

Familiar Size Effects on Reaction Time: When Congruent is Better

**Authors:**

Carmen Fisher, Irene Sperandio

**Affiliations of Authors:**

School of Psychology, University of East Anglia, Norwich Research Park, Norwich, NR4  
7TJ, United Kingdom.

**Corresponding Author:**

Correspondence should be addressed to Irene Sperandio at the above address.

E-mail: [i.sperandio@uea.ac.uk](mailto:i.sperandio@uea.ac.uk)

**Word count:**

**8,828**

### **Abstract**

Familiar size is known to influence our perception of object's size and distance. In this study, we examined whether or not simple reaction times (RTs) are also affected by prior knowledge of objects' size. In a series of experiments, participants were asked to respond as quickly as possible to briefly presented images of familiar objects, equated for luminance and retinal size. The effects of familiar size and object animacy on RTs were investigated under natural (Experiment 1) and reduced (Experiment 2) viewing conditions. Restricted viewing conditions were introduced to manipulate the availability of depth cues. A systematic effect of familiar size on RTs was considered for progressively 'shrunk' (Experiment 3) and 'enlarged' (Experiment 4) objects on the screen with respect to their familiar size. Measures of perceived size were also taken by means of a manual estimation task (Experiment 5). Results showed an effect of animacy on simple RTs: participants were faster to respond to images of animals than non-animals. An effect of familiar size on simple RTs was also observed under reduced viewing conditions only: objects shown closer to their real-world size were detected significantly more quickly than those further from their familiar size. However, this familiar-size advantage did not reflect perceived size. Hence, simple RTs under reduced viewing conditions are modulated by the degree of compatibility between physical size and long-term representations of size.

Keywords; Real-world size, Perceived size, Animacy, Simple reaction times, Restricted view

### **Public significance**

In this study, we demonstrate that knowledge about object's size can influence reaction times such that we are faster to detect briefly presented images of objects that match their familiar size. For example, a key shown at its familiar size is responded to more quickly than a bus shown at the same size. We also observe that animacy has an effect on our reaction times, such that we are faster to detect images of animals than non-animals. We argue that both familiar size and animacy are automatically processed by the brain in early stages of visual processing and might involve specialised neural networks.

## Familiar Size Effects on Reaction Time: When Congruent is Better

**Glossary**

**Familiar size** = Stored knowledge of the size of an object.

**Actual real-world size** = Physical size of the object in the external world.

**Perceived size** = Our perceptual experience of an object's size

**Retinal size** = Size of an image produced on the retina, measured in degrees of visual angle.

**Physical size** = The actual size of an object as presented on the screen

**Normative size** = The preferred size for object representation

Familiar objects are often experienced at different distances and visual angles, producing different representations of the same object on the retina. These variations in distance and angle could be seen as problematic in the process of establishing object recognition. However, it is commonly understood that despite changes in viewing conditions, the observer is still able to identify certain properties of the object, such as its familiar size. Familiar size can be defined as the previously stored knowledge about the size of an object that constructs an understanding of the object's distance. This past experience provides important information about object's size which can help the observer in making size judgments (e.g. Slack, 1956; Schiffman, 1967). For example, previous research has demonstrated that participants were just as accurate at judging the size of a familiar object from direct view, as recalling its size from memory (Bolles & Bailey, 1956).

More recently, Konkle and Oliva (2007; 2011) have used a range of methods, including drawing, ranking and size adjustment tasks, to assess the effects of previous knowledge of an object's size on size perception. These studies led to the same conclusion: our perception and memory for objects reflect a normative size that is strongly related to the real-world size, also referred to as the 'canonical visual size' (Konkle & Oliva, 2007; Konkle & Oliva, 2011). In other words, the authors demonstrated that through experience we retain information about the typical size of objects and this influences how we prefer and expect to perceive their size. Therefore, even though the sensory information about an object changes

due to variations in viewing conditions, our perception of object's size is influenced by internal representations of its familiar size.

Interestingly, the same researchers implemented a Stroop-like task in another study to show that familiar size can also affect the speed of processing of real-world objects (Konkle & Oliva, 2012a). Their task involved the presentation of image pairs that consisted of various familiar objects displayed next to each other on a screen. One image was always presented at a larger physical (actual) size compared to the other image. The images corresponded to either big or small objects in the real world. Participants were asked to indicate which image had the biggest or smallest physical size from the pairing, a task that was unrelated to the knowledge of familiar size. It was observed that congruency in real-world size of the two images had an influence on the speed of response, such that there was an advantage in RT for those images whose physical size was congruent with the real-world size. For example, when presenting a big elephant and a small snail (congruent trial) responses were faster than when presenting a big snail and a small elephant (incongruent trial). This finding shows that familiar size is an automatic property of object representation that can affect the speed of processing of real-world objects. In keeping with these findings, other studies have demonstrated that some properties of objects are processed in an automatic fashion by the observer and as such can influence the early visual processing of objects. For example, Grill-Spector & Kanwisher (2005) reported that participants were able to identify (i.e. sail boat vs. ship) and categorise (i.e. boat vs. house) objects with the same level of speed and accuracy as for detection alone. Hence, object identity and categorisation seem to take place early on in the processing of visual information.

The influence of perceived size on speed of processing has been explored by means of a simple reaction time (RT) approach by Sperandio and colleagues (Sperandio, Savazzi, Gregory, & Marzi, 2009). In this study, participants were asked to react as fast as possible to

briefly presented visual stimuli of different physical size but subtending the same retinal angle as a result of their different viewing distance. Stimuli consisted of plain dots or pictures of tennis balls matched in luminance. It was found that simple RTs were significantly affected by perceived size only when familiar object stimuli (i.e. tennis ball) were presented. Therefore, simple RTs reflect perceived size even when the retinal angle subtended by the stimuli is constant: participants responded faster to stimuli perceived as larger (i.e. big tennis ball at far distance) than stimuli perceived as smaller (i.e. small tennis ball at near distance). While the effect of *retinal* size on simple RT, whereby bigger stimuli on the retina produce faster responses, is well-established in the literature (e.g. Osaka, 1976), Sperandio et al's (2009) work suggested for the first time the existence of a relationship between *perceived* size and simple RTs. Importantly, the effect of perceived size on RTs was enhanced by the observer's familiarity with the objects. The finding that RTs are a measure of perceived size was further supported by other studies using visual illusions, in which illusory larger stimuli produced faster RTs than illusory smaller stimuli, although their retinal image was always the same (Sperandio, Savazzi, & Marzi, 2010; Plewan, Weidner, & Fink, 2012; Savazzi, Emanuele, Scalf, & Beck, 2012). Taken together, these previous studies demonstrate that speed of processing can be affected by perceived size and prior knowledge about the size of objects in the world. However, it is still unclear how as well as the extent to which familiar size affects simple RTs.

Here, we measured simple RTs in response to images of real-world objects with different familiar size (e.g. small or big). The images were briefly presented on the screen and subtended the same visual angle and luminance. Participants were asked to react to the images as soon as they saw them appear on the screen (Experiments 1-4). In the first two experiments, two cognitive dimensions of the stimuli were manipulated: familiar size (big vs small) and animacy (animate vs inanimate). Except for Experiment 1, where natural

conditions of viewing were used, all the other experiments were carried out under reduced viewing conditions to manipulate the number of depth cues available (Holway & Boring, 1941) and enhance the effect of familiar size on perception (Epstein, 1963; Gogel, 1969; Ittelson, 1951). To examine if simple RTs are systematically affected by familiar size, off-sized versions of familiar objects (i.e. physically shrunken or enlarged on the screen with respect to their known size) with an increasing level of incongruence between their familiar and physical size were presented in Experiments 3 and 4. Finally, a manual size estimation task was used in Experiment 5 to measure the perceived size of familiar objects to establish if any effects of familiar size on RTs are due to perceived size, as previously reported (e.g. Sperandio et al., 2009). Given the evidence for an influence of both familiar and perceived size on speed of processing, we expected to find an effect of familiar size on simple RTs (Experiments 1-4) and that this effect would reflect how big an object appeared to the observer (Experiment 5).

### **Experiment 1: The influence of Familiar Size and Animacy on simple RTs**

In the first experiment, images of familiar objects were presented using a simple RT paradigm. Participants were asked to press a designated button as soon as they detected an image appearing on the screen. Stimuli were randomly presented from a selection of images chosen for their familiar size (small or big) and animacy (animate or inanimate).

Previous research has demonstrated that familiar size can speed up participants' performance in judging the physical size of an object compared to its pair but only when the object's familiar size pairing was congruent with the physical size presented on a screen (Konkle & Oliva, 2012a). To put it simply, participants were faster to respond to a small

apple (11° of visual angle) presented with a big piano (18°) (congruent pairing), compared to a big rubber duck (18°) presented with a small couch (11°) (incongruent pairing).

The aim of Experiment 1 was to verify if the same effect can be generalized to simple RTs. Simple RTs can be considered as one of the most basic measures of speed of processing, whereby participants are reporting their conscious perception of the onset of a target by means of stereotyped (basic) speeded responses (Johnson et al., 1985).

Therefore, if familiar size is an automatic property of object representation (Konkle & Oliva, 2012a), then one would expect simple RTs to be faster in response to those stimuli that are physically closer to their familiar size. For example, participants should be faster to detect a picture of a mouse than a picture of an elephant, when both image sizes subtend the same visual angle.

Another cognitive dimension investigated in the present experiment was animacy. Although a large body of evidence has supported the idea of different neural mechanisms dedicated to animate and inanimate object categories (e.g. Caramazza & Shelton, 1998; Cichy, Pantazis, & Oliva, 2014; Konkle & Caramazza, 2013; Mahon, Anzellotti, Schwarzbach, Zampini, & Caramazza, 2009; McMullen & Purdy, 2006, for reviews see Gerlach, 2007; Martin, 2007), much less consensus exists as to whether or not there is an advantage for animate/living objects over inanimate/non-living objects. On the one hand, there is evidence that images of animals are detected more quickly by the observers (e.g. Li, VanRullen, Koch, & Perona, 2002; New, Cosmides, & Tooby, 2007). On the other hand, no difference across the two categories has also been reported (Praß, Grimsen, König, & Fahle, 2013; VanRullen & Thorpe, 2001). Here, we examined the effect of animacy on simple RTs. One should note that although simple RTs do not require object categorization or identification, it has been reported that detection and categorization occur simultaneously,

namely as soon as the observers detect an object, they already know its category (Grill-Spector & Kanwisher, 2005; Mack, Gauthier, Sadr, & Palmeri, 2008). Therefore, one might expect to find an effect of animacy on simple RTs.

Animacy is frequently examined in association with familiar size (e.g. Konkle & Oliva, 2007, 2011, 2012; Gabay, Leibovich, Henik, & Gronau, 2013). Interestingly, there is fMRI evidence to indicate a tripartite organisation of neural activity for object representation such that objects are functionally organized into three cortical zones that preferentially respond to: i) large objects; ii) small objects; iii) animals (Konkle & Caramazza, 2013). This finding suggests that the representation of animate stimuli should be independent of familiar size. Therefore, it is conceivable that we may find an effect of animacy and familiar size on RTs but not an interaction between these two factors.

The images presented in this experiment were controlled to have the same overall luminance and aspect ratio. Controlling for these factors was deemed necessary as it is well known that luminance and physical size affect RT, such that bigger and brighter stimuli on the retina typically produce faster responses (e.g. Osaka, 1976; Pins & Bonnet, 1996).

## Method

### Participants

Twenty-four participants (2 males), ranging in age from 18 to 44 years ( $M=21$   $SD = 5.52$ ) took part in the experiment. The sample size for this and all following experiments was deemed to be appropriate to attain a moderate effect size with  $\alpha = .05$  and power = .80, according to calculations performed in G\*Power (Faul, Erdfelder, Lang, & Buchner, 2007). Participants were all right-handed. In this and all following experiments, participants had normal or corrected-to-normal vision, gave informed consent prior to testing and received



course credits or payment for their time. All methods were compliant with the rules and regulations of the Psychology Ethics committee of the University of East Anglia.

### **Apparatus**

Participants sat in a dimly lit room with their head on a chin rest placed 57 cm away from a PC monitor. Visual stimuli were presented on a DELL screen (17 inches) with a screen resolution of 1280 x 1024. The stimuli and the psychophysical experiments were programmed in E-Prime version 2.0 software (Psychology Software Tools, Pittsburgh, USA). A Konica Minolta LS-100 luminance photometer was used to measure luminance of the screen and stimuli. Stimuli were displayed on a grey background with a luminance of 128 cd/m<sup>2</sup>. A black fixation cross (font Arial, size 16 pt) was presented at the centre of the screen. Size and luminance of the stimuli were adjusted using GNU image manipulation program, version 2.8.6 (available at: [www.GIMP.com](http://www.GIMP.com)).

### **Stimuli**

Stimuli consisted of colour images of real objects matched in luminance and approximate aspect ratio. To this end, the average luminance of each image was adjusted to correspond to 40 cd/m<sup>2</sup> and the size of each image was scaled to fit inside a 6x6 cm frame, producing a visual angle of 6° when viewed at the 57 cm distance. Images were selected according to their familiar size, using a range of sizes similar to that employed by Konkle and Caramazza (2013). Those images classified as 'small' had a familiar size that ranged between 0.8 cm and 50 cm ( $M= 10.52$  cm,  $SD=10.47$ ), while those classified as 'big' had a familiar size that ranged between 76 cm and 30,000 cm ( $M=1405.43$  cm,  $SD= 4642.84$ ). Depending on the image, the maximum size could have been in height, width or length. For example, an elephant which is biggest by its length, measuring from its head to its bottom around 500 cm on average, would be placed in the 'big' category (note that information about actual size was

collected from various internet sources). Half of the images were animate objects and the other half were inanimate (Figure 1a). The 'animate' condition included pictures of animals, while the 'inanimate' included non-living objects (Konkle & Caramazza, 2013). Images were compiled from different sources, including the Normative BOSS collections V1&V2 (Brodeur, Dionne-Dostie, Montreuil, & Lepage, 2010), Animacy x Size database (Konkle & Caramazza, 2013), Big and Small database (Konkle & Oliva, 2012b), POPORO database (Kovalenko, Chaumon, & Busch, 2012), Unique objects database (Brady, Konkle, Gill, Oliva, & Alvarez, 2013) as well as a variety of self-sourced images. For full details of image sources for this and following experiments, see Table 1 in Supplementary Materials. Headphones were given to participants to deliver a warning signal. A button response box was used to record responses.

### **Design and Procedure**

The layout of each trial consisted of a black cross presented on the screen for 1000 ms, followed by a 1000 Hz 'beep' sound (1000 ms). After this, a random interval ranging between 400 and 600 ms was introduced before an image of a familiar object or a blank screen ('catch trial') was presented for 80 ms. A period of 2000 ms was given to allow for a response (Figure 1b). Participants were instructed to respond as fast as possible to the onset of any stimulus image by pressing a designated button on a response box and to refrain from responding on catch trials. A practise block was included to ensure participants familiarized themselves with the task. There were 30 trials for each of the four conditions of stimulus presentation (small/big x animate/inanimate) plus 18 catch trials, presented in two blocks, yielding 276 trials in total. Participants were offered breaks at regular intervals to prevent fatigue. Stimuli were presented in a random fashion.

-----Please insert Figure 1 about here -----

## Results

In this and in the following experiments, anticipations (RTs < 140 ms) and delayed responses (RTs > 650 ms) in relation to stimulus onset, were excluded from the analyses (Sperandio et al., 2009). In this experiment, all participants had high accuracy rates, ranging from 89.44 to 99.64 % ( $M=94.81$ ,  $SD=0.03$ ).

A 2x2 repeated measures ANOVA was conducted on the RT data with Familiar size (small vs. big) and Animacy (animate vs. inanimate) as main factors. A significant main effect of Animacy was found ( $F(1, 23)=16.778$ ,  $p<.001$ ,  $\eta^2=.422$ ); participants were significantly slower to respond to inanimate ( $M=210.56$ ,  $SD=26.73$ ) compared to animate objects ( $M=206.64$ ,  $SD=25.80$ ). However, neither the main effect of Familiar size ( $F(1, 23)=0.726$ ,  $p=.403$ ,  $\eta^2=.031$ ; big objects:  $M=208.04$ ,  $SD=24.84$ ; small objects:  $M=209.16$ ,  $SD=27.76$ ) nor its interaction with Animacy ( $F(1, 23)=1.39$ ,  $p=.25$ ,  $\eta^2=.057$ ) were significant (Figure 2a). Therefore, contrary to our expectations, we did not find an effect of Familiar size on RTs. However, in line with previous findings supporting early categorisation of objects (Clarke, Taylor, Devereux, Randall, & Tyler, 2013), we observed an advantage for animate compared to inanimate objects on the speed of response. To rule out the possibility that the effect of animacy on RTs was simply related to differences in spatial frequency between animate and inanimate objects (Harel & Bentin, 2009; Viggiano, Costantini, Vannucci, & Righi, 2004), an analysis of the low-level statistical features to compare the spatial frequency spectrum between the two semantic categories was conducted, using the Natural Image Statistical Toolbox for MATLAB (Bainbridge & Oliva, 2015). As it turned out, there was no difference in terms of spatial frequency between animate and inanimate images ( $p=0.86$ ) (see Supplementary Materials, Table 5).

**Experiment 2: The influence of Familiar Size and Animacy on simple RTs, under reduced viewing conditions**

In Experiment 2, the stimuli and design remained the same as in the previous experiment. However, restricted viewing conditions were created by means of a dark room, a reduction tunnel and a monocular pinhole (Holway & Boring, 1941; Sperandio et al., 2009). As established by Holway and Boring (1941), under these viewing conditions the availability of visual cues about distance information is greatly reduced. Specifically, the monocular viewing condition removes binocular cues (e.g. vergence and retinal disparity), pinhole vision impairs the observers' ability to accommodate, and the use of a dark room combined with a reduction tunnel eliminates contextual cues and additional light sources. Reducing these depth cues results in a decrease of depth perception, forcing the visual perceptual system to rely more on retinal size information (Holway & Boring, 1941; Sperandio et al., 2009). Therefore, under such circumstances, perceptual judgments tend to reflect retinal size rather than perceived size (i.e. the product of distance information and retinal size).

Interestingly, Sperandio et al (2009) showed that RTs to stimuli of constant retinal size were governed by perceived size only when participants were presented with images of familiar objects (i.e. tennis balls) rather than unfamiliar plain shapes (i.e. circles). However, these effects were extinguished when reduced viewing conditions were implemented: RTs to the familiar object simply reflected the retinal size of the image. It should be noted, however, that in Sperandio et al.'s (2009) study only one familiar object was used and the retinal size was manipulated, making it impossible to establish any effects of familiar size on RTs. Given that several previous studies have shown that under reduced viewing conditions, perception relies more heavily on familiar size information (e.g. Gogel, 1969; Epstein, 1963; Ittelson,

1951; Schiffman, 1967; Slack, 1956), we repeated Experiment 1 under such restricted viewing conditions to verify whether or not familiar size can influence RTs.

## **Method**

### **Participants**

Twenty-six participants took part in the experiment. However, the data of two participants were removed due to low accuracy (>20% of errors) and technical difficulties. Those included in the final sample (24) ranged in age from 18 to 38 years ( $M=22$ ,  $SD= 4.93$ ). Four of the participants were left-handed, 9 were left-eye dominant and 6 of them were males.

### **Apparatus**

To generate reduced viewing conditions, participants performed the task in an otherwise dark room and viewed the stimuli through a reduction tunnel and a 1 mm pinhole with their dominant eye. The screen's background was changed to black (0.01 cd/m<sup>2</sup> of luminance) and the colour of the fixation cross was changed to white. Participants' eye dominance was assessed using the Dolman's method (also known as "hole-in-the-card" test; e.g. Cheng, Yen, Lin, Hsia, & Hsu, 2004). Participants wore pinhole glasses in which all the holes but the most centrally located one were covered with black tape. They then looked into a tube of 8 cm in diameter and 60 cm in length (i.e. 'reduction tunnel'). The tube led to the computer screen where the images were displayed. Headphones were given to participants to deliver a warning signal. A button response box was used to record responses.

### **Stimuli**

The stimuli used in this experiment were the same as in the previous experiment, however, due to changes in viewing conditions stimuli now subtended a visual angle of  $5.7^\circ$  instead of  $6^\circ$ .

### **Design and Procedure**

The experimental design was the same as in the previous experiment.

Participants were seated 60 cm away from the screen of the computer, in front of the end of the reduction tunnel. They were asked to wear a pair of pinhole glasses that had one central 1x1 mm aperture aligned with the dominant eye. The participant was also asked to wear headphones and to place their hand on the response button.

### **Results**

The participants' accuracy ranged from 88.37 to 99.64 % ( $M= 96.62$ ,  $SD=0.03$ ). A 2x2 repeated measures ANOVA was conducted on the RT data with Familiar size (small vs. big) and Animacy (animate vs inanimate) as main factors. A significant main effect of Familiar size was found ( $F(1, 23) =13.711$ ,  $p=.001$ ,  $\eta^2 = .373$ ); participants were significantly faster at responding to Small ( $M= 275.25$ ,  $SD= 51.50$ ) compared to Big ( $M=280.31$ ,  $SD= 53.14$ ) familiar objects. A significant main effect of Animacy was also found ( $F(1, 23) =20.695$ ,  $p<.001$ ,  $\eta^2 = .474$ ); as observed in Experiment 1, participants were significantly slower at responding to inanimate ( $M=281.47$ ,  $SD=51.66$ ) compared to animate objects ( $M= 274.08$ ,  $SD= 52.84$ ). However, the two-way interaction did not reach significance ( $F(1, 23) =0.805$ ,  $p=0.379$ ,  $\eta^2 = .034$ ) (Figure 2b).

Under reduced viewing conditions, we replicated the effect of animacy observed in Experiment 1; RTs in response to pictures of animals were faster than non-animals. This suggests that that advantage in processing animate stimuli is independent to the changes in

viewing conditions, as the animacy effect on RT was observed both under natural (Experiment 1) and reduced (Experiment 2) viewing conditions.

We also found an effect of size, such that there was an advantage in RTs for small objects compared to big objects, indicating that RTs are modulated by familiar size but only under restricted conditions of observation, when depth cues are removed.

-----Please insert Figure 2 about here -----

### **Is the effect of familiar size on simple RTs proportionate?**

The main aim of the following two experiments was to verify whether or not simple RTs are systematically affected by familiar size when off-sized versions of familiar objects (i.e. bigger or smaller on the screen compared to their size in the real-world) are viewed under restricted conditions of observation. In Experiment 3, familiar objects of normal-size to progressively under-size, were shown on the screen with constant aspect ratio. In Experiment 4, familiar objects of normal-size to progressively over-size were shown on the screen with constant aspect ratio. Objects were considered to be normal-sized if their physical size on the screen corresponded to a visual angle that was consistent with the typical size of the object viewed at a typical distance. For example, a candy would need to be placed at 60 cm of viewing distance to subtend  $5.7^\circ$  of visual angle, while a double-decker bus would need to be placed 122 m away from the participant's eyes to generate the same retinal image size (Figure 3). As such, it was hypothesized that detection time would be systematically modulated by the increased incongruence between familiar and physical size of the stimuli presented on the screen.

As familiar size was the primary concern of the present investigation, the effect of animacy will not be explored further in the subsequent experiments.

-----Please insert Figure 3 about here -----

### Experiment 3: From normal to under-sized stimuli

#### Method

##### Participants

Twenty-nine participants took part in the experiment. However, only the data of 28 participants were included in the analysis. One participant was removed due to technical difficulties. Those included in the final sample, ranged in age from 18 to 23 years ( $M=19$ ,  $SD=1$ ). Five of the participants were left-handed, five were left eye-dominant and seven of them were males.

##### Apparatus

The same apparatus as in Experiment 2 was used here.

##### Stimuli

Stimuli were selected according to the level of incongruence between their familiar size and physical size presented on the screen. Based on this criterion, three intervals were determined: 1 (normal-sized stimuli), 1/10 (under-sized stimuli) and 1/100 (under-sized stimuli). Objects were chosen such that they would fit within a hand (x1), be half the size of a person (1/10) and be bigger than a person (1/100); a similar criterion for stimulus selection was used by Konkle and Caramazza (2013). Those images classified as '1' had a familiar size that ranged between 5 cm and 7 cm ( $M= 6.61$  cm,  $SD= 0.92$ ), those classified as '1/10' had a familiar size that ranged between 50 cm and 70 cm ( $M= 68.09$  cm,  $SD= 10.94$ ) and those classified as '1/100' had a familiar size that was greater than or equal to 500 cm ( $M= 5912$  cm,  $SD= 8165$ ). As in the previous experiments, all images were of constant physical size ( $5.7^\circ$ ) and luminance ( $30$  cd/m<sup>2</sup>). Therefore, only '1' stimuli were the same size as in the real world (normal-sized), whereas '1/10' and '1/100' stimuli were smaller than in the real world (under-



sized). Some of the images used in this experiment were previously used in Experiments 1 and 2 (see Table 2 in Supplementary material for details about image source).

### **Design and Procedure**

Participants performed the experiment under restricted viewing conditions, as described in Experiment 2. The experimental design was similar to Experiment 1 except for the number of trials: there were 30 trials for each of the three size intervals (1 vs. 1/10 vs. 1/100), plus 14 catch trials, presented in two blocks, yielding 208 trials in total.

### **Results**

The participants' accuracy ranged from 85 to 100% ( $M=96.08$ ,  $SD=0.04$ ). As Mauchly's test of sphericity was significant ( $\chi^2(2) = 9.386$ ,  $p = .009$ ), the Greenhouse-Geisser value was reported. One-way repeated measures ANOVA was conducted on the RT data with Familiar size (1 vs. 1/10 vs. 1/100) as the main factor. The ANOVA showed a main effect of Familiar size on RTs ( $F(1.535, 41.443) = 4.553$ ,  $p = .024$ ,  $\eta^2 = .144$ ). Post-hoc tests with Bonferroni correction revealed that this effect was mainly driven by a difference in RTs between the two extreme conditions. Participants were significantly faster to respond to objects presented at their true size ( $M = 286.02$ ,  $SD = 62.97$ ) compared to objects that were presented at 1/100th of their familiar size ( $M = 294.89$ ,  $SD = 62.92$ ) ( $p_{corr} = .018$ ). However, the differences between 1/10<sup>th</sup> objects ( $M = 289.29$ ,  $SD = 59.66$ ) and those shown at their true size ( $p_{corr} = .421$ ), along with 1/10<sup>th</sup> compared to 1/100<sup>th</sup> objects ( $p_{corr} = .399$ ) were not significant. Although some of the comparisons did not reach significance, visual inspection of figure 4a clearly shows a linear trend, which was supported by a linear contrast analysis ( $F(1, 27) = 8.927$ ,  $p = .006$ ,  $\eta^2 = .248$ ). Additionally, a Pearson's correlation coefficient ( $r$ ) was calculated between RT for each image averaged across participants and log-transformed real-

world size (See Table 4 in Supplementary Materials for details). The correlation revealed a positive relationship between RTs and real-world size ( $r(88)=.334, p=.001$ ), (Figure 5A).

These results demonstrate that simple RTs are affected by familiar size in a systematic manner when the discrepancy between physical size and familiar size increases proportionally.

-----Please insert Figure 4 about here -----

-----Please insert Figure 5 about here -----

### **Experiment 4: From normal to over-sized stimuli**

#### **Method**

##### **Participants**

Twenty-eight participants took part in this experiment (3 male, 7 left eye-dominant and 3 left handed). They ranged in age from 18 to 55 years ( $M=23, SD=9.42$ ).

##### **Apparatus**

As in Experiment 3, images of real objects were selected based on the proportion of their familiar size to their physical size presented on the screen and were categorised as: x1 (normal-sized stimuli), x2 (over-sized stimuli) and x10 (over-sized stimuli). As it was not possible to directly mirror the different intervals in familiar size as in Experiment 3, due to the fact that very small objects (e.g. tip of a pen) seen as enlarged can become unfamiliar and difficult to recognise, we simply chose 3 different size intervals that would not include objects smaller than 0.4 cm in the real world.

Those images classified as '1' had a familiar size that ranged between 5 cm and 7 cm ( $M= 6.58$  cm,  $SD=0.89$ ), those classified as 'x2' had a familiar size that ranged between 2.5

cm and 3.5 cm ( $M= 2.96$  cm,  $SD= 0.52$ ) and those classified as 'x10' had a familiar size that ranged between 0.4 cm and 1.6cm ( $M= 0.96$  cm,  $SD= 0.33$ ).

Again, all images were balanced for luminance (30 cd/m<sup>2</sup>) and were scaled to subtend 5.7° of visual angle. Therefore, only '1' stimuli were the same size as in the real world (normal-sized), whereas 'x2' and 'x10' stimuli were bigger than in the real world (over-sized), some images were the same as those used in the previous experiment (see Table 3 in Supplementary Materials).

### Design and Procedure

See Experiment 3.

### Results

Accuracy was high, ranging from 88.94 to 99.52% ( $M=97.28$ ,  $SD= 0.03$ ). A one-way repeated measures ANOVA was conducted with Familiar size (1 vs. x2 vs. x10) as the main factor. The ANOVA revealed no main effect of Familiar size on RTs ( $F(2, 54) = 1.332$ ,  $p = .273$ ,  $\eta^2 = .047$ ; x1:  $M=290.64$ ,  $SD=41.96$ ; x2:  $M= 294.92$ ,  $SD=44.07$ ; x10:  $M=293.12$ ,  $SD=42.09$ ). Similarly, the linear contrast analysis did not reach significance ( $F(1,27) = 1.241$ ,  $p = .275$ ,  $\eta^2 = .044$ ) (Figure 4b). As for Experiment 3, a Pearson's correlation coefficient ( $r$ ) was calculated between RTs for each image averaged across participants and the log-transformation of each object's size. RT did not correlate with real-world size ( $r(88) = -0.001$ ,  $p = 0.496$ ) (Figure 5B).

This lack of effect on RTs for small objects shown at a magnified size could be attributed to a reduced discrepancy (i.e. less variability) between physical and familiar size, which was not the case for Experiment 3 where the size range was much wider. Specifically, while the range difference between the two extreme categories in Experiment 3 was 4.51 in log-units, this difference in Experiment 4 was only 1.25 in log-units. An alternative explanation for the results could be that participants were unable to correctly identify some of

the magnified objects, especially those from the smallest category (x10). In fact, while we are used to see large objects as small on the retina (as in Experiment 3), we rarely experience tiny objects as big on the retina, unless they are held very close to the eyes (in Experiment 4, x10 objects would need to be placed at 6 cm of viewing distance from the eyes to generate 5.7° of visual angle). Therefore, magnified small objects could have been treated by the visual system as unfamiliar images.

### **Experiment 5: The effect of familiar size on manual estimates.**

The aim of this experiment was to establish if the effect of familiar size on RTs, observed under reduced viewing conditions (Experiments 2 and 3) could be attributed to perceived size, as previous studies have demonstrated that simple RTs are faster in response to objects that are perceived as bigger even when their retinal image is constant (Sperandio et al., 2009; Sperandio, Savazzi, & Marzi, 2010; Plewan, Weidner, & Fink, 2012; Savazzi, Emanuele, Scalf, & Beck, 2012).

In Experiment 5, participants judged the perceived size of stimuli under restricted conditions of observation. Therefore, if the effects of familiar size on RTs reported in Experiments 2 and 3 are due to perceived size, then one might expect to find that objects shown at their familiar size will be perceived as bigger compared to those objects presented at a size that is incongruent with their known size. In other words, those objects that generated faster RTs (i.e. stimuli congruent with prior knowledge about real-world size) should also be estimated as larger than those objects that generated slower RTs (i.e. incongruent stimuli) in agreement with the findings discussed above where perceived larger objects are responded to more quickly than perceived smaller objects

### **Participants**

Thirty participants took part in this experiment, however, two were removed due to failure to comply with the task instructions and technical issues. The remaining participants (7 left handed, 7 left eye-dominant and 7 males) ranged in age from 18 to 44 years ( $M=20$ ,  $SD=4.79$ ).

### **Apparatus**

The same images and apparatus as described in Experiment 3 were used here. Measurements of manual estimations were taken in millimeters by the experimenter using a pair of digital callipers. Two little points were drawn on the forefinger and thumb of the participant's right hand and served as markers for the measurements.

### **Design and Procedure**

Participants were asked to estimate the size of the object presented on the screen using their thumb and forefinger. Perceived size was measured by means of manual size estimation. Manual size estimation has been widely used in the literature to record perceived size (e.g. Haffenden & Goodale, 1998; Sperandio, Lak, & Goodale, 2012; Westwood & Goodale, 2003), even under reduced viewing conditions (Marotta & Goodale, 2001). Its effectiveness and sensitivity in measuring changes in size perception have been previously demonstrated (e.g. Franz, 2003; Franz & Gegenfurtner, 2008). The advantages of using MSE are that the fingers return to the starting position (i.e. pinch returns to an estimate of zero) and the estimate can be made without looking at the hand, making this measure more 'implicit' and less prone to memory effects and anchoring biases than other forms of perceptual reports (e.g. Bolles & Bailey, 1956). Images of familiar objects were divided into two blocks, depending on their dimension properties (width vs. height), which determined the orientation of the hand during the manual estimation task. For example, an elongated object such as a glue stick presented vertically, would meet the criterion for maximum height but not width. There were

41 images that met the criterion for maximum width and 46 images that met the criterion for maximum height. The remaining images that met both criteria ( $N = 3$ ) were placed into the width category to even out the number of trials with respect to the height category. Prior to testing, participants were instructed on how to perform the manual estimation task according to the 'width' or 'height' block and were asked to complete a practice session involving both hand orientations. Participants used their dominant hand and eye to perform the task. The two experimental blocks (width vs. height) were presented in counterbalanced order. Stimuli within each block were randomly presented.

At the beginning of each block, participants received instructions about the relevant dimension to be judged (width vs. height). A typical trial consisted of a cross appearing on the screen, followed by an image of an object. The participant then manually estimated the object. Once the participant was happy with his/her judgment, the experimenter recorded the manual estimation using the digital callipers. In between trials, participants were required to rest their hand on the table with their right forefinger and thumb pinched together. A break was given between the two blocks.

The experiment consisted of 30 trials for each of the three size intervals (1 vs. 1/10 vs. 1/100), yielding 90 trials in total.

## Results

A one-way ANOVA with Familiar size (1 vs. 1/10 vs. 1/100) as main factor was carried out on manual estimates. As Mauchly's test of sphericity was significant ( $\chi^2(2) = 22.249, p < .001$ ), the Greenhouse-Geisser value was reported.

The effect of Familiar size on manual estimates was significant ( $F(1.27, 34.285) = 11.905, p = .001, \eta^2 = .306$ ). Post-hoc tests with Bonferroni correction revealed that the difference between x1 ( $M = 57.29, SD = 15.77$ ) and 1/100 ( $M = 62.99, SD = 17.59$ ) was

significant ( $p_{corr}=.003$ ), along with the difference between 1/10 ( $M=58.88$ ,  $SD=16.15$ ) and 1/100 ( $p_{corr}=.009$ ). Although the comparison between x1 and 1/10 was only approaching significance ( $p_{corr}=.074$ ), a proportionate relationship between the means of each size condition can be observed in figure 6 and is supported by the linear contrast analysis, which was significant ( $F(1, 27) = 13.867$ ,  $p=.001$ ,  $\eta p^2 = .339$ ).

-----Please insert Figure 6 about here -----

These results show that the perceived size of real-world objects can be influenced by their familiar size, such that when objects were presented at the same retinal size, participants perceived those objects that are known to be big as larger than those that are known to be small. These findings are consistent with previous studies on the effects of familiar size on perceived size and distance under restricted conditions of observation. For example, it has been demonstrated that when photographs of a golf ball and a baseball were presented at the same retinal size and distance, under reduced viewing conditions, observers perceived the baseball to be bigger and further away than the golf ball (Ono, 1969). Similarly, observers judged coins of small familiar size as closer and smaller than their familiar bigger counterparts, despite being presented at the same retinal size and distance (Epstein & Baratz, 1964). These previous reports, along with our results, support the hypothesis that familiar size influences our perception of objects' size, when depth cues are removed. However, contrary to our expectations, the direction of the effect of familiar size suggests that perceived size cannot account for the RT advantage reported in the above experiments. It rather reflects the level of congruence between stored representation of size and the actual object's size, whereby consistency across these two sources of information determines faster RTs.

## General Discussion

Previous research has suggested that simple RTs reflect perceived size, particularly when the visual stimulus consists of a familiar object (Sperandio et al., 2009). In this study, we investigated the extent to which familiar object size influences simple RTs and perceived size. Our findings showed that whilst under regular viewing conditions RTs were unaffected by familiar size (Experiment 1), we did observe an effect of familiar size on RTs under reduced viewing conditions, whereby familiar small objects were responded to faster than familiar big objects of constant retinal size and luminance (Experiment 2). This implies that the availability of reliable depth cues given by natural viewing conditions (Experiment 1) counteracts familiar size as a pictorial depth cue. In contrast, under reduced viewing conditions (Experiment 2), familiar size acts as a depth cue and affects the speed of processing. This effect appeared to be proportionate to the degree of incongruence between familiar and physical size of the stimuli shown on the screen, at least when the incongruence was big enough to influence detection time (Experiments 3 and 4) or it did not concern magnified small objects (Experiment 4). Under reduced viewing conditions, an effect of familiar object size on perceived size, as measured with a manual estimation task, was also observed. Contrary to our expectations, however, manual estimates were smaller for familiar small objects than familiar big objects, suggesting that the RT advantage reported here was unrelated to perceived size (Experiment 5). Finally, the effect of animacy was investigated in the first two experiments. As it turned out, participants were faster to detect animate compared to inanimate objects, under both full (Experiment 1) and reduced viewing conditions (Experiment 2). In the ensuing discussion, we will consider the effects of animacy and familiar size on RTs and their possible neural underpinnings.

### Animacy effect



Faster and more accurate responses to animate than inanimate object categories are typically reported in the literature. For example, New, Cosmides, and Tooby (2007) provided evidence for an attentional bias for animals. Using a change detection task, where participants were asked to detect if and where a change occurred in a scene, they observed faster and more accurate responses to changes to those scenes involving animals as targets compared to non-animals. Importantly, this preferential attention to animate images was unrelated to low-level perceptual features (e.g. luminance, eccentricity and pixel size) as well as by how interesting the targets were judged to be, suggesting that this behavioural advantage was indeed the result of a category-driven process (New et al., 2007). Similarly, Proverbio, Del Zotto, and Zani (2007) reported a RT advantage for animal compared to non-animal stimuli in a choice detection task. By recording ERPs, they found that this behavioural advantage was associated with stronger responses in the right occipital-temporal cortex (Proverbio et al., 2007). Animacy effects have also been observed in several eye movement studies in which participants were instructed to make a saccade as fast as possible to a target (i.e. saccadic choice task). Using this paradigm, faster and more accurate categorizations were reported for scenes containing animate than inanimate objects (Crouzet, Joubert, Thorpe, & Fabre-Thorpe, 2012; Crouzet, Kirchner, & Thorpe, 2010; Kirchner & Thorpe, 2006).

A great deal of neuropsychological and neuroimaging research suggests that these category-specific biases might rely upon separate neural networks for animate and inanimate objects (e.g. Chouinard & Goodale, 2010; Griffiths, Marslen-Wilson, Stamatakis, & Tyler, 2013; Konkle & Caramazza, 2013; Chao, Weisberg, & Martin, 2002; Gerlach, 2007). For instance, a dissociation of performance between animate and inanimate objects following brain lesions typically in the temporal lobe is well-documented in the literature, demonstrating that conceptual knowledge is organized in specific semantic categories that can be selectively impaired (e.g. Caramazza & Shelton, 1998; Sacchett & Humphreys, 1992;

Warrington & McCarthy, 1983; Warrington & Shallice, 1984; for a review see Mahon & Caramazza, 2009). Interestingly, based on the evidence that domain-specific semantic deficits seem to more frequently involve the recognition of animals/living things compared to any other domain, an evolutionary advantage for a separate and rapid animate/living pathway which operates at the early stages of the object identification process and links into the emotion regulation system has been suggested (e.g. Caramazza & Shelton, 1998; Mormann et al., 2011; New et al., 2007). This hypothesis has been further supported by a combined fMRI and MEG study, which demonstrated that specialised networks for object categorisation can be observed at the early stages of visual processing (Cichy et al., 2014). Specifically, the earliest separable response for animate compared to inanimate objects was registered at 157 ms after stimulus onset and was related to activity within the inferior temporal cortex (Cichy et al., 2014).

Taken together, these findings provide converging evidence for the existence of separate mechanisms for animate and inanimate object categories, which might explain the animacy effects described above. In line with previous findings, our participants were faster to simply detect the onset of images of animals compared to non-animals in the current investigation, providing further support to the idea that animate objects are attended to more quickly than inanimate objects. To the best of our knowledge, this is the first demonstration of an animacy effect on simple RTs. Notably, the animacy effect observed here was unrelated to low-level visual features of the stimuli, such as differences in spatial frequency between the two semantic categories, which argues against the hypothesis that the behavioural advantage for animate objects relies merely on underlying statistical properties of the images (Harel & Bentin, 2009; Viggiano, Costantini, Vannucci, & Righi, 2004). Instead, our findings might be indicative of an evolutionary advantage for animate objects, as originally suggested by Caramazza and Shelton (1998).

### **Familiar size effect**

Behavioral effects of familiar size have also been reported in the literature. For example, it has been established that the perceived size of an object is influenced by its familiar size, such that individuals are more accurate at estimating the size of familiar than non-familiar objects (e.g. Bolles & Bailey, 1956; Slack, 1956). This effect of object familiarity is evident particularly under reduced viewing conditions, where the usual cues to distance, such as vergence, accommodation, and retinal disparity, are restricted (e.g. Gogel, 1969; Epstein, 1963; Ittelson, 1951; Schiffman, 1967).

Moreover, it has been shown that familiar size can affect speed of processing. In a Stroop-like paradigm, pairs of familiar objects were shown and participants were asked to indicate which one of the two images was visually bigger. Responses were faster if the difference in size between the two images was congruent to the true size difference between the familiar objects, e.g. a big elephant and a small mouse would be responded to more quickly than a small elephant and a big mouse (Konkle & Oliva, 2012a). By the same token, Gabay, Leibovich, Henik, & Gronau (2013b) demonstrated in a priming study involving objects of equal retinal size but different familiar size as primes and integers as targets that participants' response time in making odd-even judgments of the target increased when the conceptual size of the object image was incongruent with the integer's value. For example, an elephant prime followed by the number 1 determined a slower response compared to a trial in which the number 1 was primed by the picture of a mouse. These findings show that a cost in response time is incurred when there is a lack of shared conceptual representations between pairs of familiar objects (Konkle & Oliva, 2012a) or prime and target stimuli (Gabay et al, 2013b) and suggest that familiar size is an automatic object property. Here, we found that simple RT, a stereotyped motor response, is also influenced by familiar size with faster responses to stimuli that are physically closer to their real-world size. One might argue that

the effect of familiar size on simple RTs depends on the degree of congruence between the physical size of the object displayed on the screen and its internal representation of size, whereby a cost is incurred for increased degrees of incongruence, which slow down RTs. In other words, it is conceivable that the difference in speed of processing (~7 ms on average for significant comparisons) between objects of different familiar-size intervals could reflect the time taken for the visual system to process the stimulus, access to the stored representation of size, and perform a computational adjustment with respect to such an internal representation, with a benefit for object images that physically match their internal representation. However, further electrophysiological research would need to be conducted to confirm this hypothesis.

Interestingly, there are instances in which response times tend to favour big objects, so that familiar big objects are responded to more quickly than small objects. One such study involved a lexical decision task, where participants were simply asked to identify if the letters presented formed a word (e.g. apple) or a non-word (e.g. lerop) (Sereno, O'Donnell, & Sereno, 2009). It was found that people were faster to classify word-stimuli when the word was a familiar big object (e.g. bus) compared to a familiar small object (e.g. pea). A similar effect was reported in Konkle and Oliva's (2012a) study described above, whereby along with the congruency finding, results revealed that familiar big objects were responded to faster than familiar small objects (Konkle & Oliva, 2012a). This is also in agreement with the RT study carried out by Sperandio et al (2009) in which participants were faster to react to perceived larger stimuli than perceived smaller stimuli of fixed retinal size, but only when images of familiar objects (i.e. tennis balls) instead of simple circles were presented, highlighting once again the importance of object familiarity on response speed. Our results indicate that the advantage of response speed did not correspond to the estimates of perceived size, which tended to increase with the real-world size. Therefore, our findings do not support the notion that "bigger is better" as sometimes reported in the literature (e.g. Konkle & Oliva,

2012a; Sereno et al., 2009; Sperandio et al., 2009; Witt & Proffitt, 2005), but rather "congruent is better", where smaller deviations from known size resulted in faster detection times.

From a neural perspective, it has been suggested that different brain areas represent objects of different familiar size in a manner that reflects an organization based upon their real-world size (Konkle & Oliva, 2012b). Using fMRI, Konkle and Oliva (2012b) showed that the occipito-temporal cortex is topographically organized according to familiar size. Particularly, the parahippocampal cortices (PHC) was preferentially activated by familiar big objects, whereas regions such as the lateral occipital (LO) complex and occipital temporal sulcus (OTS) were sensitive to familiar small objects. This activity was resistant to both low-level (e.g. changes in retinal size) and high-level (e.g. abstract concepts of size) effects (Konkle & Oliva, 2012b). More recently, Konkle and Caramazza (2013) has provided evidence for a tripartite distinction between response preferences for animals, big objects, and small objects in the occipital-temporal cortex. In particular, big objects were reported to preferentially activate medial regions (including PHC), whereas small objects activated more lateral regions of the occipital temporal cortex, such as the inferior temporal gyrus (ITG). Images of animals also preferentially activated more lateral regions in the occipital temporal cortex (Konkle & Caramazza, 2013). Interestingly, this tripartite organization of the ventral stream relies on separate sub-networks as demonstrated in a resting-state study (Konkle & Caramazza, 2016). Correspondingly, it has been shown that a patient with inferior occipitoparietal cortex damage was more accurate at judging the distance of familiar (e.g. plastic bottle) than unfamiliar objects (e.g. cube), suggesting that her depth perception, which was severely impaired by the lesion, could to some extent benefit from familiar size (Berryhill, Fendrich, & Olson, 2009).

To conclude, familiar size can influence the time it takes to process and respond to familiar stimuli with an advantage for those stimuli subtending retinal image sizes that are closer to their long-term representations of size. This behavioural advantage might reflect the spatial organization of neural response preferences as dictated by real-world size

### **Conclusion**

In summary, here we demonstrate for the first time that a relationship exists between simple RTs, a stereotyped speeded response to stimulus onset, and familiar size, such that a simple RT advantage was recorded for objects shown closer to their 'true' size. The familiar size effect on RTs did not reflect perceived size but rather the degree of congruency to a stored 'normative' size. Furthermore, we have shown that simple RTs were modulated by animacy with faster detection times for images of animals than non-animals. Taken as a whole, these findings suggest that both familiar size and animacy are automatic features of object representations that might take place at an early stage in the processing of visual information and might be sub-served by distinct neural processes.

## References

- Berryhill, M., Fendrich, R., & Olson, I. (2009). Impaired distance perception and size constancy following bilateral occipitoparietal damage. *Experimental Brain Research*, *194*(3), 381–93. <https://doi.org/10.1007/s00221-009-1707-7>
- Bolles, R. C., & Bailey, D. E. (1956). Importance of object recognition in size constancy. *Journal of Experimental Psychology*, *51*(3), 222–5. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/13306868>
- Brady, T. F., Konkle, T., Gill, J., Oliva, A., & Alvarez, G. a. (2013). Visual long-term memory has the same limit on fidelity as visual working memory. *Psychological Science*, *24*(6), 981–90. <https://doi.org/10.1177/0956797612465439>
- Brodeur, M. B., Dionne-Dostie, E., Montreuil, T., & Lepage, M. (2010). The bank of standardized stimuli (BOSS), a new set of 480 normative photos of objects to be used as visual stimuli in cognitive research. *PLoS ONE*, *5*(5). <https://doi.org/10.1371/journal.pone.0010773>
- Caramazza, A., & Shelton, J. (1998). Domain-specific knowledge systems in the brain: the animate-inanimate distinction. *Journal of Cognitive Neuroscience*, *10*, 1–34.
- Chao, L. L., Weisberg, J., & Martin, A. (2002). Experience-dependent modulation of category-related cortical activity. *Cerebral Cortex*, *12*(5), 545–551. <https://doi.org/10.1093/cercor/12.5.545>
- Cheng, C. Y., Yen, M. Y., Lin, H. Y., Hsia, W. W., & Hsu, W. M. (2004). Association of ocular dominance and anisometric myopia. *Investigative Ophthalmology & Visual Science*, *45*(8), 2856–60. <https://doi.org/10.1167/iovs.03-0878>
- Cichy, R. M., Pantazis, D., & Oliva, A. (2014). Resolving human object recognition in space and time. *Nature Neuroscience*, *17*(3), 455–62. <https://doi.org/10.1038/nn.3635>
- Clarke, A., Taylor, K. I., Devereux, B., Randall, B., & Tyler, L. K. (2013). From perception to conception: how meaningful objects are processed over time. *Cerebral Cortex*, *23*(1), 187–197. <https://doi.org/10.1093/cercor/bhs002>
- Crouzet, S. M., Joubert, O. R., Thorpe, S. J., & Fabre-Thorpe, M. (2012). Animal Detection Precedes Access to Scene Category. *PLoS ONE*, *7*(12), 1–10. <https://doi.org/10.1371/journal.pone.0051471>
- Crouzet, S. M., Kirchner, H., & Thorpe, S. J. (2010). Fast saccades toward faces: face detection in just 100 ms. *Journal of Vision*, *10*(4), 1–17. <https://doi.org/10.1167/10.4.16>
- Epstein, W. (1963). The influence of assumed size on apparent distance. *The American Journal of Psychology*, *76*(2), 257–265. Retrieved from <http://www.jstor.org/stable/1419162>
- Epstein, W., & Baratz, S. S. (1964). Relative Size in Isolation As a Stimulus for Relative Perceived Distance. *Journal of Experimental Psychology*, *67*(6), 507–513. <https://doi.org/10.1037/h0043588>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. <https://doi.org/10.3758/BF03193146>
- Franz, V. H. (2003). Manual size estimation: A neuropsychological measure of perception? *Experimental Brain Research*, *151*(4), 471–477. <https://doi.org/10.1007/s00221-003->

1477-6

- Franz, V. H., & Gegenfurtner, K. R. (2008). *Grasping visual illusions: Consistent data and no dissociation. Cognitive Neuropsychology* (Vol. 25).  
<https://doi.org/10.1080/02643290701862449>
- Gabay, S., Leibovich, T., Henik, A., & Gronau, N. (2013a). Size before numbers: Conceptual size primes numerical value. *Cognition*, *129*(1), 18–23.  
<https://doi.org/10.1016/j.cognition.2013.06.001>
- Gabay, S., Leibovich, T., Henik, A., & Gronau, N. (2013b). Size before numbers: Conceptual size primes numerical value. *Cognition*, *129*(1), 18–23.  
<https://doi.org/10.1016/j.cognition.2013.06.001>
- Gerlach, C. (2007). A review of functional imaging studies on category specificity. *Journal of Cognitive Neuroscience*, *19*(2), 296–314. <https://doi.org/10.1162/jocn.2007.19.2.296>
- Gogel, W. C. (1969). The effect of object familiarity on the perception of size and distance. *The Quarterly Journal of Experimental Psychology*, *21*(3), 239–247.  
<https://doi.org/10.1080/14640746908400218>
- Grill-Spector, K., & Kanwisher, N. (2005). As soon as you know it is there , you know what it is. *Psychological Science*, *16*(2), 152–160.
- Haffenden, A. M., & Goodale, M. A. (1998). The effect of pictorial illusion on prehension and perception. *Journal of Cognitive Neuroscience*, *10*(1), 122–136.  
<https://doi.org/10.1162/089892998563824>
- Harel, A., & Bentin, S. (2009). Stimulus type, level of categorization, and spatial-frequencies utilization: implications for perceptual categorization hierarchies. *Journal of Experimental Psychology. Human Perception and Performance*, *35*(4), 1264–73.  
<https://doi.org/10.1037/a0013621>
- Holway, A. H., & Boring, E. G. (1941). Determinants of apparent visual size with distance variant. *The American Journal of Psychology*, *54*(1), 21–37.
- Humphreys, G. W., Price, C. J., & Riddoch, M. J. (1999). From objects to names: A cognitive neuroscience approach. *Psychological Research*, *62*(2–3), 118–130.  
<https://doi.org/10.1007/s004260050046>
- Ittelson, W. H. (1951). Size as a cue to distance: Static localization. *The American Journal of Psychology*, *64*(1), 54–67. <https://doi.org/10.1177/036354657700500303>
- Johnson, R. C., McClearn, G. E., Yuen, S., Nagoshi, C. T., Ahern, F. M., & Cole, R. E. (1985). Galton's data a century later. *American Psychologist*, *40*(8), 875–892.  
<https://doi.org/10.1037/0003-066X.40.8.875>
- Kirchner, H., & Thorpe, S. J. (2006). Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. *Vision Research*, *46*(11), 1762–1776.  
<https://doi.org/10.1016/j.visres.2005.10.002>
- Konkle, T., & Caramazza, A. (2013). Tripartite organization of the ventral stream by animacy and object size. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *33*(25), 10235–42. <https://doi.org/10.1523/JNEUROSCI.0983-13.2013>
- Konkle, T., & Caramazza, A. (2016). The Large-Scale Organization of Object-Responsive Cortex Is Reflected in Resting-State Network Architecture. *Cerebral Cortex (New York, N.Y. : 1991)*, 1–13. <https://doi.org/10.1093/cercor/bhw287>
- Konkle, T., & Oliva, A. (2007). Normative representation of objects: Evidence for an



- ecological bias in object perception and memory. *Proceedings of the 29th Annual Meeting of the ....* Retrieved from [http://cvcl.mit.edu/Papers/KonkleOliva\\_CogSci07.pdf](http://cvcl.mit.edu/Papers/KonkleOliva_CogSci07.pdf)
- Konkle, T., & Oliva, A. (2011). Canonical visual size for real-world objects. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(1), 23–37. <https://doi.org/10.1037/a0020413>
- Konkle, T., & Oliva, A. (2012a). A familiar-size Stroop effect: real-world size is an automatic property of object representation. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(3), 561–9. <https://doi.org/10.1037/a0028294>
- Konkle, T., & Oliva, A. (2012b). A real-world size organization of object responses in occipitotemporal cortex. *Neuron*, *74*(6), 1114–24. <https://doi.org/10.1016/j.neuron.2012.04.036>
- Kovalenko, L. Y., Chaumon, M., & Busch, N. A. (2012). A pool of pairs of related objects (POPORO) for investigating visual semantic integration: Behavioral and electrophysiological validation. *Brain Topography*, *25*(3), 272–284. <https://doi.org/10.1007/s10548-011-0216-8>
- Li, F. F., VanRullen, R., Koch, C., & Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(14), 9596–9601. <https://doi.org/10.1073/pnas.092277599>
- Long, B., Konkle, T., Cohen, M. A., & Alvarez, G. A. (2016). Mid-level perceptual features distinguish objects of different real-world sizes. *Journal of Experimental Psychology: General*, *145*(1), 95–109. <https://doi.org/10.1037/xge0000130>
- Mack, M. L., Gauthier, I., Sadr, J., & Palmeri, T. J. (2008). Object detection and basic-level categorization: sometimes you know it is there before you know what it is. *Psychonomic Bulletin & Review*, *15*(1), 28–35. <https://doi.org/10.3758/PBR.15.1.28>
- Mahon, B. Z., Anzellotti, S., Schwarzbach, J., Zampini, M., & Caramazza, A. (2009). Category-Specific Organization in the Human Brain Does Not Require Visual Experience. *Neuron*, *63*(3), 397–405. <https://doi.org/10.1016/j.neuron.2009.07.012>
- Mahon, B. Z., & Caramazza, A. (2009). Concepts and categories: a cognitive neuropsychological perspective. *Annual Review of Psychology*, *60*(1), 27–51. <https://doi.org/10.1146/annurev.psych.60.110707.163532>
- Marotta, J. J., & Goodale, M. A. (2001). Role of familiar size in the control of grasping. *Journal of Cognitive Neuroscience*, *13*(1), 8–17. <https://doi.org/10.1162/089892901564135>
- Martin, A. (2007). The Representation of Object Concepts in the Brain. *Annual Review of Psychology*, *58*(1), 25–45. <https://doi.org/10.1146/annurev.psych.57.102904.190143>
- McMullen, P. A., & Purdy, K. S. (2006). Category-specific effects on the identification of non-manipulable objects. *Brain and Cognition*, *62*(3), 228–240. <https://doi.org/10.1016/j.bandc.2006.06.002>
- Mormann, F., Dubois, J., Kornblith, S., Milosavljevic, M., Cerf, M., Ison, M., ... Koch, C. (2011). A category-specific response to animals in the right human amygdala. *Nature Neuroscience*, *14*(10), 1247–1249. <https://doi.org/10.1038/nn.2899>
- New, J., Cosmides, L., & Tooby, J. (2007). Category-specific attention for animals reflects ancestral priorities, not expertise. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(42), 16598–603. <https://doi.org/10.1073/pnas.0703913104>

- Ono, H. (1969). Apparent distance as a function of familiar size. *Journal of Experimental Psychology*, 79(1), 109–115. <https://doi.org/10.1037/h0026879>
- Osaka, N. (1976). Reaction time as a function of peripheral retinal locus around fovea: effect of stimulus size. *Perceptual and Motor Skills*, 42(43), 603–606. <https://doi.org/10.2466/pms.1976.43.2.603>
- Pins, D., & Bonnet, C. (1996). On the relation between stimulus intensity and processing time : Pieron's law and choice reaction time. *Perception*, 58(3), 390–400.
- Plewan, T., Weidner, R., & Fink, G. R. (2012). The influence of stimulus duration on visual illusions and simple reaction time. *Experimental Brain Research*, 223(3), 367–75. <https://doi.org/10.1007/s00221-012-3265-7>
- Praß, M., Grimsen, C., König, M., & Fahle, M. (2013). Ultra Rapid Object Categorization: Effects of Level, Animacy and Context. *PLoS ONE*, 8(6). <https://doi.org/10.1371/journal.pone.0068051>
- Proverbio, A. M., Del Zotto, M., & Zani, A. (2007). The emergence of semantic categorization in early visual processing: ERP indices of animal vs. artifact recognition. *BMC Neuroscience*, 8, 24. <https://doi.org/10.1186/1471-2202-8-24>
- Sacchett, C., & Humphreys, G. W. (1992). Calling a squirrel a squirrel but a canoe a wigwam: a category-specific deficit for artefactual objects and body parts. *Cognitive Neuropsychology*, 9(1), 73–86. <https://doi.org/10.1080/02643299208252053>
- Savazzi, S., Emanuele, B., Scalf, P., & Beck, D. (2012). Reaction times and perceptual adjustments are sensitive to the illusory distortion of space. *Experimental Brain Research*, 218(1), 119–128. <https://doi.org/10.1007/s00221-012-3012-0>
- Schiffman, R. H. (1967). Size-Estimation of familiar objects under informative and reduced conditions of viewing. *The American Journal of Psychology*, 80(2), 229–235. <https://doi.org/10.1177/036354657700500303>
- Sereno, S. C., O'Donnell, P. J., & Sereno, M. E. (2009). Size matters: bigger is faster. *Quarterly Journal of Experimental Psychology (2006)*, 62(6), 1115–1122. <https://doi.org/10.1080/17470210802618900>
- Slack, C. W. (1956). Familiar size as a cue to size in the presence of conflicting cues. *Journal of Experimental Psychology*, 52(3), 194–198. <https://doi.org/10.1037/h0044669>
- Sperandio, I., Lak, A., & Goodale, M. A. (2012). Afterimage size is modulated by size-contrast illusions. *Journal of Vision*, 12(2), 1–10. <https://doi.org/10.1167/12.2.18>
- Sperandio, I., Savazzi, S., Gregory, R. L., & Marzi, C. A. (2009). Visual reaction time and size constancy. *Perception*, 38(11), 1601–1609. <https://doi.org/10.1068/p6421>
- Sperandio, I., Savazzi, S., & Marzi, C. A. (2010). Is simple reaction time affected by visual illusions? *Experimental Brain Research*, 201(2), 345–50. <https://doi.org/10.1007/s00221-009-2023-y>
- VanRullen, R., & Thorpe, S. J. (2001). Is it a bird? Is it a plane? Ultra-rapid visual categorisation of natural and artificial objects. *Perception*, 30(6), 655–668. <https://doi.org/10.1068/p3029>
- Viggiano, M. P., Costantini, A., Vannucci, M., & Righi, S. (2004). Hemispheric asymmetry for spatially filtered stimuli belonging to different semantic categories. *Cognitive Brain Research*, 20(3), 519–524. <https://doi.org/10.1016/j.cogbrainres.2004.03.010>
- Warrington, E. K., & McCarthy, R. (1983). Category specific access dysphasia. *Brain*, 106,

859–878.

Warrington, E. K., & Shallice, T. (1984). Category Specific Semantic Impairments. *Brain*, *107*, 829–854.

Westwood, D. A., & Goodale, M. A. (2003). A haptic size-contrast illusion affects size perception but not grasping. *Experimental Brain Research*, *153*(2), 253–259.  
<https://doi.org/10.1007/s00221-003-1599-x>

Witt, J. K., & Proffitt, D. R. (2005). See the ball, hit the ball: Apparent ball size is correlated with batting average. *Psychological Science*, *16*(12), 937–938.  
<https://doi.org/10.1111/j.1467-9280.2005.01640.x>

*Figure 1.* Stimuli and trial sequence: (A) Example of stimulus display for each experimental condition in Experiments 1 and 2. Thirty stimuli were chosen depending on their familiar size (small vs. big) and were either animate or inanimate objects. All images were adjusted to match in average luminance ( $30 \text{ cd/m}^2$ ) and fit inside a frame of fixed size ( $6^\circ \times 6^\circ$ ). The aspect ratio of each image was maintained to prevent distortion. From Left to Right: small inanimate (key), small animate (butterfly), big inanimate (lorry) and big animate (rhinoceros). (B) Example of experimental trial sequence and timing. At the beginning of each trial a fixation cross was presented for 1000 ms followed by a warning signal. Next, a random interval (400-600 ms) was introduced, followed by a stimulus of 80 ms. Then, a blank screen was presented until the response button was pressed or the time limit of 2000 ms was reached. During catch trials, the stimulus was replaced by a blank screen and participants were asked to refrain from responding.

*Figure 2.* Results of Experiments 1 and 2: (A) Experiment 1. Mean RTs (ms) as a function of Familiar size and Animacy, under regular viewing conditions. (B) Experiment 2. Mean RT (ms) as a function of Familiar size and Animacy, under reduced viewing conditions. Error bars in both graphs represent within-subjects 95% confidence intervals (WSCl) (Masson & Loftus, 2003).

*Figure 3.* Relationship between familiar size and distance for a specified visual angle. The dashed line represents the viewing distance used in Experiments 3-5 (i.e. 60 cm). The distance (indicated by the vertical lines) and the size of the images are proportional to the typical size-distance relationship in order to subtend the visual angle tested in Experiment 3 (i.e.  $5.7^\circ$ ). Exemplars from each category of familiar size are depicted (i.e. x1, 1/10, and 1/100).

*Figure 4.* Results of Experiment 3 and 4: (A) Experiment 3 Mean RTs (ms) as a function of familiar size, under reduced viewing conditions (left to right: increasing familiar size). The asterisk denotes a significant difference ( $p_{corr} < 0.05$ ) between means based on Bonferroni corrected t-tests. (B) Experiment 4. Mean RTs (ms) as a function of familiar size, under reduced viewing conditions (left to right: decreasing familiar size). Error bars in both graphs represent within-subjects 95% confidence intervals (WSCIs) (Masson & Loftus, 2003).

*Figure 5.* Correlation between the mean RT and actual real-world size of the stimuli used in (A) Experiment 3 and (B) Experiment 4. The x-axis corresponds to the log-transformed size of the object (originally in centimetres) and the y-axis corresponds to the mean RT (ms) for each image.

*Figure 6.* Results of Experiment 5. Mean manual size estimation (cm) for the three categories of familiar size (increasing familiar size from left to right) with estimations collapsed across hand orientation. Error bars represent within-subjects 95% confidence intervals (WSCIs) (Masson & Loftus, 2003). The asterisk denotes a significant difference ( $p_{corr} < 0.01$ ) between means based on Bonferroni corrected t-tests.

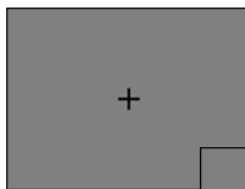
Figure 1

a

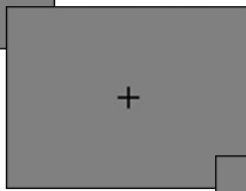


b

Fixation:  
1000 ms



Cue:  
1000 ms



Random interval:  
400-600 ms



Target:  
80 ms



Reponse:  
2000 ms

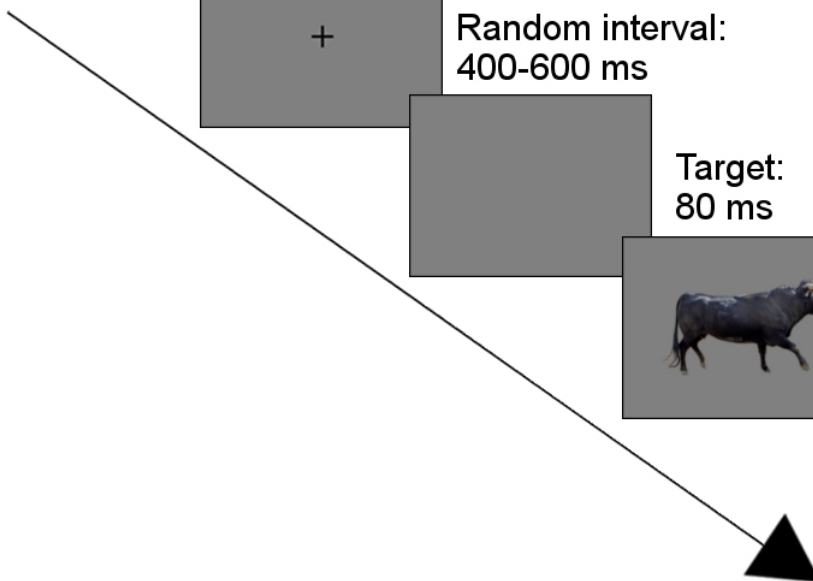


Figure 2

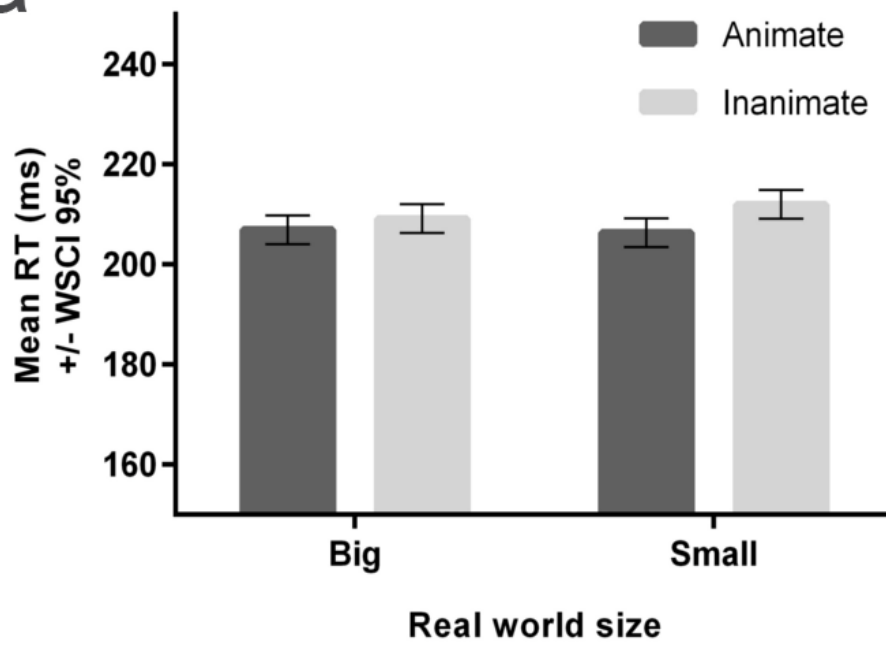
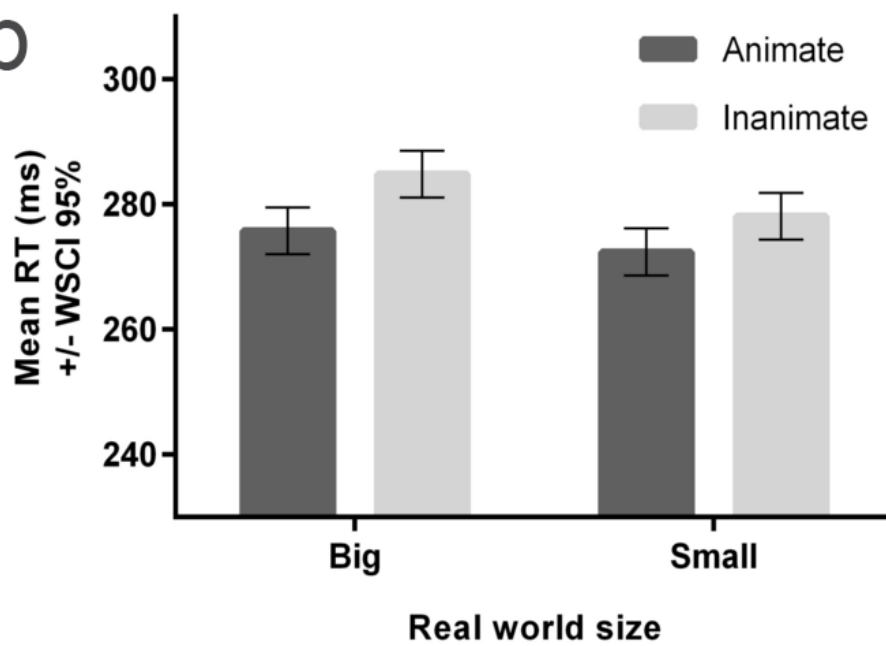
**a****b**

Figure 3

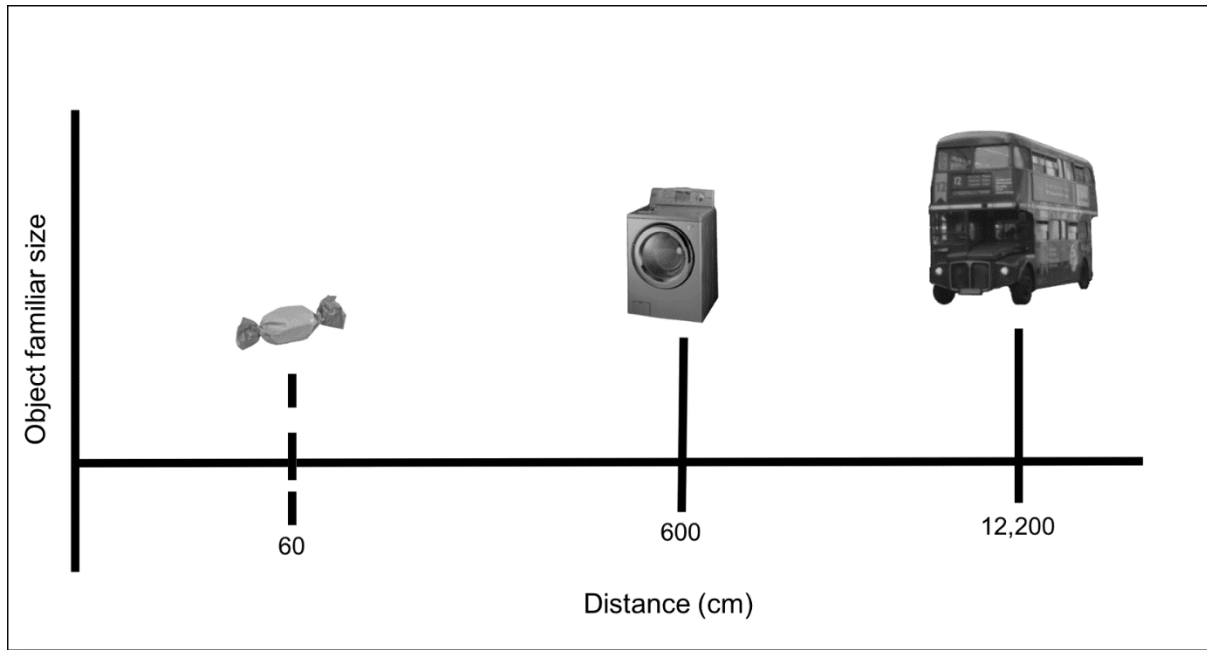




Figure 4

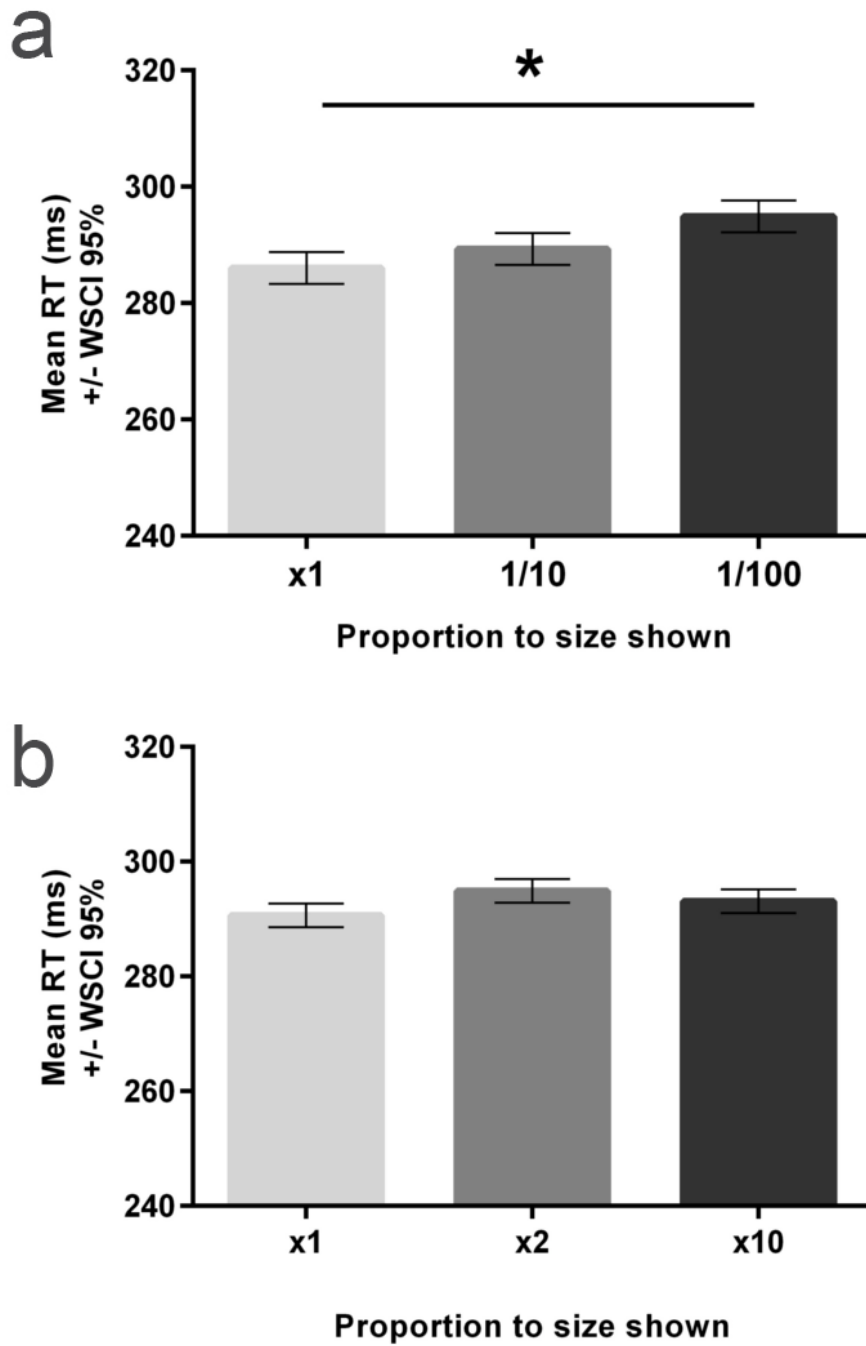


Figure 5

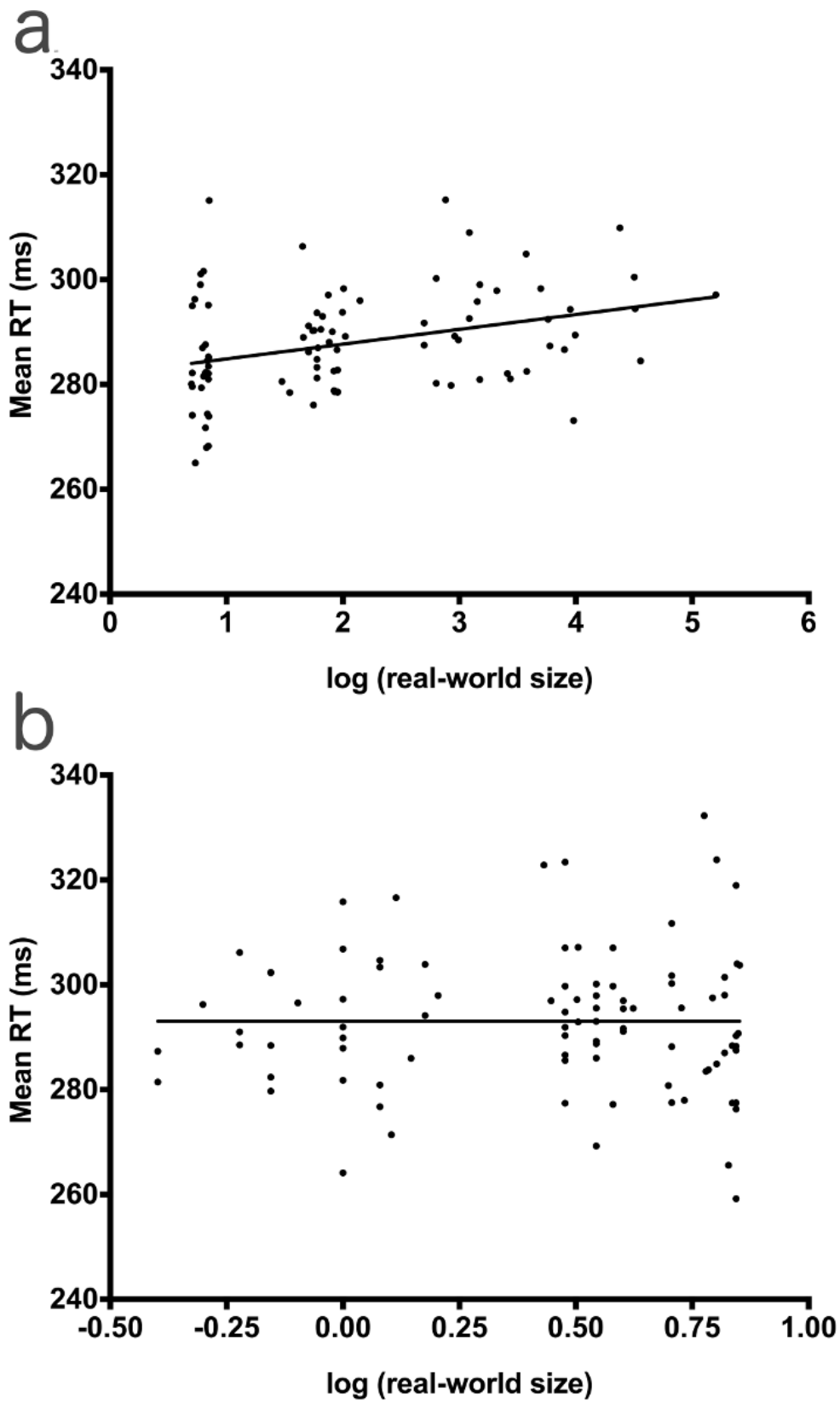


Figure 6

