RESEARCH ARTICLE

# Influence of global and local features on parallel object identification [version 1; referees: 4 approved with reservations]

Alvydas Šoliūnas

Institute of Biosciences, Vilnius University, Vilnius, LT-10257, Lithuania

**v1** **First published:** 24 May 2018, **7**:660 (https://doi.org/10.12688/f1000research.14468.1)
**Latest published:** 24 May 2018, **7**:660 (https://doi.org/10.12688/f1000research.14468.1)

## Abstract

**Background:** The present study concerns parallel and serial processing of visual information, or more specifically, whether visual objects are identified successively or simultaneously in multiple object stimulus. Some findings in scene perception demonstrate the potential parallel processing of different sources of information in a stimulus; however, more extensive investigation is needed.

**Methods:** We presented one, two or three visual objects of different categories for 100 ms and afterwards asked subjects whether a specified category was present in the stimulus. We varied the number of objects, the number of categories and the type of object shape distortion (distortion of either global or local features).

**Results:** The response time and accuracy data corresponded to data from a previous experiment, which demonstrated that performance efficiency mostly depends on the number of categories but not on the number of objects. Two and three objects of the same category were identified with the same accuracy and the same response time, but two objects were identified faster and more accurately than three objects if they belonged to different categories. Distortion type did not affect the pattern of performance.

**Conclusions:** The findings suggest the idea that objects of the same category can be identified simultaneously and that identification involves both local and global features.

## Keywords

visual perception, object identification, multiple objects, global features

**Open Peer Review**

**Referee Status:** ? ? ? ?

| | Invited Referees 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| **version 1** published 24 May 2018 | ? report | ? report | ? report | ? report |

1 **Dan McCarthy**, Brown University, USA

2 **Thiago Monteiro de Paiva Fernandes**, Federal University of Paraíba, Brazil

3 **Denis V. Yavna**, Southern Federal University, Russian Federation

4 **Mohd Izzuddin Hairol**, National University of Malaysia (UKM), Malaysia

**Discuss this article**

Comments (0)

**Corresponding author:** Alvydas Šoliūnas (alvydas.soliunas@gf.vu.lt)

**Author roles: Šoliūnas A**: Conceptualization, Data Curation, Formal Analysis, Investigation, Software, Writing – Original Draft Preparation

**Competing interests:** No competing interests were disclosed.

**Grant information:** The author(s) declared that no grants were involved in supporting this work.



**How to cite this article:** Šoliūnas A. **Influence of global and local features on parallel object identification [version 1; referees: 4 approved with reservations]** *F1000Research* 2018, **7**:660 (https://doi.org/10.12688/f1000research.14468.1)

**First published:** 24 May 2018, **7**:660 (https://doi.org/10.12688/f1000research.14468.1)

## Introduction

Visual information is processed in the brain in both parallel and serial. This is true not only for anatomical reason (e.g. parallel neuronal pathways from the retina; Merigan & Mounsell, 1993) but also for perceptual phenomena. An example of parallel perceptual processing is target detection in a visual search task. Features such as color, orientation, motion, spatial frequency and stereodepth are detected in parallel (Enns & Rensink, 1991; Gilchrist *et al.*, 1997; Treisman & Gelade, 1980; Wolfe, 1994), whereas ambiguous figures and stimuli in a binocular rivalry are perceived in serial (Alais & Blake, 2004; Leopold & Logothesis, 1999). The present study concerns the identification of visual objects in multiple object stimuli and whether objects are identified in parallel or serial mode.

In scene perception, we have empirical findings that demonstrate the parallel processing of a scene category, i.e. gist perception together with perception of objects in a scene (Brandman & Peelen, 2017; Gagne & MacEvoy, 2014; Hollingworth & Henderson, 1999; Joubert *et al.*, 2007; Joubert *et al.*, 2008). Rousselet *et al.* (2002, 2004a) demonstrate parallel processing of several scenes presented simultaneously. Rousselet & colleagues (2004b) argue, based on neurophysiological findings that neurons of infero-temporal cortex with large receptive field could encode the identity of several objects in parallel and theoretically simultaneous identification of several objects is possible. On the other hand, scene identification influences identification of objects in the scene and the identification of objects influences identification of a scene (Davenport & Potter, 2004; Davenport, 2007; Joubert *et al.*, 2007; Joubert *et al.*, 2008; Mack & Palmeri, 2010). Such influences suppose successive perceptual processes.

Much less research has been conducted on interaction between objects presented simultaneously without scene context. Gronau *et al.* (2008) (see also Auckland *et al.*, 2007; Green & Hummel, 2006) demonstrated the facilitation effect of semantically and spatially related objects on their identification in a study in which two semantically related or unrelated objects were presented in congruent or incongruent spatial relation.

If objects are identified successively, we can expect to observe a direct dependence of reaction time on the number of objects. Successive identification permits interaction between objects that could result in the nonlinear dependence of reaction time on the number of objects when reaction time depends on factors such as the similarity of objects and belongingness to the same or different categories. Our recent study demonstrated that object identification time depends much more on the number of categories than on the number of objects upon multiple object stimulus (Soliunas *et al.*, 2018). One, two or three objects (pictures of 10 categories of man-made objects) were presented simultaneously for 100 ms and then followed by a name of a category. Subjects were asked to answer whether objects of this category were present in a stimulus. Performance accuracy and reaction time did not depend on the number of objects if the objects belonged to the same category.

The present study is further verification of the hypothesis that objects of the same category are identified simultaneously. For better control of the global features of objects, new categories were selected and new objects were produced. A shape is a possible feature of an object that could enable parallel identification. To verify this possibility, we manipulated the global and local features of objects by distorting them. We predicted that the parallel identification of objects would be observed for intact objects, but not for globally distorted objects.

## Methods

### Subjects

All subjects were invited to participate in the study by personal discussion. Between February and May 2015, a total of 58 volunteer students from Vilnius University agreed to took part in the experiment (44 females and 14 males; 20–22 years of age). Each subject had normal or corrected-to-normal vision and verbally confirmed that had no prior experience with psychophysical testing of a similar nature. The subjects were not informed about the specific goals of this particular experiment. All subjects signed an informed consent form approved by the Lithuanian Bioethics Committee (consent form No. ASI12, approval No. 158200-13-578-173; issued by the Vilnius Region Ethics Committee of Biomedical Research, Vilnius, Lithuania). All subjects took part in one experimental session.

### Stimuli

A total of 62 objects of 10 categories (shoe, cap, clock, ashtray, cup, table, telephone, vase, mirror, and kettle/teapot) were selected from internet search engines in such a way that the shape of an object was not the exclusive feature of particular category. When selecting objects and creating stimuli of multiple objects, seven outline shapes were taken into account: “8”-shaped, circle, ellipse, square, elongated, triangle and “L”-shaped. Objects of each category had at least three outline shapes and each outline shape had at least three categories of objects.

All objects were transformed into grayscale pictures and resized in such a way that fitted into a 100×100-pixel area. There were eight types of stimuli that varied in the number of objects (one, two or three), number of categories (one, two or three), and number of outline shapes (one or two) (Figure 1): i) “1-1” stimuli (one object); ii) “1-2” stimuli (one category, two objects of different shape); iii) “1-3” stimuli (one category, three objects of two shapes); iv) “2-2s” stimuli (two categories, two objects of the same shape); v) “2-2d” stimuli (two categories, two objects of different shapes); vi) “2-3s” stimuli (two categories, two objects of the same category and the same shape and third object of different category and different shape); vii) “2-3d” stimuli (two categories, two objects of the same category but different shapes and third object of different category but the same shape as one of the two objects of the first category); and viii) “3-3” stimuli (three categories, three objects of two shapes). In total, 10 stimuli of each type were created, giving 80 stimuli altogether (Supplementary File 1).

The objects were placed into a 200×200-pixel area around a fixation point that was located at the center of this area. Stimuli

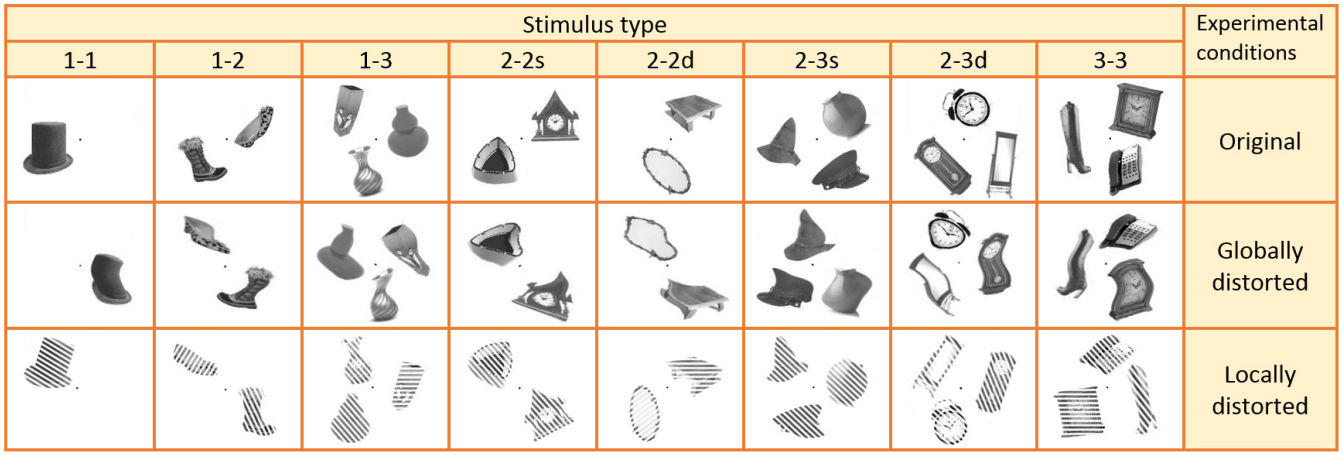


**Figure 1. Example of stimulus.** Columns represent different types of stimulus; rows represent experimental conditions. Each individual column represents the same objects under three experimental conditions.

were presented at the center of screen on the white background and subjects did not see the limits of the 200×200-pixel stimulus area. Distance between the subject's eyes and the screen was 60 cm, and consequently the angular size of the 200×200-pixel stimulus was 8°×8°. The orientation of a particular object was not constant across stimuli and could rage between −45° and +45° with respect to natural (vertical or horizontal) orientation.

Stimuli were presented under three experimental conditions: i) original; ii) locally distorted; and iii) globally distorted. The "original" condition corresponds to the presentation of stimuli described above. Locally distorted stimuli were created by partially masking the original stimuli with white stripes: 9-pixel-wide stripes with 9-pixel gaps (Figure 1). This procedure partially or completely eliminates some local features of objects but basically preserve the outline shape. The smaller the features, the higher the probability of elimination. Globally distorted objects were created by applying *Whirl and pinch* and *Ripple* functions in the image editor GIMP 2.8.10 (Kimball *et al.*, 2013). The same values of these functions distort objects of different shapes to different degrees; we therefore had to apply different values of these functions to more-or-less subjectively equalize the assessed degree of distortion in different objects. *Whirl and pinch* values ranged from −80 to +80 for elongated and rounded shapes and from −200 to +200 for more angular shapes, and *pinch amount* ranged from −1 to +1. *Ripple* values ranged from 40 to 70. The applied global distortion procedure affects outline shape and to a lesser degree the local elements of object. The smaller the elements, the less distortion there is.

To reduce memorization of stimuli during experiment, the orientation and location of particular object in particular stimulus varied across conditions.

## Procedure

The experiment was performed at the Department of Neurobiology and Biophysics, Vilnius University. Experimental sessions were conducted during daytime (the precise time of the day was not controlled) in a room with natural daylight illumination.

Stimulus presentation and data registration were controlled by E-Prime v.2.0 (Psychology Software Tools, Inc., 2012) experiment generator running on Windows OS. Stimuli were presented on the screen of 19-inch CRT monitor running at 85 Hz frame-rate and 1024×768 resolution. The subject's head was not fixed but they were instructed to hold the same distance (about 60 cm) from the display during experiment.

Before the experimental session, subjects performed practice session that consisted of 16 trials (two trials of each stimulus type). Only original stimuli were presented during practice.

The trial procedure of experimental session is shown in Figure 2. A fixation point was presented at the center of screen for 306 ms and the subjects were asked to keep their eyes focused on the fixation point during the test stimulus presentation. Appearance of the fixation point was followed by a 106-ms blank interval and then a test stimulus was displayed for 106 ms (i.e. for 9 frames of the CRT monitor) under "original" conditions and for 200 ms (i.e. 17 frames) under "locally distorted" and "globally distorted" conditions. The longer stimulus exposition duration under the two conditions with distorted stimuli were chosen to equalize response accuracy under all conditions. The test stimulus was followed by a 35-ms blank interval and then a masking pattern was displayed for 306 ms (we used backward masking procedure to control the time available for object identification). The masking pattern was an 8°×8° square of chaotic pattern. After 35 ms blank interval, a probe-word was presented. The probe-word was a name of a category written in lowercase Arial font, 2° height. Subjects had to decide whether an object defined by a probe-word was present or not on a given trial by pressing the "1" or "2" key on the right side of a keyboard. One half of subjects received the instruction to press the "1" for Yes and "2" for No, whereas the other half received inverse instruction. Subjects had four seconds to make their decision. The response time

(the duration from onset of probe-word to the keypress event) and accuracy were recorded for each trial. The response initialized the next trial with a 106-ms delay.

The order of experimental conditions was randomized across participants. There were 60-s rest intervals between conditions. Each condition consisted of 160 trials, i.e. 80 stimuli were presented twice in random order. Altogether, 480 stimuli were presented in the experimental session, with eight types of stimuli presented randomly under each condition. The whole experimental session lasted about 30 min.

### Statistical analysis

The chi-square goodness of fit test confirmed normal distribution of experimental data. The reaction time and response accuracy data were analyzed in two-way ANOVA for stimulus type ("1-1", "1-2", "1-3", "2-2", "2-3", and "3-3") and experimental conditions (original, locally distorted, and globally distorted). Initially, there were eight stimulus types, but as there were no statistical differences between performance for the "2-2s" and "2-2d" stimuli, we merged these results into one group "2-2". For the same reason, we merged results of "2-3s" and "2-3d" into one group "2-3". Newman–Keuls post hoc test was applied to assess the significance of differences between means. All statistical analysis was performed using Statistica v.7 software (StatSoft Inc., 2004).

## Results

The results of the experiment are presented in Figure 3 and Dataset 1. Two-way (stimulus type and experimental conditions) ANOVA indicated significant main effects of: stimulus type ($F(5,27822) = 281.7$, $P < 0.0001$ for reaction time (RT) data and $F(5,27822) = 203.4$, $P < 0.0001$ for response accuracy); experimental conditions ($F(2,27822) = 27.7$, $P < 0.0001$ for RT data and $F(2,27822) = 49.5$, $P < 0.0001$ for response accuracy). The interaction of the two factors was not significant ($F(10,27822) = 1.7$, $P = 0.079$ for RT and $F(10,27822) = 1.2$, $P = 0.272$ for accuracy data), which means that stimulus distortion did not change the pattern of performance that was observed for original stimuli. The significant main effect of the experimental conditions indicates that the RT was shorter (730 ms) and the accuracy was higher (86,4%) for original stimuli than for locally or globally distorted stimuli (760 ms and 80.9% for locally distorted stimuli and 762 ms and 82.1% for globally distorted stimuli), but this finding is not notable because the duration of stimulus exposition was different under different

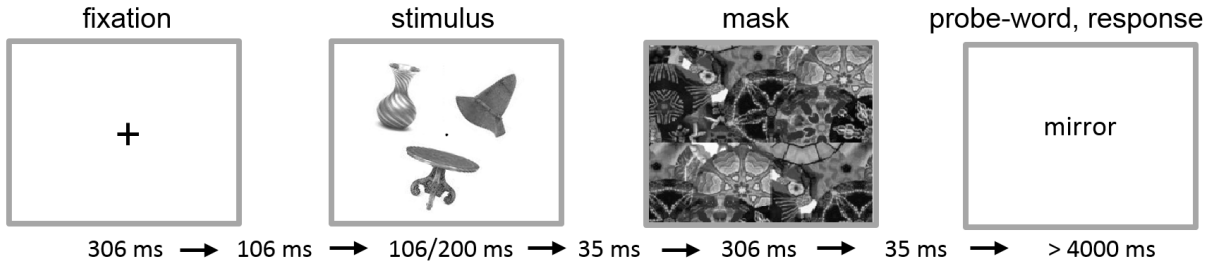


**Figure 2. Example trial of experiment.**

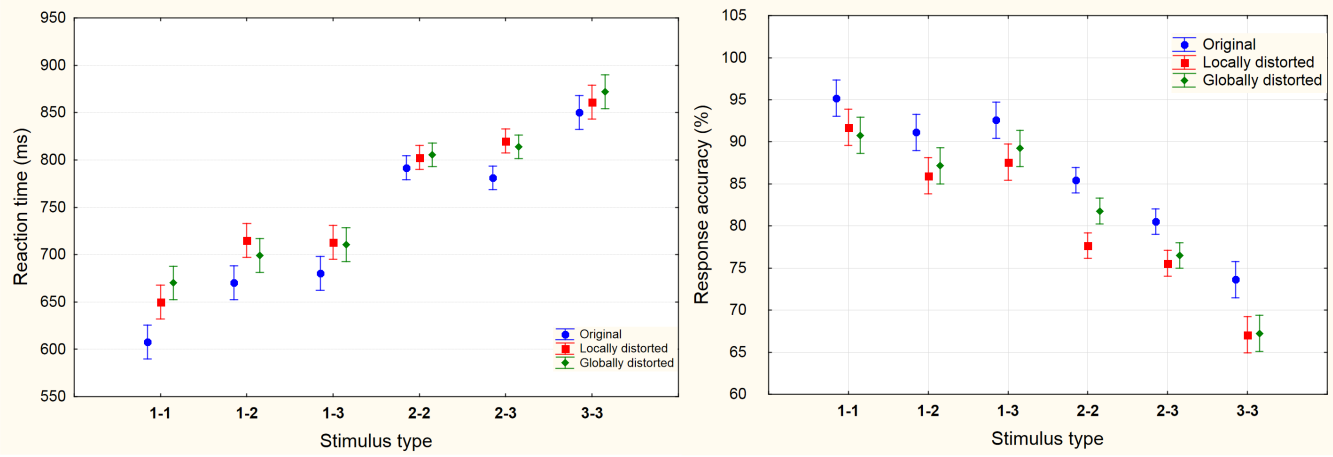


**Figure 3. Dependence of reaction time and response accuracy on stimulus type and experimental conditions (distortion type).** Mean values are presented with 95% confidence intervals. Stimulus type: the first digit represents the number of categories in the stimulus, the second digit represents the number of objects (e.g. "1-3" represents three objects of one category).

conditions and the absolute values of performance under different conditions are irrelevant. What we are interested in is the dependency of the identification of objects on the number of objects and on the number of categories.

Figure 3 reveals the influence of the number of objects and the influence of the number of categories on object identification. For RT data, we can see four statistically different levels of performance (we should stress again that we compare values between "stimulus types" but not between conditions): the shortest RT is for "1-1" stimuli, with longer RTs for "1-2" and "1-3" stimuli, even longer RTs for "2-2" and "2-3" stimuli, and the longest RT for "3-3" stimuli. There is no significant difference between "1-2" and "1-3" cases (P = 0.391, 0.876, and 0.329 for "Original", "Locally distorted", and "Globally distorted" conditions, respectively, according to analysis using the Newman–Keuls post hoc test) and between "2-2" and "2-3" cases (P = 0.363, 0.442 and 0.472 for "Original", "Locally distorted", and "Globally distorted" conditions, respectively). The same four levels of performance were found under all conditions.

For accuracy data, we can see a similar pattern of performance: the highest accuracy is in "1-1" case, with lower accuracy in "1-2" and "1-3" stimulus types, even lower accuracy in the "2-2" and "2-3" stimulus type, and the lowest accuracy in the "3-3" stimulus type. Here we can see two deviations from this rule: accuracy was higher in "2-2" case than in "2-3" case under "Original" (P < 0.01) and under "Globally distorted" (P < 0.01) conditions.

Summarization of the performance in relation to the number of objects and on the number of categories without differentiating experimental conditions is presented in Figure 4.

The dependence of performance effectiveness on the number of categories is more clearly expressed than the dependence of performance effectiveness on the number of objects. For one-category stimuli, there was no difference in RT and accuracy whether two or three objects were presented. For two-category stimuli, there was no difference in RT whether two or three objects were presented, but accuracy was higher in the case of two objects. We can state that "the more categories, the poorer the performance", but not that "the more objects, the poorer the performance", because it depends on whether objects belong to the same or to different categories.

**Dataset 1. Response time and accuracy data**

http://dx.doi.org/10.5256/f1000research.14468.d204177

"1-1", "1-2", "2-2s", "2-2d", "1-3", "2-3s", "2-3d", "3-3" are different types of stimuli where the first digit between quotation marks indicates the number of categories and the second digit indicates the number of objects. Original, original conditions; locally d., locally distorted conditions; globally d., globally distorted conditions; RT, reaction time; %, percentage accuracy. (Šoliūnas, 2018).

## Discussion

The present study is continuation of our previous investigation (Soliunas *et al.*, 2018) described in the Introduction, which findings suggested that the objects of the same category could be identified in parallel mode. The experiment described here further tested this hypothesis.

The first important result is the replication of the principal findings of the previous experiment, despite the fact that a different set of stimuli were used (all stimuli were newly created) and a different group of subjects took part in the experiment. These findings indicate that the identification of objects in multiple object stimuli basically depends on the number of categories present, but not on the number of objects. It further supports the suggestion that objects of the same category are identified simultaneously.

The second aim of the study was to search for the features of stimuli that could enable the parallel identification of objects of the same category, i.e. searching the "category" features that are

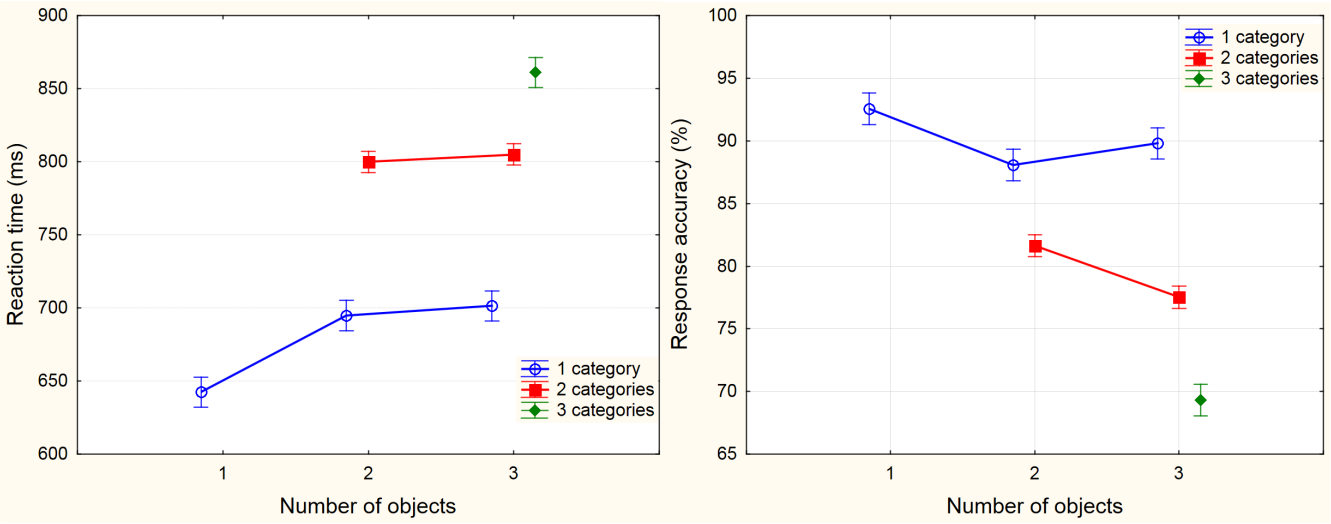


**Figure 4. Dependence of reaction time and response accuracy on the number of categories and the number of objects.** Mean values are presented with 95% confidence intervals.

identified in parallel. One set of stimuli had more distorted local features and the other set of stimuli had a more distorted global features. We predicted that the parallel identification of objects could be based on global features, therefore the distortion of outline shape should result in a dependency of response time on the number of objects independent of whether objects belong to the same or to different categories. The results of the experiment did not support our hypothesis. Both types of distortion had an effect only on the absolute level of performance accuracy, and to reach the same accuracy level as with intact stimuli, the exposition time for distorted stimuli was doubled. Distortion of global or local features did not change the pattern of task performance (i.e. the dependency of performance effectiveness on the number of objects and on the number of categories). At this point we can only suggest that both local and global features are used to identify the category of objects in a multiple-object environment.

It is too early to conclude that objects of the same category are identified in parallel in natural settings based on the findings of this study. Further investigations are required to test this suggestion. Here we can only speculate about the possible processes of identification of multiple objects. Our findings suggest the following scenario. As the identification of one object was faster and more accurate than identification of two objects of the same category, it is possible that the visual system first identifies one category. Additional time is required to identify other objects of this category but this could be done in parallel mode because this second stage could be regarded as "detection stage" instead of the first "identification stage". Many studies and theories state that the detection of an object's presence is a faster process than identification of the object's category (Biederman, 1987; de la Rosa *et al*., 2011; Kobylka *et al*., 2017; Marr, 1982; Nakayama *et al*., 1995; but see Green, 1992; Grill-Spector & Kanwisher, 2005, which found no difference in the performance time between identification and detection tasks). In our case, the visual system should detect whether an object such as "shoe" (object of the first identified category) is present or not, and all shoes are detected simultaneously if they are present. Later follows the next "identification stage" when the next category is identified. It remains unclear what kind of object features are processed during the detection and identification of objects.

A somewhat similar two-stage processing model of several simultaneously presented pictures was suggested by Potter & Fox (2009). They presented up to four photographs simultaneously in a rapid serial visual presentation procedure and subjects had to either detect a verbally denoted target or memorize pictures and perform a recognition test after each sequence of pictures. The two stages for visual processing suggested by the authors were: fast global processing of all pictures in a stimulus that is sufficient for target detection, and slower serial processing that is required for object recognition.

In summary, the presented experimental data support the hypothesis that visual objects of the same category are identified in parallel in multiple object stimuli. As the distortion of global or local features do not influence the performance pattern, we can suggest that both global and local features are processed during identification of object category. The number of simultaneously presented objects was restricted to three items in our experiment, therefore the further research is needed with higher number of objects.

## Data and software availability

**Dataset 1. Response time and accuracy data.** "1-1", "1-2", "2-2s", "2-2d", "1-3", "2-3s", "2-3d", "3-3" are different types of stimuli where the first digit between quotation marks indicates the number of categories and the second digit indicates the number of objects. Original, original conditions; locally d., locally distorted conditions; globally d., globally distorted conditions; RT, reaction time; %, percentage accuracy. DOI: 10.5256/f1000research.14468.d204177 (Šoliūnas, 2018).

Free alternatives for E-prime software: PsychoPy (Peirce, 2007); DXMD (Forster & Forster, 2003).


## Competing interests

No competing interests were disclosed.

## Grant information

The author declared that no grants were involved in supporting this work.


## Supplementary material

**Supplementary File 1. Stimuli used in the experiment.**

Click here to access the data.

## References


Alais D, Blake R, eds.: **Binocular Rivalry.** MIT Press: Cambridge. 2004.
Reference Source

Auckland ME, Cave KR, Donnelly N: **Nontarget objects can influence perceptual processes during object recognition.** *Psychon Bull Rev.* 2007; **14**(2): 332–337.
PubMed Abstract | Publisher Full Text

Biederman I: **Recognition-by-components: A theory of human image understanding.** *Psychol Rev.* 1987; **94**(2): 115–147.
PubMed Abstract | Publisher Full Text

Brandman T, Peelen MV: **Interaction between Scene and Object Processing Revealed by Human fMRI and MEG Decoding.** *J Neurosci.* 2017; **37**(32):

7700–7710.
PubMed Abstract | Publisher Full Text

Davenport JL: **Consistency effects between objects in scenes.** *Mem Cognit.* 2007; **35**(3): 393–401.
PubMed Abstract | Publisher Full Text

Davenport JL, Potter MC: **Scene Consistency in Object and Background Perception.** *Psychol Sci.* 2004; **15**(8): 559–564.
PubMed Abstract | Publisher Full Text

de la Rosa S, Choudhery RN, Chatziastros A: **Visual object detection, categorization, and identification tasks are associated with different time courses and sensitivities.** *J Exp Psychol Hum Percept Perform.* 2011; **37**(1): 38–47.
PubMed Abstract | Publisher Full Text

Enns JT, Rensink RA: **Preattentive recovery of three-dimensional orientation from line drawings.** *Psychol Rev.* 1991; **98**(3): 335–351.
PubMed Abstract | Publisher Full Text

Forster KI, Forster JC: **DMDX: A Windows display program with millisecond accuracy.** *Behav Res Methods Instrum Comput.* 2003; **35**(1): 116–124.
PubMed Abstract | Publisher Full Text

Gagne CR, MacEvoy SP: **Do simultaneously viewed objects influence scene recognition individually or as groups? Two perceptual studies.** *PLoS One.* 2014; **9**(8): e102819.
PubMed Abstract | Publisher Full Text | Free Full Text

Gilchrist ID, Humphreys GW, Neumann H, *et al.*: **Luminance and edge information in grouping: A study using visual search.** *J Exp Psychol Hum Percept Perform.* 1997; **23**(2): 464–480.
PubMed Abstract | Publisher Full Text

Green M: **Visual search: detection, identification, and localization.** *Perception.* 1992; **21**(6): 765–77.
PubMed Abstract | Publisher Full Text

Green C, Hummel JE: **Familiar interacting object pairs are perceptually grouped.** *J Exp Psychol Hum Percept Perform.* 2006; **32**(5): 1107–1119.
PubMed Abstract | Publisher Full Text

Grill-Spector K, Kanwisher N: **Visual recognition: as soon as you know it is there, you know what it is.** *Psychol Sci.* 2005; **16**(2): 152–60.
PubMed Abstract | Publisher Full Text

Gronau N, Neta M, Bar M: **Integrated contextual representation for objects' identities and their locations.** *J Cogn Neurosci.* 2008; **20**(3): 371–388.
PubMed Abstract | Publisher Full Text

Hollingworth A, Henderson JM: **Object identification is isolated from scene semantic constraint: evidence from object type and token discrimination.** *Acta Psychol (Amst).* 1999; **102**(2–3): 319–343.
PubMed Abstract | Publisher Full Text

Joubert OR, Rousselet GA, Fize D, *et al.*: **Processing scene context: fast categorization and object interference.** *Vision Res.* 2007; **47**(26): 3286–97.
PubMed Abstract | Publisher Full Text

Joubert OR, Fize D, Rousselet GA, *et al.*: **Early interference of context congruence on object processing in rapid visual categorization of natural scenes.** *J Vis.* 2008; **8**(13): 11.1–18.
PubMed Abstract | Publisher Full Text

Kimball S, Mattis P, GIMP Development Team: **GNU Imagine Manipulation Program.** 2013.
Reference Source

Kobylka F, Persike M, Meinhardt G: **Object Localization Does Not Imply Awareness of Object Category at the Break of Continuous Flash Suppression.** *Front Hum Neurosci.* 2017; **11**: 312.
PubMed Abstract | Publisher Full Text | Free Full Text

Leopold DA, Logothesis NK: **Multistable phenomena: changing views in perception.** *Trends Cogn Sci.* 1999; **3**(7): 254–264.
PubMed Abstract | Publisher Full Text

Mack ML, Palmeri TJ: **Modeling categorization of scenes containing consistent versus inconsistent objects.** *J Vis.* 2010; **10**(3): 11.1–11.
PubMed Abstract | Publisher Full Text

Marr D: **Vision: A Computational Investigation into the Human Representation and Processing of Visual Information.** W H Freeman: New York. 1982.
Reference Source

Merigan WH, Maunsell JH: **How parallel are the primate visual pathways?** *Annu Rev Neurosci.* 1993; **16**: 369–402.
PubMed Abstract | Publisher Full Text

Nakayama K, He ZI, Shimojo S: **Visual surface representation: A critical link between lower-level and higher-level vision.** In SM Kosslyn & DN Osherson (Eds.). *An invitation to cognitive science: Visual cognition*, MIT Press: Cambridge. 1995; 1–70.
Reference Source

Peirce JW: **PsychoPy--Psychophysics software in Python.** *J Neurosci Methods.* 2007; **162**(1–2): 8–13.
PubMed Abstract | Publisher Full Text | Free Full Text

Potter MC, Fox LF: **Detecting and remembering simultaneous pictures in a rapid serial visual presentation.** *J Exp Psychol Hum Percept Perform.* 2009; **35**(1): 28–38.
PubMed Abstract | Publisher Full Text | Free Full Text

Psychology Software Tools, Inc: **[E-Prime 2.0].** 2012.
Reference Source

Rousselet GA, Fabre-Thorpe M, Thorpe SJ: **Parallel processing in high-level categorization of natural images.** *Nat Neurosci.* 2002; **5**(7): 629–30.
PubMed Abstract | Publisher Full Text

Rousselet GA, Thorpe SJ, Fabre-Thorpe M: **Processing of one, two or four natural scenes in humans: the limits of parallelism.** *Vision Res.* 2004a; **44**(9): 877–94.
PubMed Abstract | Publisher Full Text

Rousselet GA, Thorpe SJ, Fabre-Thorpe M: **How parallel is visual processing in the ventral pathway?** *Trends Cogn Sci.* 2004b; **8**(8): 363–70.
PubMed Abstract | Publisher Full Text

Šoliūnas A: **Dataset 1 in: Influence of global and local features on parallel object identification.** *F1000Research.* 2018.
Data Source

Soliunas A, Pleskaciauskas A, Daktariunas A: **Multiple object categorization and effect of spatial frequencies.** *PeerJ Preprints.* 2018; **6**: e26666v1.
Publisher Full Text

StatSoft, Inc., Tulsa, USA: **[Statistica 7.0].** 2004.
Reference Source

Treisman AM, Gelade G: **A feature-integration theory of attention.** *Cogn Psychol.* 1980; **12**(1): 97–136.
PubMed Abstract | Publisher Full Text

Wolfe JM: **Guided search 2.0 A revised model of visual search.** *Psychon Bull Rev.* 1994; **1**(2): 202–238.
PubMed Abstract | Publisher Full Text

# Open Peer Review

## Current Referee Status: ? ? ? ?

**Version 1**

Referee Report 28 June 2018

https://doi.org/10.5256/f1000research.15749.r35056

? **Mohd Izzuddin Hairol**
Optometry & Vision Sciences Programme, Faculty of Health Sciences, National University of Malaysia (UKM), Kuala Lumpur, Malaysia

The article investigated whether identification performance of objects are performed serially or in parallel. The author shows that reaction time and accuracy of object identification are affected by the number of categories of objects, but not the number of objects. There are a few issues that the author need to address.

Parts of the Introduction seem incompletely discussed/elaborated. For example, in para 3, the elaboration of the previous research would be better if they were related back clearly to the purpose of the current study.

Insufficient details in Methods

What do the authors mean by corrected-to-normal vision? How was it assessed?

What was the size of each pixel? Since the stimulus size was restricted to 100x100 pixels, the angular size of the objects, particularly if they had details that might be important for recognition, might have influence response accuracy.

Among the 10 categories of objects that were chosen, were they tested to be equally identifiable? Some objects might be harder to identify than others, therefore could affect reaction time and accuracy.

The author mentioned that “the longer stimulus exposition duration under the two conditions with distorted stimuli were chosen to equalize response accuracy under all conditions” (page 3). Well if response accuracy had been equalize then of course one would not find any difference in duration time between original, locally-distorted and globally distorted objects?

A repeated-measures ANOVA would be more suitable to analyse the data.

There are several English language issues that the author need to address, as have been pointed out by other reviewers. A glaring example is the first sentence in the article: “Visual information is processed in the brain in both parallel and serial.” Parallel and serial WHAT?

The implications of the study need to be better discussed.

The current study is undoubtly of interest but all of these concerns need to be addressed.

**Is the work clearly and accurately presented and does it cite the current literature?**
Partly

**Is the study design appropriate and is the work technically sound?**
Partly

**Are sufficient details of methods and analysis provided to allow replication by others?**
Yes

**If applicable, is the statistical analysis and its interpretation appropriate?**
No

**Are all the source data underlying the results available to ensure full reproducibility?**
Yes

**Are the conclusions drawn adequately supported by the results?**
Partly

***Competing Interests:*** No competing interests were disclosed.

***Referee Expertise:*** visual perception, visual psychophysics

**I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.**

Referee Report 25 June 2018

https://doi.org/10.5256/f1000research.15749.r35055

? **Denis V. Yavna**
Department of Psychophysiology and Clinical Psychology, Academy of Psychology and Educational Sciences, Southern Federal University, Rostov-on-Don, Russian Federation

This paper is aimed to answer whether the identification of objects and its category in an multi-object stimuli are performed serially or in parallel. It is shown that the number of categories of simultaneously presented objects, but not the number of objects per se, affects the time and accuracy of the identification of a category. The author demonstrates that if several objects belong to a smaller number of categories, then the categories are easier to identify than when several objects belong to a larger number of categories. It is assumed that objects of one category are identified in parallel mode.

The main problem of this study is that it is much more possible to guess the right answer when there are more than one image of one category that are presented simultaneously comparing to stimuli that consist of images of different categories. Suppose there are 3 objects belonging to the same category. If an

observer identifies at least one object, they are more likely to give the correct answer than in a situation where there are two categories: a priori, they have less chance to detect an object of one of these categories. It seems that accounting for a priori probability is necessary. The simplest thing that comes to mind is subtraction of the total probability of random guessing from an empirical estimate of the probability for each type of stimulus (1-1, 1-2, etc).

I agree with Dan McCarthy concerning the issue of the statistical analysis used by the author. Probably, for a more correct (and powerful in a statistical sense) comparison of test results performed through repeated testing of observers in similar tasks, it is necessary to apply a method that gives special attention to differences in performance in different tasks and not to the *individual differences* in speed and accuracy of identification; e. g. repeated measures ANOVA is more preferable. And of course, in the presence of conclusions based on the adoption of the null hypothesis, power analysis is desired. Probably this[1] work containing tables with power estimations for the 3x6 design, will be useful for author.

I also would like to note that the use of the chi-square test to check the normality of distribution is not recommended by many authors (e. g. [2] p. 94]); but this note is not fundamental.

Undoubtedly, the work is of great interest and should be indexed, but a significant revision is required.

**References**

1. Potvin P, Schutz R: Statistical power for the two-factor repeated measures ANOVA. *Behavior Research Methods, Instruments, & Computers*. 2000; **32** (2): 347-356 Publisher Full Text
2. Zar JH: Biostatistical analysis (Fifth edition). *Upper Saddle River: Pearson Prentice Hall*. 2010.

**Is the work clearly and accurately presented and does it cite the current literature?**
Partly

**Is the study design appropriate and is the work technically sound?**
Partly

**Are sufficient details of methods and analysis provided to allow replication by others?**
Yes

**If applicable, is the statistical analysis and its interpretation appropriate?**
No

**Are all the source data underlying the results available to ensure full reproducibility?**
Yes

**Are the conclusions drawn adequately supported by the results?**
Partly

***Competing Interests:*** No competing interests were disclosed.

**I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.**

Referee Report 15 June 2018

https://doi.org/10.5256/f1000research.15749.r35054

? **Thiago Monteiro de Paiva Fernandes**
Department of Psychology, Perception, Neuroscience and Behavior Laboratory, Federal University of Paraíba, João Pessoa, Brazil

This article reports the influence of global and local features on parallel object identification using one, two or three visual objects. Author's efforts to conduct a proper study needs to be mentioned/recognised. The author provided a clear and testable hypothesis and the article is straightforward.

**1) The introduction is somewhat unfocused. The author rapidly switch from discussing theoretical and physiological aspects of visual perception, but do not strongly compare their assumptions with alternative theories. Although it is already brief, I reckon the Introduction could be made more concise and should be edited to improve flow.**

- There are a considerable number of obvious and/or dubious statements (e.g., Visual information is processed in the brain in both parallel and serial *what?*; Features such as color, orientation, motion, spatial frequency and stereodepth are detected in parallel). Although it is true that we can not ignore the complexity of retinal information, seems like the author ignored the boundary completion. When one take part in some assumption (e.g., parallel; apologise if I got it wrong), it is fundamental that both sides (parallel vs serial) could be briefly discussed.

- The author reported some *old* references to support assumptions. Some of them appear to have been poorly chosen/incorrect. Maybe the author could update them. I reckon this will strengthen the arguments.

- Scene context discussion (one or two sentences) could be interesting to improve Introduction.

**2) Overall, some parts of the article is difficult to understand due to the insufficient English language quality (examples are listed below) and many typographical errors throughout the text. I suggest that the authors enlist an editor/colleague to improve the English.**

English quality examples:

"Visual information is processed in the brain in both parallel and serial."
- What? Processing? Yes, I understood, but seems like we need a complement.

"Much less research"
- What about "Research on interaction between objects... are underreported..." ?

"For better control of the global features of objects, new categories were selected and new objects were produced"

- It is not clear what the authors' intent is in this paragraph. Further, it is not clear if the authors are making a prediction if producing new objects would affect parallel identification.

**3) The absence of relevant information in the Methods section is worrisome. There is a lack of detail about study procedures.**

- The subjects were assessed for neuropsychiatric disorders?

- The subjects were assessed for substance use?

- "normal or corrected-to-normal"; Do you mean 20/20? Using eye chart? Freiburg visual acuity?

- Don't you think that the sample characteristics biased the results? (e.g., only young subjects; significant gender differences; menstrual cycle influence; cognitive performance). If you disagree, please update Subjects subsection.

- "All *of the* objects were transformed into grayscale pictures and resized in such a way ... 100x100-pixel"
I see your point. However, the readers also need to understand your point. Why grayscale? Why 100 x 100?

- Luminance of the monitor screen?

- Why 60 cm? This was based on previous studies?

- Why you chose 106-ms blank interval? Seems like an odd number if you do not explain.

Further, in my opinion some of methods could be reorganised into more relevant sub-sections.

**4) Statistical analysis needs to be reformulated. I reckon the Bayes Factor would be an interesting approach. Nevertheless, I recognise this may take a long time to update. My main concern is about the ANOVA. I see your point for merging results (2-2s and 2-2). However, it seems worrisome. I suggest to you performing the analysis without merge the results (multivariate analysis).**

- Why the use of chi-square (fit test) for normal distribution?

- Why the Newman-Keuls post-hoc (and not the REGWQ)?

- Effect size? Confidence Interval? Please provide such additional information.

I agree with the other reviewer about the possibility of Type I and/or Type II error. Thus, you will need to reorganise this section.

**5) I am unsure what implications are to be drawn from the study at present. I mean, you need to better address this on Discussion.**

- Although the author can not drawn physiological conclusions, it should be - at least - mentioned.

- You supported the hypothesis "that visual objects of the same category are identified in parallel..." I see your point. However, this needs to be better addressed. The Discussion should have more references.

I am providing some references to help in the organisation.

**References**

1. Munneke J, Brentari V, Peelen MV: The influence of scene context on object recognition is independent of attentional focus.*Front Psychol*. 2013; **4**: 552 PubMed Abstract | Publisher Full Text
2. Sastyin G, Niimi R, Yokosawa K: Does object view influence the scene consistency effect?. *Atten Percept Psychophys*. 2015; **77** (3): 856-66 PubMed Abstract | Publisher Full Text
3. Beauchamp MS, Lee KE, Haxby JV, Martin A: Parallel visual motion processing streams for manipulable objects and human movements.*Neuron*. 2002; **34** (1): 149-59 PubMed Abstract
4. Meyer AS, Ouellet M, Häcker C: Parallel processing of objects in a naming task.*J Exp Psychol Learn Mem Cogn*. 2008; **34** (4): 982-7 PubMed Abstract | Publisher Full Text
5. Bullier J, Nowak LG: Parallel versus serial processing: new vistas on the distributed organization of the visual system.*Curr Opin Neurobiol*. 1995; **5** (4): 497-503 PubMed Abstract
6. Sasaki Y: Processing local signals into global patterns.*Curr Opin Neurobiol*. 2007; **17** (2): 132-9 PubMed Abstract | Publisher Full Text
7. Rauss K, Pourtois G, Vuilleumier P, Schwartz S: Effects of attentional load on early visual processing depend on stimulus timing.*Hum Brain Mapp*. 2012; **33** (1): 63-74 PubMed Abstract | Publisher Full Text

**Is the work clearly and accurately presented and does it cite the current literature?**
Yes

**Is the study design appropriate and is the work technically sound?**
Yes

**Are sufficient details of methods and analysis provided to allow replication by others?**
Partly

**If applicable, is the statistical analysis and its interpretation appropriate?**
No

**Are all the source data underlying the results available to ensure full reproducibility?**
Yes

**Are the conclusions drawn adequately supported by the results?**
Yes

***Competing Interests:*** No competing interests were disclosed.

***Referee Expertise:*** Visual perception; Cognition; Psychophysics

**I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.**

Referee Report 05 June 2018

https://doi.org/10.5256/f1000research.15749.r34363

? **Dan McCarthy**
Brown University, Providence, RI, USA

The current study investigated whether multiple objects are processed serially or in parallel by testing the impact of global and local feature distortions on identification performance. Specifically, participants identified whether an object belonging to a particular category was present in a display consisting of 1-3 objects. Local distortion was applied by masking images with stripes masking the background color of the display, thereby occluding local object features while preserving the global shape outline. Global distortion was accomplished by warping objects in an image editor, allowing full view of the object despite deformations in individual features and the global object outline.

Results indicate that such distortion led to slowed responses and decreased accuracy in identifying an object category, but this is attributed on differences in the exposure duration (~2x longer viewing time for distorted stimuli to achieve similar performance to originals). Increasing the number of categories present in a display was associated with lower accuracy and slower response times when 2 or more objects were shown. No reaction time differences were observed in the key contrast between the 2 or 3 object displays, yet accuracy was higher with 2 vs. 3 objects in the 2 category case.

Based on these findings, the author argues that multiple objects can be identified simultaneously, but only when they belong to the same category. The lack of performance differences between locally and globally distorted stimuli was interpreted as evidence that object identification relies on both local and global features.

While it is written clearly and the question is interesting as other work indeed suggests the plausibility of multiple object identification in parallel, I have statistical and methodological concerns related to the interpretation of these results that I list below:

1. A two-way ANOVA is not the appropriate statistical test given my understanding that all participants completed the same task. A 2-way repeated measures ANOVA should be used instead. This may explain the extremely large F-values and within-groups degrees of freedom (i.e., DFwithin = 27822 for all main effects and the interaction). I am unsure how this value was obtained. The correct DFwithin should be 285 for stimulus type ((6 groups - 1)*(58 participants - 1) = 5*57 = 285), 114 for distortion ((3 groups - 1)*(58 participants - 1) = 2*57 = 114), and 570 for the interaction (5*2*57 = 570). This test should be redone correctly and accurately reported to confirm the reported results hold.

2. Another concern is that the primary conclusion for parallel object identification within a single category is based on a failure to reject the null. No power analysis is reported to indicate the chance of a Type II error. A powerful approach to address evidence in favor of the null would be to use Bayes Factors (for examples, see Wagenmakers, 2007[1]; Jarosz & Wiley, 2014[2]). This can be easily implemented in JASP (Love et al., 2015[3]) - freely available, user friendly software that the author could use to conduct a Bayesian version of the two-way repeated measures test discussed above that would allow a direct test of the likelihood that the null hypothesis (no RT difference between identifying 2 or 3 objects of the same category) is a better model of the data than the alternative (a difference). Jarosz & Wiley (2014) is an excellent resource for how to interpret the results of such a test and what strength of evidence favoring either hypothesis may be.

3. The use of occluding bars for local distortion is somewhat odd given that the same 'whirl and pinch' effects could be applied to the interior of objects without distorting the global boundary. A brief description of why these disparate methods were chosen or examples in previous literature would be helpful for the reader.

I look forward to seeing a revised version of this work once these concerns have been properly

addressed.

**References**

1. Wagenmakers E: A practical solution to the pervasive problems ofp values. *Psychonomic Bulletin & Review*. 2007; **14** (5): 779-804 Publisher Full Text
2. Jarosz A, Wiley J: What Are the Odds? A Practical Guide to Computing and Reporting Bayes Factors. *The Journal of Problem Solving*. 2014; **7** (1). Publisher Full Text
3. Love J, Selker R, Marsman M, Jamil T, VerhagenAJ, Ly A, *et al.*: JASP (Version 0.6.6). *[Computer software].*2015.

**Is the work clearly and accurately presented and does it cite the current literature?**
Partly

**Is the study design appropriate and is the work technically sound?**
Partly

**Are sufficient details of methods and analysis provided to allow replication by others?**
Yes

**If applicable, is the statistical analysis and its interpretation appropriate?**
No

**Are all the source data underlying the results available to ensure full reproducibility?**
Yes

**Are the conclusions drawn adequately supported by the results?**
Partly

***Competing Interests:*** No competing interests were disclosed.

***Referee Expertise:*** Form perception, motion perception, attention, visually-guided action, synesthesia

**I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.**