$ ). Mosaic among Likely dissimilarities were correlated between monkeys and humans although this did not reach statistical significance (Figure 4C;  $r = 0.60$ ,  $p = 0.06$ ). Finally, Occluded among Likely dissimilarities were not correlated (Figure 4C;  $r = 0.31$ ,  $p = 0.37$ ).

We then asked if these trends in dissimilarity hold at the individual level. To this end, we plotted the average dissimilarity for mosaic among likely searches for each participant against the average dissimilarity for occluded among likely searches (Figure 4D). Values above the  $y = x$  line would confirm the amodal completion effect. Nearly all individual participants, except one human and one monkey, showed the effect (Figure 4D). To further compare individual monkey and human participants, we computed a similarity score between pairs of participants as the correlation between the dissimilarities across the searches. The resulting pairwise dissimilarities



**Figure 4. Monkeys experience rapid amodal completion like humans**

(A) Occluded objects are rapidly and automatically completed during visual search. In the key comparison, search for an occluded display (left) among “likely completions” (middle) as distractors is hard, implying that the two displays have low dissimilarity. By contrast, searching for a “mosaic” display (right) among likely completions (middle) is easy, implying that these two displays are dissimilar. The shapes at the bottom show the full set of stimuli tested in the experiment. (B) Average dissimilarities for occluded-likely and mosaic-likely searches for humans (blue, left) and monkeys (green, right). Asterisks indicate statistical significance using a t-test across the dissimilarities of 10 shapes tested (\*\*\* is  $p < 0.0005$ ; \*\*\*\* is  $p < 0.00005$ ). (C) Effect strength for each condition. Dissimilarities in visual search for monkeys across all searches tested plotted against the corresponding values for humans. The correlation between all points is indicated on the top left ( $r = 0.72$ ). Correlations for each group of pairs are also shown separately: occluded among likely (red dots) and mosaic among likely (ochre dots). Asterisks represent statistical significance (\*\*\* is  $p < 0.0005$ ). Correlation values without an asterisk are not statistically significant. (D) Effect strength for each participant. Average dissimilarities of each participant for occluded-likely searches plotted against the dissimilarities for the corresponding mosaic-likely searches. Blue dots represent the 16 human participants; green dots represent 3 monkeys. The dashed line indicates the  $y = x$  line. It can be seen that nearly all human and monkey participants showed a larger dissimilarity for mosaic-likely searches compared to occluded-likely searches, confirming the amodal completion effect. (E) Inter-participant consistency for all pairs of participants (both human and monkey), calculated as the pairwise correlation between participants’ object dissimilarities. Large values indicate high agreement between that pair of subjects.

are shown in [Figure 4E](#). This revealed no clear grouping of monkey or human participants, again underscoring the fact that monkeys show very similar responses to amodal completion as humans.

We conclude that monkeys experience amodal completion in visual search, just like humans.

#### Experiment 4. Do monkeys show mirror confusion like humans do?

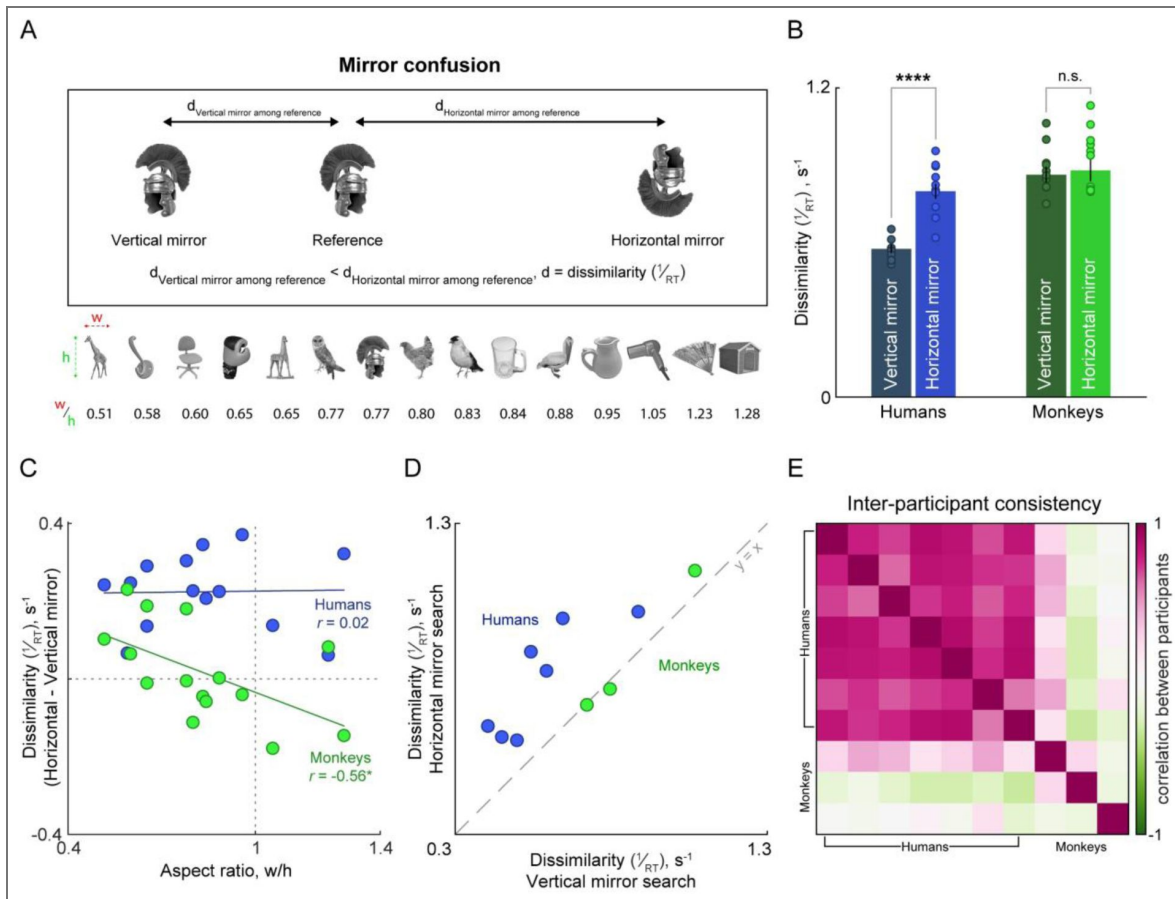
Humans tend to confuse vertical mirror images compared to horizontal mirror images, presumably because we frequently see vertical images across changes in viewpoint ([Rollenhagen and Olson, 2000](#); [Corballis, 2018](#)). Mirror confusion has been observed in high-level visual cortex in monkeys ([Rollenhagen and Olson, 2000](#)) as well as in humans ([Corballis, 2018](#)). However, there has been only two studies showing that macaque monkeys show low accuracy in learning to discriminate vertical mirror images of artificial shapes ([Riopelle et al., 1964](#); [Washburn, 1993](#)).

We sought to systematically investigate this issue by asking if humans and monkeys experience mirror confusion in visual search involving natural images. We selected 15 natural objects and created horizontal and vertical mirror images of each ([Figure 5A](#)). If participants experience vertical mirror confusion, they should be slower to respond on trials when the target is a vertical mirror image compared to when the target is a horizontal image. In terms of dissimilarity, the dissimilarity between vertical mirror images should be smaller than the dissimilarity between horizontal mirror images ([Figure 5A](#)).

Humans were highly accurate during visual search involving mirror images (accuracy, mean  $\pm$  s.d.: 95%  $\pm$  2%, chance is 25%) and their reaction times were also highly consistent ( $r_{sh} = 0.92$ ,  $p_{sh} < 0.001$ ,  $n = 30$  search pairs). They were also more accurate on horizontal compared to vertical mirror image searches (accuracy, mean  $\pm$  s.d.: 93%  $\pm$  3% for vertical mirror searches; 99%  $\pm$  1% for horizontal mirror searches;  $p < 0.0005$ , sign-rank test across mean accuracy across 15 objects). This effect was more pronounced when comparing the search dissimilarity, obtained by taking the reciprocal of the average response times ([Figure 5B](#); dissimilarity, mean  $\pm$  s.e.m.: 0.57  $\pm$  0.01 s<sup>-1</sup> for vertical mirror searches; 0.80  $\pm$  0.02 s<sup>-1</sup> for horizontal mirror searches;  $p < 0.00005$ , paired t-test across 15 natural objects). Thus, in keeping with previous observations, humans showed strong and robust vertical mirror confusion during visual search.

Next, we asked whether monkeys showed mirror confusion. Monkeys performed the task itself well above chance (accuracy, mean  $\pm$  s.d.: 67%  $\pm$  6% across three monkeys; chance is 25%, see example trials in [Supplementary Video S1](#)), and their reaction times were also consistent ( $r_{sh} = 0.81$ ,  $p_{sh} < 0.001$ ,  $n = 30$  search pairs). Their accuracy did not differ significantly between vertical and horizontal mirror searches (accuracy, mean  $\pm$  s.d.: 66%  $\pm$  11% for vertical mirror searches; 69%  $\pm$  3% for horizontal mirror searches;  $p = 0.38$ , sign-rank test across mean accuracy across 15 objects). We also observed no significant difference in their dissimilarities for vertical vs horizontal mirror searches ([Figure 5B](#); dissimilarity, mean  $\pm$  s.e.m.: 0.86  $\pm$  0.03 s<sup>-1</sup> for vertical mirror searches; 0.88  $\pm$  0.04 s<sup>-1</sup> for horizontal mirror searches;  $p = 0.59$ , paired t-test across 15 natural objects). Thus, monkeys do not experience mirror confusion during visual search like humans do.

Because prior studies with monkeys have demonstrated mirror-confusion ([Riopelle et al., 1964](#); [Washburn, 1993](#); [Rollenhagen and Olson, 2000](#)), we dug deeper into our data to investigate if the monkey behaviour was driven by any stimulus-dependent factors. We found that monkeys do show mirror confusion for some objects but not others. The effect seemed to be strongly driven by the aspect ratio (width/height): a plot of the difference in dissimilarity (horizontal minus vertical) against the aspect ratio for each object revealed a very interesting difference in behaviour between humans and monkeys ([Figure 5C](#)). Humans showed a consistent mirror confusion independent of aspect ratio ([Figure 5C](#),  $r = 0.02$ ,  $p = 0.935$ ). By stark contrast, monkeys showed a strong negative correlation ([Figure 5C](#),  $r = -0.56$ ,  $p = 0.029$ ). In other words, for narrow objects



**Figure 5. Monkeys show no mirror confusion unlike humans**

(A) Perception of mirror confusion of natural images in visual search for humans. Below are the 15 stimuli arranged according to the aspect ratio: calculated as the ratio of the width to the height. (B) Mean  $\pm$  s.e.m. of the dissimilarities for humans and monkeys across the two types of trials, vertical mirror among reference: the target was a vertical mirror image embedded among reference displays, horizontal mirror among reference: the target was a horizontal mirror image embedded among reference displays, asterisks indicate significance of the Student's t-test checking for differences between the means where \*\*\*\* is  $p < 0.0005$ , and n.s. is  $p > 0.05$  (paired samples), data points are unique trials (averaged across all subjects and repetitions). (C) *Effect strength for each search.* Difference in dissimilarity between horizontal and vertical mirror conditions plotted against the aspect ratio of the 15 stimuli tested for humans (blue) and monkeys (green). The  $r$  values are the Pearson correlation and coloured lines are the best-fitting lines for each group. (D) *Effect strength for each participant.* Mean dissimilarities for vertical mirror among reference and horizontal mirror among reference pairs for each subject, 7 humans and 3 monkeys. (E) Inter-participant consistency for all pairs of participants (both human and monkey), calculated as the pairwise correlation between participants' object dissimilarities. Large values indicate high agreement between that pair of subjects.

they had higher dissimilarity for horizontal mirror images compared to vertical mirror images, whereas the trend was reversed for wide objects. This suggests that they were using very different features to search for oddball targets compared to humans.

To elucidate whether mirror confusion is present in individual participants, we plotted the dissimilarity for each participant for horizontal mirror searches versus vertical mirror searches (Figure 5D). Here, vertical mirror confusion would be indicated by values above the  $y = x$  line. It can be seen that all human participants showed a mirror confusion effect whereas the effect is weak or absent across the three monkeys. Finally, to investigate how similar participants are to each other in their responses, we calculated the inter-participant consistency as before. This revealed a remarkable grouping of human and monkey participants: human participants were more similar to each other in their responses, whereas monkeys were dissimilar to humans and to each other in their responses (Figure 5E).

We conclude that monkeys do not show mirror confusion during visual search as humans do.

## Experiment 5. Do monkeys show an advantage for global shapes?

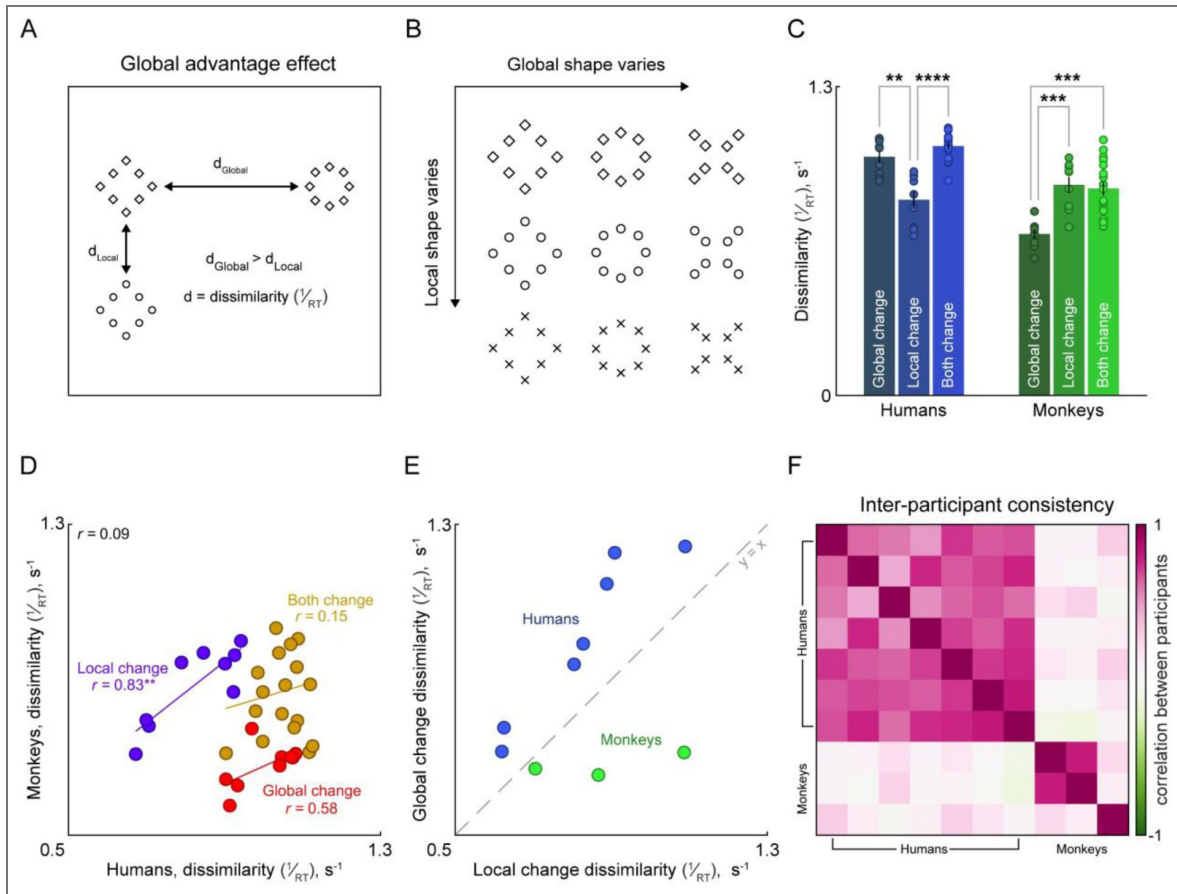
Humans are faster to report shape at the global level compared to the local level (Navon, 1977; Kimchi, 1992). In visual search, this has been demonstrated as faster responses to changes in global shape compared to local shape (Jacob and Arun, 2020), as depicted in Figure 6A. We therefore set out to investigate whether monkeys show global advantage during visual search like humans do.

To this end, we created hierarchical shapes with either a diamond, circle, or 'x' shape at the global, local or both levels (Figure 6B). We created search arrays involving all pairs of these shapes: these search arrays could therefore be divided into those that contained only a change in global shape, a change in local shape, or both global and local shape changes.

Humans were highly accurate on this task (accuracy, mean  $\pm$  s.d.: 96%  $\pm$  1%) and their reaction times were also highly consistent ( $r_{sh} = 0.93$ ,  $p_{sh} < 0.001$ ,  $n = 36$  search pairs). They were also more accurate to find global compared to local shape changes (accuracy, mean  $\pm$  s.d.: 97%  $\pm$  0.2% for global change searches; 91%  $\pm$  0.4% for local change searches;  $p < 0.05$  using sign-rank test comparing 9 global vs local change searches). Humans showed significantly smaller dissimilarities for searches with local compared to global shape differences, confirming that they are more sensitive to global shape changes (Figure 6C; dissimilarity, mean  $\pm$  s.d.: 1.00  $\pm$  0.02 s<sup>-1</sup> for searches with only global change; 0.82  $\pm$  0.03 s<sup>-1</sup> for searches with only local change,  $p < 0.0005$ , paired t-test for global vs local change across 9 searches). Interestingly, searches with both global and local change were even faster than searches involving only one type of change, suggesting that humans integrate both local and global shape during search (Figure 6C; dissimilarity, mean  $\pm$  s.e.m.: 1.05  $\pm$  0.06 s<sup>-1</sup> for searches with both change;  $p < 0.00005$  comparing both change with local-only changes and  $p = 0.09$  comparing both change with global-only changes, unpaired t-test). Thus, consistent with previous studies, humans showed a robust global advantage during visual search.

Next, we asked whether monkeys showed a global advantage. Monkeys performed poorly on visual search involving these hierarchical stimuli (accuracy, mean  $\pm$  s.d.: 56%  $\pm$  15% across three monkeys; chance is 25%) but their reaction times were highly consistent ( $r_{sh} = 0.85$ ,  $p_{sh} < 0.001$ ,  $n = 30$  search pairs). They also showed an accuracy difference between global and local shape changes in favour of local shape changes (accuracy, mean  $\pm$  s.d.: 31%  $\pm$  0.1% for global change searches; 66%  $\pm$  1.9% for local change searches;  $p = 0.003$  using sign-rank test comparing 9 global vs local change searches).

However, monkeys showed a dramatically different pattern from humans. This is illustrated by one of the monkeys doing the task in Supplementary Video S1. Monkeys showed a significantly smaller dissimilarity for global shape changes compared to local shape changes (Figure 6C; dissimilarity, mean  $\pm$  s.e.m.: 0.67  $\pm$  0.01 s<sup>-1</sup> for searches with only global change; 0.88  $\pm$  0.03 s<sup>-1</sup> for searches with only local change;  $p < 0.0005$ , paired t-tests comparing 9 global and local searches). Even for searches with both global and local changes, monkeys' performance was better than



**Figure 6. Monkeys show a local advantage unlike humans**

(A) Humans show a global advantage when viewing hierarchical stimuli, such as a circle made of diamonds. Specifically, for humans, shapes differing only at the global level are perceived as more dissimilar than shapes differing only at the local level (i.e.  $d_{Global} > d_{Local}$ ). (B) Hierarchical stimuli used for the visual search experiment in humans and monkeys. An example video of one of the monkeys performing this task is available in [Supplementary Video S1](#). (C) Dissimilarity between pairs of hierarchical stimuli during visual search, in humans (left group, blue) and monkeys (right group, green). Dissimilarity is shown separately for three groups of search pairs: pairs differing only in global shape (left bars), differing only in local shape (middle bars) and differing in both global and local shape (right bars). Dots centred around each bar represent average dissimilarity for each individual participant. Asterisks indicate statistical significance obtained from an unpaired t-test comparing the mean dissimilarities in each group ( $n = 9$  for global change and local change,  $n = 18$  for both change,  $**$  is  $p < 0.005$ ,  $***$  is  $p < 0.0005$ ,  $*****$  is  $p < 0.00005$ , non-significant comparisons are not marked). (D) Effect strength for each condition between humans and monkeys. Average dissimilarity for each search pair in monkeys plotted against the corresponding values in humans. The overall correlation between all search pairs is shown on the top left ( $r = 0.09$ ). Correlations for each group of pairs are also shown separately: pairs differing only in global shape (red dots), in local shape (purple dots) and differing in both global and local shape (ochre dots). Asterisks represent statistical significance ( $**$  is  $p < 0.005$ ). Correlation values without an asterisk are not statistically significant. (E) Effect strength for each participant. Average dissimilarity for each search pair with a global shape difference plotted against the dissimilarity for the corresponding local shape difference for individual participants (humans, blue dots; monkeys, green dots). It can be seen that all humans find global shape differences consistently larger than local shape differences, whereas all monkeys show exactly the opposite pattern. (F) Inter-participant consistency for all pairs of participants (both human and monkey), calculated as the pairwise correlation between participants' object dissimilarities. Large values indicate high agreement between that pair of subjects.

searches with global-only changes but no better than their performance on searches with local-only changes (Figure 6C [↗](#); dissimilarity, mean  $\pm$  s.e.m.:  $0.87 \pm 0.02 \text{ s}^{-1}$  for searches with both global and local change;  $p < 0.0005$  for both vs global-only changes;  $p = 0.73$  for both vs local-only changes, unpaired t-test). Thus, monkeys showed markedly smaller sensitivity to changes in global shape and were more sensitive to local shape changes, exactly opposite to humans.

To investigate these differences between monkeys and humans further, we plotted the search dissimilarity in monkeys for each search condition against the corresponding values obtained from human participants (Figure 6D [↗](#)). This revealed very interesting patterns. There was no overall correlation between human and monkey dissimilarities across all 36 search pairs tested (Figure 6D [↗](#),  $r = 0.09$ ,  $p = 0.59$ ). Interestingly, human and monkey dissimilarities were correlated only for searches involving local-only shape differences (Figure 6D [↗](#);  $r = 0.83$ ,  $p < 0.005$  for local-shape changes) but not for global-only searches or both global & local change searches (Figure 6D [↗](#);  $r = 0.58$ ,  $p = 0.09$  for global-only searches;  $r = 0.15$ ,  $p = 0.54$  for global & local searches). The poor correlation for global-change-only searches might be due to poor accuracy in those searches for monkeys. Furthermore, dissimilarities for both-change searches also do not agree as both humans and monkeys appear to use different information to make their decisions. Thus, monkeys seem to prioritize information at the global and local levels differently from humans when both types of information is present, but process information at the local level similarly to humans.

To elucidate individual differences, we plotted the average dissimilarity for each participant for global changes against the corresponding dissimilarity for local changes (Figure 6E [↗](#)). This too revealed a remarkable difference: all human participants showed greater dissimilarity for global changes, whereas all monkeys showed greater dissimilarity for local changes.

This striking difference between humans and monkeys was also evident upon calculating inter-participant consistency as before. Here, we found that human participants were highly consistent with each other and behaved quite differently from monkeys. Likewise, all three monkeys were more consistent with each other than with human participants.

We conclude that monkeys show a local advantage, unlike humans.

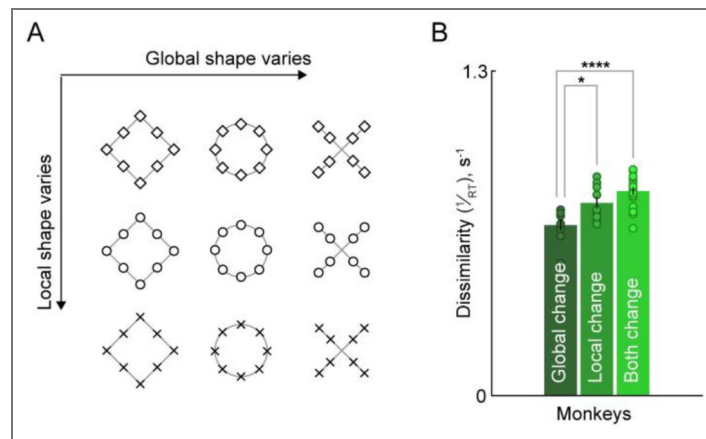
## Experiment 6: Why do monkeys show a local advantage?

The local advantage observed in monkeys could be due to two reasons. One possibility is that monkeys cannot integrate spatially disconnected local elements into a larger global shape. Alternatively, they might prioritize information at smaller spatial scales compared to humans. To distinguish between these possibilities, we performed an additional experiment. We took the same shapes as used earlier but added connecting gray lines between the local shapes so that the global shape was evident without the requirement of grouping (Figure 7A [↗](#)).

All three monkeys were well above chance on this task (accuracy, mean  $\pm$  s.d.:  $59\% \pm 12\%$  across three monkeys, chance is 25%) and their reaction times were consistent ( $r_{sh} = 0.80$ ,  $p_{sh} < 0.001$ ,  $n = 30$  search pairs). As in Experiment 5, they also showed an accuracy difference between global and local shape changes in favour of local shape changes (accuracy, mean  $\pm$  s.d.:  $43\% \pm 07\%$  for global change searches;  $62\% \pm 13\%$  for local change searches;  $p = 0.02$  using sign-rank test comparing 9 global vs local change searches).

Here too, monkeys showed a larger dissimilarity for local changes compared to global changes (Figure 7B [↗](#), mean  $\pm$  s.e.m.:  $0.68 \pm 0.02 \text{ s}^{-1}$  for searches with only global change;  $0.77 \pm 0.02 \text{ s}^{-1}$  for searches with only local change;  $p < 0.05$ , paired t-test across 9 search pairs). Even for pairs differing in both local and global shape, monkeys behaved as if they only processed local shape: the dissimilarity for shapes differing at both levels was larger than global-only changes, but similar to local-only changes (average dissimilarity, mean  $\pm$  s.e.m.:  $0.81 \pm 0.01 \text{ s}^{-1}$  across 18 searches with both global and local change;  $p = 0.091$  for local change vs both change;  $p < 0.00005$  for global change vs both change; unpaired t-test).

Thus, local advantage in monkeys is present even when local elements are explicitly connected together, suggesting that their local advantage comes from a higher priority for local shape rather than an inability to group local elements. It would be interesting to explore these differences



**Figure 7. Monkeys show a local advantage even for connected shapes**

(A) Hierarchical stimuli used for this experiment. Here, we connected local elements within each global shape with a gray line to facilitate grouping into a global shape. (B) Average dissimilarity between pairs of hierarchical stimuli during visual search in monkeys, shown separately for three groups of search pairs: pairs differing only in global shape (*left bars*), differing only in local shape (*middle bars*) and differing in both global and local shape (*right bars*). Dots centred around each bar represent average dissimilarity for each individual participant. Asterisks indicate statistical significance obtained from an unpaired t-test comparing the mean dissimilarities in each group ( $n = 9$  for global change and local change,  $n = 18$  for both change, \* is  $p < 0.05$ , \*\*\*\* is  $p < 0.00005$ , non-significant comparisons are not marked).

further in future work.

## Discussion

Here, we compared perception in monkeys and humans by testing them on a range of high-level perceptual phenomena. Monkeys showed similar coarse object representations, Weber's law and rapid amodal completion just like humans. However, they did not show mirror confusion or a global advantage, unlike humans. Our results reveal both qualitative similarities and differences between human and monkey perception and ultimately provide for an unbiased and nuanced evaluation of macaque monkeys as a model for human vision. Below we discuss these findings in relation to the existing literature.

### Relation to previous approaches

There is a long and rich literature of studies on visual perception in non-human primates and comparisons to human vision (Matsuno and Fujita, 2009 [↗](#)). Comparative studies have included diverse non-human primate species including marmosets, baboons, macaques and even chimpanzees. Because of the difficulty of obtaining any kind of report from animals, each study usually has only tested one specific aspect of visual perception in non-human primates. These tests often involve extensive training of animals on complex tasks. Such extensive training could potentially modify the underlying visual representations or the decision-making process.

Our study represents a novel alternative to study animal perception. We trained animals on an oddball search task that could be performed on any set of images. We then selected a range of perceptual phenomena that could be tested by choosing different sets of images. This approach was successfully used recently to investigate deep networks (Jacob et al., 2021b [↗](#)) and proposed as a general framework for cognition (Taylor et al., 2022 [↗](#)). For animal cognition, our approach has the major advantage that animals need not be retrained each time to test each additional aspect of perception. As a result, our approach also avoids the danger of extensive training modifying the underlying representations.

An interesting alternative approach to understanding perception in animals is a free-choice paradigm, pioneered by Logothetis and colleagues to study the neural correlates of subjective percepts in monkeys (Logothetis and Schall, 1989 [↗](#); Leopold and Logothetis, 1996 [↗](#)). In these studies, animals were first trained to report unambiguous stimuli first where only the correct response is rewarded, and then presented in a small fraction of trials containing ambiguous binocular rivalry stimuli where both responses are rewarded. The animals' response in the ambiguous trials were highly systematic and were taken as indicative of their subjective percept. Such free choice paradigms could be extended to other aspects of perception such as those tested in our study. For instance, to test mirror confusion, monkeys could be trained on a same-different task involving pairs of completely different images where they are rewarded only for making correct responses and then presented in a small fraction of trials with vertical-mirror and horizontal-mirror pairs of images where either response would be rewarded. If monkeys made more "same" responses for vertical-mirror compared to horizontal-mirror pairs, this would confirm that they experience mirror confusion. We seriously considered such free-choice paradigms but eventually decided to use visual search for several reasons. First, free-choice paradigms can only have a small fraction of trials of interest, limiting our ability to test many aspects of perception. Second, free choice paradigms work best if all trials are similar in difficulty, but even small variations in difficulty could make animals change their response strategies when they encounter hard trials. Third, free-choice paradigms also involve extensive training to achieve extremely high levels of performance. In our experience, monkeys are highly sensitive to task difficulty and can often switch their task strategy fluidly from following a rule to making random choice responses. By contrast, visual search in our experience was a much more natural task for monkeys to learn and can be performed at a high throughput rate. Importantly, we were easily able to elicit large variations in response times without affecting the overall accuracy and reward

rate, thereby reducing the animals' incentive to switch task strategies and operate according to a stable decision rule. We therefore propose that visual search could be a highly effective task to study animal perception across multiple species.

## Relation of each finding to previous studies

We now discuss each individual finding in our study with previous studies. Our first finding is that monkeys and humans have similar object representations in visual search. In other words, searches that were hard for humans were hard for monkeys as well (Figure 2). This finding is consistent with a number of previous studies showing similar object categorization behaviour (Rajalingham et al., 2015, 2018), and similar object representations at the neural level (Kriegeskorte et al., 2008b) in monkeys and humans. In fact, neural representations in monkey IT cortex are highly predictive of visual search dissimilarity in humans (Sripati and Olson, 2010; Zhivago and Arun, 2014). These studies together suggest that macaque monkeys and humans have similar coarse object representations.

Our second finding is that monkeys show Weber's law for luminance, just like humans (Figure 3). Weber's law has been demonstrated in humans for luminance discrimination for intensity values within a broad range (Hecht, 1924; Pramod and Arun, 2014), line length (Ono, 1967; Pramod and Arun, 2014) and sound pressure levels (Pardo-Vazquez et al., 2019). While Weber's Law has been demonstrated in monkeys for numerosity (Nieder and Miller, 2003; Jordan and Brannon, 2006) and line length (Tudusciuc and Nieder, 2010), to our knowledge our study is the first to demonstrate Weber's law for luminance discrimination in monkeys.

Our third finding is that monkeys experience rapid and automatic amodal completion in visual search, just like humans (Figure 4). There is only one similar study where monkeys trained to categorize line length respond to an occluded line as being longer (Fujita, 2001). By contrast, we demonstrate using a wide range of abstract shapes that monkeys find it difficult to find an occluded target among likely completions compared to a mosaic target among likely completions, just like humans do (Rensink and Enns, 1998; Cherian and Arun, 2024). Our study is the first to demonstrate that monkeys experience rapid, automatic amodal completion in the same way as humans.

Our fourth finding is that monkeys do not show mirror confusion, unlike humans (Figure 5). This disagrees with previous reports of mirror confusion in monkeys both in behaviour (Riopelle et al., 1964; Washburn, 1993) as well as in their inferior temporal cortex (Rollenhagen and Olson, 2000). There could be several reasons for this discrepancy. The task in the two previous studies involved training monkeys to discriminate between two mirror images that were either horizontal mirror reflections or vertical mirror reflection, and the images used were letters or symbols. By contrast, monkeys in our study did not receive any special shape discrimination training and were tested on images of real animate and inanimate objects. Reconciling these differences will involve testing monkeys on both natural objects as well as abstract shapes. Our finding also stands in contrast with the observation of mirror confusion in single neurons in object-selective and face-selective cortex in macaque monkeys (Rollenhagen and Olson, 2000; Freiwald and Tsao, 2010). We speculate that monkeys might have such representations but might not necessarily use them to guide visual search. These are interesting possibilities for future study.

Our final finding is that monkeys show a local advantage while processing stimuli containing global and local shape, unlike humans who show a global advantage (Figure 6). This local advantage persisted even when we connected the local elements, suggesting that the local advantage may not arise from differences in perceptual grouping. There is mixed evidence in the literature regarding this effect in non-human primates, with many studies reporting a local advantage and some a global advantage. Baboons and capuchin monkeys trained on shape matching tasks show local advantage (Deruelle and Fagot, 1997; Fagot and Deruelle, 1997; Spinozzi et al., 2006). Interestingly, chimpanzees trained on oddball search have a local advantage during visual search but a global advantage when the local elements are connected (Fagot and Tomonaga, 1999). By contrast, macaque monkeys continued to show a local

advantage in our study even when the local elements are connected. When tested on the same stimuli and task, chimpanzees showed a global advantage, but macaque monkeys showed a local advantage (Hopkins and Washburn, 2002 [↗](#)). Thus, there could be differences even between non-human primates. Our findings are discordant with previous studies showing that macaque monkeys trained on shape matching have a global advantage (Tanaka and Fujita, 2000 [↗](#); Hopkins and Washburn, 2002 [↗](#)) and with a report of global advantage present in the macaque high-level visual cortex (Sripati and Olson, 2009 [↗](#)). Once again, we surmise that perhaps macaques have such representations in some brain region but may not necessarily use them for visual search. Evaluating this possibility will require testing for global/local advantage at the behavioural and neural levels simultaneously.

## Implications for macaques as a model system for vision

Our study represents a first comprehensive evaluation of visual perception in monkeys on multiple perceptual phenomena and revealed important similarities as well as differences between monkey and human vision. The qualitative similarities with human vision reconfirms that macaque monkeys are indeed good models. However, the qualitative differences observed in our study also demonstrate important limitations of macaque monkeys as a model for human vision. However, we must be cautious in interpreting these similarities and differences, because of the important distinction between performance and competence (Firestone, 2020 [↗](#)). In other words, macaque monkeys and humans may have similar visual representations, but their overt behaviours might be different due to the way in which they utilize these representations to perform tasks. This could explain the contradictory findings between our study and previous work: for instance, there is mirror confusion and global advantage in the high-level visual areas of macaque monkeys, but we did not observe these effects in our study in behaviour. Of course, these results come from testing neural representations in one group of macaques and testing behaviour in another, so reconciling these differences will require careful testing and interpretation.

## Conclusions

Our study represents a novel and general approach to characterize a variety of visual systems, similar to the signature testing approach proposed for cognition (Taylor et al., 2022 [↗](#)). We propose that visual search can be a powerful paradigm to test a variety of animals since is a natural, ecologically valid task with an objective performance measure. Visual search dissimilarities can be directly compared to neural dissimilarities, enabling comparing neural representations (Sripati and Olson, 2010 [↗](#); Zhivago and Arun, 2014 [↗](#)). These can also be compared with dissimilarities in neural network activations (Jacob et al., 2021b [↗](#)) or computer vision algorithms (Pramod and Arun, 2016b [↗](#)). This enables powerful comparisons between a variety of visual systems for core perceptual phenomena, which we believe will yield fundamental insights into the nature of vision and perception.

## Materials and methods

All human experiments were performed according to a protocol approved by the Institutional Human Ethics Committee at the Indian Institute of Science (IHEC # 10/20.07.2022). All animal experiments were performed according to a protocol approved by the Institutional Animal Ethics Committee of the Indian Institute of Science and by the Committee for the Purpose of Control and Supervision of Experiments on Animals, Government of India (V-11011(3)/15/2020-CPCSEA-DADF).

## Humans

All participants were recruited from the Indian Institute of Science community, had normal or corrected-to-normal vision and were monetarily compensated. In addition, they were unaware of the experimental goals prior to completing the tasks. Experiments 2, 4 and 5 were conducted on the same setup as used to test monkeys, whereas Experiments 1 & 3 were obtained from previous studies (see below).

We recruited 16 participants for Experiment 2 (6 female, 10 male; ages mean  $\pm$  s.d.:  $26.1 \pm 3.6$  years), 7 participants for Experiment 4 (4 female, 3 male; ages mean  $\pm$  s.d.:  $26.7 \pm 2.2$  years) and 7 participants for Experiment 5 (4 female, 3 male; ages mean  $\pm$  s.d.:  $23.7 \pm 3.0$  years).

For Experiments 1 & 3, we used previously collected data from human participants performing a similar oddball search task (Mohan and Arun, 2012 [↗](#); Cherian and Arun, 2024 [↗](#)).

## Monkeys

Three male adult bonnet macaques (*macaca radiata*; all aged  $\sim$ 8 years, laboratory designations: *Di*, M1; *Ju*, M2; and *Co*, M3) were tested in this study. The monkeys were under fluid restriction during data collection but had unlimited access to food. These monkeys were trained on same-different tasks and fixation tasks previously (Jacob et al., 2021a [↗](#)) and were trained to perform the oddball search task for this study. All three animals easily learned the task in less than two weeks.

## Apparatus

We used a touchscreen-based behaviour station for our data collection for monkeys (Jacob et al., 2021a [↗](#)). Briefly, monkeys sat at a juice spout in front of a touchscreen (15" capacitive touchscreen, 1366 x 768 pixels, Elo Touch Solutions Inc. Model 1593L RevB) at an eye-to-screen distance of 23 cm. For humans, we replicated the setup on a table such that humans could sit comfortably in front of the screen with the same eye-to-screen distance as monkeys. All experiments were presented using NIMH MonkeyLogic software (Hwang et al., 2019 [↗](#)) running on MATLAB (R2020b, The Mathworks Inc.). Monkeys were rewarded with juice delivered through a custom reward circuit and peristaltic pump that allowed for precise dosing.

## Task

Both humans and monkeys performed the same oddball visual search task (Figure 1A [↗](#)). At the start of the trial, a yellow button (a circle with a radius of  $4^\circ$ , degree of visual angle) appeared on the screen against a black background. Subjects had to hold this button to initiate the trial within 10 seconds. After this, four stimuli (three identical items and one oddball item) appeared on the screen (see Figure 1B [↗](#) and 1C [↗](#)), arranged circularly around the left side of the hold button. Stimuli were presented  $15^\circ$  away from the hold button centre and  $36^\circ$  apart from each other. For Experiment 2, stimuli appeared in a row aligned to the screen centre and spaced  $9^\circ$  apart. Participants had to touch any target to complete the trial; touches to the oddball target were considered a correct response. On making a correct response, a rising auditory tone was played, and monkeys were rewarded with juice. Correct trials were followed by a short inter-trial interval (200 ms) and errors by a slightly longer inter-trial interval (500 ms).

For two experiments (Experiments 1 and 3), human data was obtained from previous studies from our lab where participants also performed a similar oddball visual search task (Figure 1D [↗](#)). Briefly, participants were seated in front a monitor and were presented with either a 16-item search array (4-by-4 grid, Experiment 1) or a 12-item search array (3-by-4 grid, Experiment 3), containing an oddball item among identical distractors. Participants had to indicate using a keypress the side on which the oddball item appeared (left vs right).

## Stimuli and search pairs

All stimuli were grayscale and scaled to subtend 5° along their longest dimension. Across the duration of each experiment, the target location was sampled equally across each of the four possible choices in the search array. Each set of unique trials (i.e., search pairs) were formed into a block when running the task. Within each block the trials were presented in random order. Thus, repetitions of unique trials were collected as blocks. Due to an unforeseen issue in the applied settings in NIMH MonkeyLogic, for some experiments (Experiments 1, 3 and 4 in monkeys), all conditions did not have equal repetitions, and we performed all analyses using all available trials for each condition. Environmental illumination was present via roof mounted lamps. Any deviations from the preceding statements are mentioned explicitly for each experiment.

### Experiment 1 – Perceptual organization

We selected four animate and four inanimate object images (Figure 2A). Unique search pairs consisted of all  ${}^8C_2 = 28$  image pairs. For each unique search pair, we created two search arrays with either item as the target and the other as the distractor for a total of 56 search pairs. Each search pair was presented 12-20 times to each monkey. This was compared to human data collected for the same stimuli and image pairs in a previously published study (Mohan and Arun, 2012).

### Experiment 2 – Weber’s Law for luminance

To test perceptual discriminability of luminance, we created circles of size 2° each and set the luminance levels based on a 16-bit scale (a.u.). Actual luminance values (in  $\text{cd/m}^2$ ) were measured post-hoc using a SpectraScan PR655 spectroradiometer (PhotoResearch Inc.). In each trial, the distractors’ luminance was either low-intensity (16384 a.u.,  $8.8 \text{ cd/m}^2$ ) or high-intensity (25600 a.u.,  $24.33 \text{ cd/m}^2$ ) and the targets’ luminance was greater than the distractor by a fixed difference in the range of 500 to 8000 a.u. (10-linearly spaced steps, ranging from 9.13 to  $21.66 \text{ cd/m}^2$  for low-intensity distractors and from 24.93 to  $45.30 \text{ cd/m}^2$  for high-intensity distractors). For each distractor luminance level, there were 10 trials with each of the targets for a total of 20 unique search pairs. Each search pair was presented 16 times to monkeys and 12 times to humans. The experiment was blocked to ensure that trials for a given distractor luminance level appeared together. We kept the rooms pitch dark to promote task performance.

### Experiment 3 – Amodal completion

As in Experiment 1, we obtained the stimuli and the relevant human data from a previous study (Cherian and Arun, 2024). To test amodal completion in monkeys, we selected 10 shapes (Figure 3A) and created 3 display conditions for each. In the ‘occluded’ displays an occluder blocked the top left quadrant of the shape from view whereas in the ‘likely’ and ‘mosaic’ displays the occluder was positioned so as to not occlude any portion of the shape. While the ‘likely’ displays showed the whole shape intact, in the ‘mosaic’ displays the top left quadrant of the shape was cropped out. For each shape we prepared two search pairs, with the ‘likely’ display always as the distractor and the ‘occluded’ or ‘mosaic’ display as the target. Thus, we created a total of 20 unique search pairs. Each search pair was presented 15-24 times to each monkey with varying target location.

### Experiment 4 – Mirror confusion

We selected images of 15 natural objects from the BOSS dataset (Brodeur et al., 2014). We labelled these images as ‘reference’ and created ‘vertical’ and ‘horizontal’ mirror images of each. Each search pair consisted of the ‘reference’ image as the distractor and either the ‘vertical’ or ‘horizontal’ mirror image as the target, for a total of 30 unique search pairs. All search pairs were presented 18-29 times to each monkey and 8 times to each human participant.

### Experiment 5 – Global advantage

Stimuli were obtained from a previous study (Jacob and Arun, 2020). Three basic shapes: ‘diamond’, ‘circle’ and ‘x’ were applied on both a global and local scale to create a set of 9 hierarchical shapes with all possible mixtures (Figure 6B). For any pair of these shapes either

the global, local or both shapes were different. We used all 36 unique pairs i.e., the 9 global same and local different (GSLD), 9 global different and local same (GDLS) and 18 global different and local different (GDLD) pairs. For each search pair we created two search arrays with either item as the target and the other as the distractor for a total of 72 total search pairs. Monkeys were presented with each of the search pairs 8 times whereas humans were presented with 4 repetitions.

### Experiment 6 – Global advantage with connected lines

Stimuli were the same as Experiment 5 except that we connected the local elements with a gray line with the same thickness as the lines making the local shapes (Figure 7A). We chose a grey colour for the connecting line to ensure local shapes remained individually distinguishable. All other experimental details were identical to Experiment 5.

### Data analyses

We analysed all data through custom scripts and functions in MATLAB. Only trials where humans or monkeys gave a valid response (correct or wrong) were included for all further analyses.

#### Response consistency

To evaluate response time consistency for humans and monkeys, we collated the response times across all participants and calculated the average response time between randomly dividing into two equal groups. For each search pair, outliers more than 3 scaled median absolute deviations from the median were removed for each search pair. Doing so improved the response consistency overall for both humans and monkeys. We confirmed that all results were qualitatively similar even without this step. We then calculated the response consistency as the split-half correlation as the average Pearson's correlation coefficient between 1000 random splits ( $r_{sh}$  with repeated sampling). The statistical significance of this correlation ( $p_{sh}$ ) was determined as follows:

$$p_{sh} = 1 - \frac{\text{number of significant } (p < 0.05) \text{ split - half correlations}}{\text{total number of random splits}}$$

#### Measuring perceptual dissimilarities

We analysed only response times (RT) from correct trials, averaged them across repetitions and took the reciprocal (1/RT) as a measure of perceptual dissimilarity. 1/RT behaves like a distance metric, varies linearly with feature differences, and can be interpreted as the underlying evidence accumulation signal driving search decisions (Arun, 2012; Pramod and Arun, 2014; Sunder and Arun, 2016).

#### Inter-participant consistency

To compare inter-individual differences across humans and monkeys, we calculated the correlation distance between every pair of participants as follows:

$$r_{i,j} = r(d_i, d_j)$$

where  $r_{i,j}$  is the Pearson's correlation between the set of dissimilarities,  $d$  (averaged across repetitions), from subjects  $i$  and  $j$ . A large correlation value implies that the two participants had a good match between their responses and vice-versa.

#### Sigmoid curve fitting

In order to compare psychometric curves, we fit a sigmoid by scaling and shifting the normal cumulative distribution function (CDF, 'normcdf' in MATLAB) as follows:

$$p = A * F(x | \mu, \sigma)$$

Where  $p$  is the predicted accuracy,  $A$  is the scaling factor, and  $F$  is the normal CDF with parameters:  $\mu$  and  $\sigma$ .  $x$  is the luminance difference and the values of  $A$ ,  $\mu$ , and  $\sigma$  were estimated using non-linear least squares fitting ('nlinfit' in MATLAB), with initial guess as max accuracy, and mean and standard deviation of the luminance levels, respectively.

Perceptual Phenomena	Monkeys	Humans
Object Relations	Yes	Yes
Weber's Law	Yes	Yes
Amodal Completion	Yes	Yes
Mirror Confusion	No	Yes
Global Advantage	No	Yes

**Table 1.** Summary of perceptual phenomena observed in monkeys vs humans

## Data availability

All data and code required to reproduce the results are available on OSF at <https://osf.io/63m5f>.

## Acknowledgements

We thank Aditya Murthy, Supratim Ray, and Sridharan Devarajan for their valuable inputs during yearly reviews. We thank Ankan Biswas and Divya Gulati for their help with the luminance measurements. We are grateful to Ramesh Venkatarayappa, Thirumala M, KV Ashok Kumar, and Ravi Kumar at the Primate Research Laboratory of the Indian Institute of Science for their excellent support with animal care.

## Additional information

### Author contributions

TC, GJ and SPA designed the experiments; TC and GJ collected the data; TC and SPA analysed and interpreted the data and wrote the manuscript with inputs from GJ.

### Funding

This research was supported by a DBT/Wellcome Trust India Alliance Senior Fellowship awarded to SPA (Grant# IA/S/17/1/503081), a Senior Research Fellowship from the Indian Council of Medical Research (ICMR) awarded to TC (3/1/3/JRF-2015/HRD-SS/30/92575/136) and a Senior Research Fellowship from the Ministry of Education (previously MHRD), Government of India, to GJ.

### Funding

Funder	Grant reference number	Author
DBT/Wellcome Trust India Alliance	IA/S/17/1/503081	Arun SP
MOHFW   DHR   Indian Council of Medical Research (ICMR)	3/1/3/JRF-2015/HRD-SS/30/92575/136	Thomas Cherian
Ministry of Education, Government of India		Georgin Jacob

### Author ORCID iDs

**Thomas Cherian:** <https://orcid.org/0000-0002-4910-6880>

**Georgin Jacob:** <https://orcid.org/0000-0001-8262-0155>

**SP Arun:** <https://orcid.org/0000-0001-9602-5066>

## Additional files

[Supplementary Video S1](#)

## References

- Agrillo C, Beran MJ, Parrish AE** (2019) Exploring the Jastrow illusion in humans (*Homo sapiens*), rhesus monkeys (*Macaca mulatta*), and capuchin monkeys (*Sapajus apella*). *Perception* **48**:367-385 <https://doi.org/10.1177/0301006619838181> | [PubMed](#)
- Arun SP** (2012) Turning visual search time on its head. *Vision Res* **74**:86-92 <https://doi.org/10.1016/j.visres.2012.04.005> | [PubMed](#)
- Barbet I, Fagot J** (2002) Perception of the corridor illusion by baboons (*Papio papio*). *Behav Brain Res* **132**:111-115 [https://doi.org/10.1016/s0166-4328\(01\)00393-x](https://doi.org/10.1016/s0166-4328(01)00393-x) | [PubMed](#)
- Behar E, Samuel D** (1982) Visual illusions in the baboon (*Papio anubis*). *Anim Learn Behav* **10**:115-118 <https://doi.org/10.3758/BF03212056>

- Bell AH, Hadj-Bouziane F, Frihauf JB, Tootell RB, Ungerleider LG (2009) Object representations in the temporal cortex of monkeys and humans as revealed by functional magnetic resonance imaging. *J Neurophysiol* **101**:688-700 <https://doi.org/10.1152/jn.90657.2008> | PubMed
- Beran MJ (2006) Quantity Perception by Adult Humans (*Homo sapiens*), Chimpanzees (*Pan troglodytes*), and Rhesus Macaques (*Macaca mulatta*) as a Function of Stimulus Organization. *Int J Comp Psychol* **19** <https://doi.org/10.46867/ijcp.2006.19.04.05>
- Beran MJ, Perdue BM, Kelly AJ, Parrish AE (2017) What's in a face (made of foods)? Comparing children's and monkey's perception of faces in face-like images of food. *Anim Behav Cogn* **4**:324-339 <https://doi.org/10.26451/abc.04.03.10.2017>
- Brodeur MB, Guérard K, Bouras M (2014) Bank of Standardized Stimuli (BOSS) phase ii: 930 new normative photos. *PLoS ONE* **9** <https://doi.org/10.1371/journal.pone.0106953> | PubMed
- Buffalo EA, Movshon JA, Wurtz RH (2019) From basic brain research to treating human brain disorders. *Proc Natl Acad Sci U S A* **116**:26247-26254 <https://doi.org/10.1073/pnas.1919895116> | PubMed
- Cherian T, Arun SP (2024) What do we see behind an occluder? Amodal completion of statistical properties in complex objects. *Atten Percept Psychophys* <https://doi.org/10.3758/s13414-024-02948-w> | PubMed
- Corballis MC (2018) Mirror-Image Equivalence and Interhemispheric Mirror-Image Reversal. *Front Hum Neurosci* **12**:140 <https://doi.org/10.3389/fnhum.2018.00140> | PubMed
- Dahl CD, Logothetis NK, Bühlhoff HH, Wallraven C (2010) The Thatcher illusion in humans and monkeys. *Proc R Soc B Biol Sci* **277**:2973-2981 <https://doi.org/10.1098/rspb.2010.0438> | PubMed
- De Valois RL, Morgan H, Snodderly DM (1974a) Psychophysical studies of monkey Vision-III. Spatial luminance contrast sensitivity tests of macaque and human observers. *Vision Res* **14**:75-81 [https://doi.org/10.1016/0042-6989\(74\)90118-7](https://doi.org/10.1016/0042-6989(74)90118-7) | PubMed
- De Valois RL, Morgan HC, Polson MC, Mead WR, Hull EM (1974b) Psychophysical studies of monkey vision—I. Macaque luminosity and color vision tests. *Vision Res* **14**:53-67 [https://doi.org/10.1016/0042-6989\(74\)90116-3](https://doi.org/10.1016/0042-6989(74)90116-3) | PubMed
- Deruelle C, Fagot J (1997) Hemispheric lateralisation and global precedence effects in the processing of visual stimuli by humans and baboons (*Papio papio*). *Laterality* **2**:233-246 <https://doi.org/10.1080/713754268> | PubMed
- Fagot J, Deruelle C (1997) Processing of global and local visual information and hemispheric specialization in humans (*Homo sapiens*) and baboons (*Papio papio*). *J Exp Psychol Hum Percept Perform* **23**:429-442 <https://doi.org/10.1037//0096-1523.23.2.429> | PubMed
- Fagot J, Tomonaga M (1999) Global and local processing in humans (*Homo sapiens*) and chimpanzees (*Pan troglodytes*): Use of a visual search task with compound stimuli. *J Comp Psychol* **113**:3-12 <https://doi.org/10.1037/0735-7036.113.1.3>
- Firestone C (2020) Performance vs. competence in human-machine comparisons. *Proc Natl Acad Sci U S A* **117**:26562-26571 <https://doi.org/10.1073/pnas.1905334117> | PubMed
- Freiwald AW, Tsao DY (2010) Functional Compartmentalization and Viewpoint Generalization Within the Macaque Face-Processing System. *Science* **330**:845-851 <https://doi.org/10.1126/science.1194908> | PubMed
- Fujita K (1996) Linear perspective and the Ponzo illusion: A comparison between rhesus monkeys and humans. *Jpn Psychol Res* **38**:136-145 <https://doi.org/10.1111/j.1468-5884.1996.tb00018.x>
- Fujita K (1997) Perception of the Ponzo illusion by rhesus monkeys, chimpanzees, and humans: Similarity and difference in the three primate species. *Percept Psychophys* **59**:284-292 <https://doi.org/10.3758/bf03211896> | PubMed
- Fujita K (2001) Perceptual completion in rhesus monkeys (*Macaca mulatta*) and pigeons (*Columba livia*). *Percept Psychophys* **63**:115-125 <https://doi.org/10.3758/bf03200507> | PubMed

- Hecht S (1924) The visual discrimination of intensity and the Weber-Fechner Law. *J Gen Physiol* **7**:235-267 <https://doi.org/10.1085/jgp.7.2.235> | PubMed
- Hiramatsu C, Goda N, Komatsu H (2011) Transformation from image-based to perceptual representation of materials along the human ventral visual pathway. *NeuroImage* **57**:482-494 <https://doi.org/10.1016/j.neuroimage.2011.04.056> | PubMed
- Hopkins WD, Washburn DA (2002) Matching visual stimuli on the basis of global and local features by chimpanzees (Pan troglodytes) and rhesus monkeys (Macaca mulatta). *Anim Cogn* **5**:27-31 <https://doi.org/10.1007/s10071-001-0121-8> | PubMed
- Hwang J, Mitz AR, Murray EA (2019) NIMH MonkeyLogic: Behavioral control and data acquisition in MATLAB. *J Neurosci Methods* **323**:13-21 <https://doi.org/10.1016/j.jneumeth.2019.05.002> | PubMed
- Jacob G, Arun SP (2020) How the forest interacts with the trees: Multiscale shape integration explains global and local processing. *J Vis* **20**:20 <https://doi.org/10.1167/jov.20.10.20> | PubMed
- Jacob G, Katti H, Cherian T, Das J, Zhivago KA, Arun SP (2021a) A naturalistic environment to study visual cognition in unrestrained monkeys. *eLife* **10**:e63816 <https://doi.org/10.7554/eLife.63816> | PubMed
- Jacob G, Pramod RT, Katti H, Arun SP (2021b) Qualitative similarities and differences in visual object representations between brains and deep networks. *Nat Commun* **12**:1872 <https://doi.org/10.1038/s41467-021-22078-3> | PubMed
- Jordan KE, Brannon EM (2006) Weber's Law influences numerical representations in rhesus macaques (Macaca mulatta). *Anim Cogn* **9**:159-172 <https://doi.org/10.1007/s10071-006-0017-8> | PubMed
- Kimchi R (1992) Primacy of wholistic processing and global/local paradigm: a critical review. *Psychol Bull* **112**:24-38 <https://doi.org/10.1037/0033-2909.112.1.24> | PubMed
- Kriegeskorte N, Mur M, Bandettini P (2008a) Representational similarity analysis - connecting the branches of systems neuroscience. *Front Syst Neurosci* **2**:4 <https://doi.org/10.3389/neuro.06.004.2008> | PubMed
- Kriegeskorte N, Mur M, Ruff DA, Kiani R, Bodurka J, Esteky H, Tanaka K, Bandettini PA (2008b) Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron* **60**:1126-1141 <https://doi.org/10.1016/j.neuron.2008.10.043> | PubMed
- Leopold DA, Logothetis NK (1996) Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature* **379**:549-553 <https://doi.org/10.1038/379549a0> | PubMed
- Logothetis NK, Schall JD (1989) Neuronal correlates of subjective visual perception. *Science* **245**:761-763 <https://doi.org/10.1126/science.2772635> | PubMed
- Matsuno T, Fujita K (2009) A comparative psychophysical approach to visual perception in primates. *Primates* **50**:121-130 <https://doi.org/10.1007/s10329-008-0128-8> | PubMed
- Mitchell S (1988) *Tao te ching: A new English version* Harper Collins.
- Mohan K, Arun SP (2012) Similarity relations in visual search predict rapid visual categorization. *J Vis* **12**:19-19 <https://doi.org/10.1167/12.11.19> | PubMed
- Navon D (1977) Forest before trees: The precedence of global features in visual perception. *Cogn Psychol* **9**:353-383 [https://doi.org/10.1016/0010-0285\(77\)90012-3](https://doi.org/10.1016/0010-0285(77)90012-3)
- Nieder A, Miller EK (2003) Coding of cognitive magnitude: compressed scaling of numerical information in the primate prefrontal cortex. *Neuron* **37**:149-157 [https://doi.org/10.1016/s0896-6273\(02\)01144-3](https://doi.org/10.1016/s0896-6273(02)01144-3) | PubMed
- Nielsen KJ, Logothetis NK, Rainer G (2006) Discrimination Strategies of Humans and Rhesus Monkeys for Complex Visual Displays. *Curr Biol* **16**:814-820 <https://doi.org/10.1016/j.cub.2006.03.027> | PubMed
- Ono H (1967) Difference threshold for stimulus length under simultaneous and nonsimultaneous viewing conditions. *Percept Psychophys* **2**:201-207 <https://doi.org/10.3758/bf03213050>

- Pardo-Vazquez JL**, Castiñeiras-de Saa JR, Valente M, Damião I, Costa T, Vicente MI, Mendonça AG, Mainen ZF, Renart A (2019) The mechanistic foundation of Weber's law. *Nat Neurosci* **22**:1493-1502 <https://doi.org/10.1038/s41593-019-0439-7> | PubMed
- Parrish AE**, Beran MJ, Agrillo C (2019) Linear numerosity illusions in capuchin monkeys (*Sapajus apella*), rhesus macaques (*Macaca mulatta*), and humans (*Homo sapiens*). *Anim Cogn* **22**:883-895 <https://doi.org/10.1007/s10071-019-01288-9> | PubMed
- Parron C**, Fagot J (2007) Comparison of Grouping Abilities in Humans (*Homo sapiens*) and Baboons (*Papio papio*) With the Ebbinghaus Illusion. *J Comp Psychol* **121**:405-411 <https://doi.org/10.1037/0735-7036.121.4.405> | PubMed
- Passingham R** (2009) How good is the macaque monkey model of the human brain?. *Curr Opin Neurobiol* **19**:6-11 <https://doi.org/10.1016/j.conb.2009.01.002> | PubMed
- Pramod RT**, Arun SP (2014) Features in visual search combine linearly. *J Vis* **14**:1-20 <https://doi.org/10.1167/14.4.6> | PubMed
- Pramod RT**, Arun SP (2016a) Object attributes combine additively in visual search. *J Vis* **16**:8 <https://doi.org/10.1167/16.5.8> | PubMed
- Pramod RT**, Arun SP (2016b) Do computational models differ systematically from human object perception?. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 1601-1609 <https://doi.org/10.1109/cvpr.2016.177>
- Quiroga RQ**, Reddy L, Kreiman G, Koch C, Fried I, Quian Quiroga R, Reddy L, Kreiman G, Koch C, Fried I (2005) Invariant visual representation by single neurons in the human brain. *Nature* **435**:1102-1107 <https://doi.org/10.1038/nature03687> | PubMed
- Rajalingham R**, Issa EB, Bashivan P, Kar K, Schmidt K, DiCarlo JJ (2018) Large-Scale, High-Resolution Comparison of the Core Visual Object Recognition Behavior of Humans, Monkeys, and State-of-the-Art Deep Artificial Neural Networks. *J Neurosci Off J Soc Neurosci* **38**:7255-7269 <https://doi.org/10.1523/jneurosci.0388-18.2018> | PubMed
- Rajalingham R**, Schmidt K, DiCarlo JJ (2015) Comparison of Object Recognition Behavior in Human and Monkey. *J Neurosci* **35**:12127-12136 <https://doi.org/10.1523/jneurosci.0573-15.2015> | PubMed
- Rensink RA**, Enns JT (1998) Early completion of occluded objects. *Vision Res* **38**:2489-2505 [https://doi.org/10.1016/s0042-6989\(98\)00051-0](https://doi.org/10.1016/s0042-6989(98)00051-0) | PubMed
- Riopelle AJ**, Rahm U, Itoigawa N, Draper WA (1964) Discrimination of Mirror-Image Patterns by Rhesus Monkeys. *Percept Mot Skills* **19**:383-389 <https://doi.org/10.2466/pms.1964.19.2.383> | PubMed
- Roelfsema PR**, Treue S (2014) Basic neuroscience research with nonhuman primates: a small but indispensable component of biomedical research. *Neuron* **82**:1200-1204 <https://doi.org/10.1016/j.neuron.2014.06.003> | PubMed
- Rollenhagen JE**, Olson CR (2000) Mirror-Image Confusion in Single Neurons of the Macaque Inferotemporal Cortex. *Science* **287**:1506-1509 <https://doi.org/10.1126/science.287.5457.1506> | PubMed
- Spinozzi G**, De Lillo C, Salvi V (2006) Local advantage in the visual processing of hierarchical stimuli following manipulations of stimulus size and element numerosity in monkeys (*Cebus apella*). *Behav Brain Res* **166**:45-54 <https://doi.org/10.1016/j.bbr.2005.06.043> | PubMed
- Sripati AP**, Olson CR (2009) Representing the forest before the trees: a global advantage effect in monkey inferotemporal cortex. *J Neurosci* **29**:7788-7796 <https://doi.org/10.1523/jneurosci.5766-08.2009> | PubMed
- Sripati AP**, Olson CR (2010) Global Image Dissimilarity in Macaque Inferotemporal Cortex Predicts Human Visual Search Efficiency. *J Neurosci* **30**:1258-1269 <https://doi.org/10.1523/jneurosci.1908-09.2010> | PubMed
- Sunder S**, Arun SP (2016) Look before you seek: Preview adds a fixed benefit to all searches. *J Vis* **16**:3 <https://doi.org/10.1167/16.15.3> | PubMed

- Tanaka H, Fujita I (2000) Global and local processing of visual patterns in macaque monkeys. *Neuroreport* **11**:2881-2884 <https://doi.org/10.1097/00001756-200009110-00010> | PubMed
- Tang H, Buia C, Madhavan R, Crone NE, Madsen JR, Anderson WS, Kreiman G (2014) Spatiotemporal Dynamics Underlying Object Completion in Human Ventral Visual Cortex. *Neuron* **83**:736-748 <https://doi.org/10.1016/j.neuron.2014.06.017> | PubMed
- Taubert J, Wardle SG, Flessert M, Leopold DA, Ungerleider LG (2017) Face Pareidolia in the Rhesus Monkey. *Curr Biol CB* **27**:2505-2509. <https://doi.org/10.1016/j.cub.2017.06.075> | PubMed
- Taylor AH, Bastos APM, Brown RL, Allen C (2022) The signature-testing approach to mapping biological and artificial intelligences. *Trends Cogn Sci* **26**:738-750 <https://doi.org/10.1016/j.tics.2022.06.002> | PubMed
- Tudusciuc O, Nieder A (2010) Comparison of length judgments and the Müller-Lyer illusion in monkeys and humans. *Exp Brain Res* **207**:221-231 <https://doi.org/10.1007/s00221-010-2452-7> | PubMed
- Washburn DA (1993) The Stimulus Movement Effect: Allocation of Attention or Artifact?. *J Exp Psychol Anim Behav Process* **19**:380-390 <https://doi.org/10.1037//0097-7403.19.4.380> | PubMed
- Weinstein B, Grether WF (1940) A comparison of visual acuity in the rhesus monkey and man. *J Comp Psychol* **30**:187-195 <https://doi.org/10.1037/h0061507>
- Weldon KB, Taubert J, Smith CL, Parr LA (2013) How the Thatcher illusion reveals evolutionary differences in the face processing of primates. *Anim Cogn* **16**:691-700 <https://doi.org/10.1007/s10071-013-0604-4> | PubMed
- Zhivago KA, Arun SP (2014) Texture discriminability in monkey inferotemporal cortex predicts human texture perception. *J Neurophysiol* **112**:2745-2755 <https://doi.org/10.1152/jn.00532.2014> | PubMed

## Peer reviews

### Reviewer #1 (Public review):

#### Summary:

The authors set out to conduct a behavioral comparison of macaque and human vision across a wide range of visual properties. Such a comparison is critical for evaluating the use of macaques as a model system for understanding human vision and the underlying neural mechanisms. This goal represents a unique endeavour since prior studies have typically focused on only highly specific tasks. While the authors found consistent coarse representational structure for objects, evidence for Weber's Law and amodal completion, there was divergence for mirror image confusion and the use of global or local image properties.

#### Strengths:

There are three major strengths of the study. First, the authors employed a behavioral paradigm (oddball search) that allowed them to test multiple perceptual phenomena without having to train the macaques on the specific type of stimuli tested. Second, humans and macaques could be tested in an identical manner. Third, the authors tested a range of different visual properties and phenomena, allowing a broader comparison between species.

There are also some weaknesses to the study (described below), but that doesn't change the fact that the authors have demonstrated and validated a novel approach for systematic and comprehensive comparisons of vision across species.

#### Weaknesses:

The weaknesses of the study arise in part because of the breadth of the work. In cases where there are divergences between the two species, it would be helpful to know what might account for such divergence, to have more depth. Is it really a species difference, or could there be a different account? For example, does the difference in mirror image confusion arise because the stimuli were objects that would have been highly familiar to the humans but not the macaques? Further, the authors often used small sets of stimuli (e.g. 8 objects only in the test of object similarity; a small set of highly specific occlusion stimuli), and how well the findings will generalize beyond those stimuli is unclear.

The authors discuss the implications of training macaques to perform specific tasks on specific stimuli in comparing across species. While I agree that extensive training in monkeys could change perception, it is important to also consider that humans have been extensively trained through the types of visual tasks we conduct throughout our lives, so I'm not sure it is universally true that the best comparison is between humans and untrained monkeys. But this just consideration just highlights the general problem of comparing across species.

<https://doi.org/10.7554/eLife.110743.1.sa2>

## Reviewer #2 (Public review):

### Summary:

The macaque monkey is often considered as the animal model of choice to study the neural correlates of visual perception, due to the close similarities to humans in terms of anatomy, physiology and behaviour (Van Essen and Dierker, 2007; DiCarlo et al., 2012; Roelfsema and Treue, 2014; Picaud et al., 2019; Van Essen et al., 2019; Hesse and Tsao, 2020). Quite some studies have been performed to compare the behaviour of macaque monkeys and humans on visual perception tasks. However, it remains difficult to compare the results of these studies as the methods that are used differ significantly between these studies. Furthermore, behavioural studies of macaque monkeys often involve extensive training as the tasks were relatively hard, making it difficult to compare the results with humans, who generally require very little training. The authors present a set of experiments to compare visual perception between macaque monkeys and humans, using the exact same behavioral task that is easy to learn and therefore requires very little training. As expected, they overall find that the two species behave similarly. However, they find a number of interesting exceptions.

### Strengths:

A major strength of the current study is the relatively large number of tasks that were tested in the same subjects. This is made possible by using the oddball visual search task, which macaque monkeys can learn very quickly. This means that few trials are sufficient to obtain a significant difference between conditions, minimizing learning effects. Although this type of task has been used in previous studies (Sablé-Meyer et al., 2021), the current manuscript makes better use of the advantages and explains them more explicitly.

In addition, the study finds a number of interesting differences between macaque monkeys and humans. In particular, while humans can dissociate horizontally mirrored images better than vertically mirrored images, monkeys show no difference between these two conditions (Experiment 4). Also, while humans dissociate images better based on the global shape of a stimulus, monkeys dissociate images better based on local elements of a stimulus (Experiments 5 and 6). Although these findings are largely a replication of previous results, they have not yet been studied together with other tasks within the same individual subjects, and the low number of trials avoids any learning effects.

### Weaknesses:

A weakness of the study is that while the objects that were used can be considered to be familiar to humans, they are not familiar to macaque monkeys.

In Experiment 4, humans can be expected to have 3D representations of familiar objects such as a Roman helmet or an office chair. Humans can therefore be expected to have view-invariant representations of these objects, predominantly for rotations around the vertical axis of the object (as movements are most common in the horizontal plane). This can explain why only humans confuse objects more often when mirrored vertically than when mirrored horizontally.

Similarly, in Experiment 5, humans can be expected to be familiar with abstract geometric shapes such as squares and circles, while monkeys likely are not. This could explain why monkeys find it hard to recognize these geometric shapes in the global shape of the stimuli, even when thin grey lines are drawn to connect the local elements that constitute the global shape (Experiment 6). Instead, the combination of local shapes can be expected to form a texture that might be more easily recognized by the monkeys.

More generally, as proposed by Fagot et al, it might well be that monkeys tend to conceive stimuli as a combination of low-level visual features, instead of as references to objects in the outside world, as humans have learned to do (Fagot et al., 2010). This line of critique would be relevant to take into account.

Another weakness could be that only three monkeys are tested, while 24 human subjects are tested. According to some theoretical work, a finding in 3 animals is not sufficient to make a claim about an animal species (Fries and Maris, 2022). However, it seems that the results are largely consistent between the different monkeys. Moreover, the results generally agree with the results from previous literature.

The conclusions by the authors are therefore largely supported by the results. Some results could be strengthened by additional experiments, or at least a more extensive discussion of the potential weaknesses.

The potential impact of the paper is significant, as a start of a comprehensive comparison of visual perception between humans and macaque monkeys, which can inspire other labs to contribute to. This comparison can also be extended to other animal species (e.g. crows and rodents), as well as to different types of artificial neural networks (Leibo et al., 2018).

<https://doi.org/10.7554/eLife.110743.1.sa1>

### **Reviewer #3 (Public review):**

Summary:

In this study, Cherian and colleagues compare visual perception between humans and monkeys using a common oddball visual search task across a battery of perceptual phenomena. By keeping the task constant and varying only the stimulus sets, the authors aim to isolate perceptual similarities and differences between species. Across six experiments, they report that monkeys and humans share similarities in coarse object representations, Weber's law, and amodal completion, but differ in mirror confusion and global/local processing.

Strengths:

A major strength of the study is the unified experimental framework. The authors designed the experiments such that the task procedures are identical across conditions and species, differing only in the images shown. This is a significant methodological advantage, as it minimizes task-related confounds that often complicate cross-species and cross-experiment

comparisons. As a result, observed similarities and differences can be more directly attributed to perceptual processes rather than differences in training or task demands. This allows for a more comprehensive evaluation of visual perception than is typical in the literature, where individual studies often focus on a single effect with specialized training. The data are carefully collected, and the analyses are systematic and appropriate for the questions posed.

#### Weaknesses:

Despite its strengths, the study is largely descriptive and provides limited mechanistic or theoretical explanation for the observed similarities and differences. While the authors document several convergences and divergences between humans and monkeys, there is relatively little discussion of why these patterns arise or how they relate to existing theories of visual processing. As a result, it is difficult to assess the broader implications or generalizability of the findings beyond the specific task and stimuli used.

Relatedly, the rationale for selecting the particular set of perceptual phenomena is not fully developed. Some tasks appear motivated by prior work comparing humans and deep neural networks, but it is unclear whether this set constitutes a representative or theoretically grounded sampling of visual perception. Without a clearer justification, it is difficult to interpret the absence or presence of specific effects (e.g., mirror confusion or global advantage) as reflecting fundamental species similarities/differences.

<https://doi.org/10.7554/eLife.110743.1.sa0>