Internal representations of the canonical real-world distance of objects

Yijin Wang*
Center for the Study of Applied Psychology,
Guangdong Key Laboratory of Mental Health and
Cognitive Science, and the School of Psychology,
South China Normal University, Guangzhou, China

Jie Gao*
Center for the Study of Applied Psychology,
Guangdong Key Laboratory of Mental Health and
Cognitive Science, and the School of Psychology,
South China Normal University, Guangzhou, China

Fuying Zhu
Center for the Study of Applied Psychology,
Guangdong Key Laboratory of Mental Health and
Cognitive Science, and the School of Psychology,
South China Normal University, Guangzhou, China

Xiaoli Liu
Center for the Study of Applied Psychology,
Guangdong Key Laboratory of Mental Health and
Cognitive Science, and the School of Psychology,
South China Normal University, Guangzhou, China

Gexiu Wang
Center for the Study of Applied Psychology,
Guangdong Key Laboratory of Mental Health and
Cognitive Science, and the School of Psychology,
South China Normal University, Guangzhou, China

Yichong Zhang
Center for the Study of Applied Psychology,
Guangdong Key Laboratory of Mental Health and
Cognitive Science, and the School of Psychology,
South China Normal University, Guangzhou, China

Zhiqing Deng
Center for the Study of Applied Psychology,
Guangdong Key Laboratory of Mental Health and
Cognitive Science, and the School of Psychology,
South China Normal University, Guangzhou, China

Juan Chen
Center for the Study of Applied Psychology,
Guangdong Key Laboratory of Mental Health and
Cognitive Science, and the School of Psychology,
South China Normal University, Guangzhou, China
Key Laboratory of Brain,
Cognition and Education Sciences,
Ministry of Education, South China Normal University,
Guangzhou, China

Citation: Wang, Y., Gao, J., Zhu, F., Liu, X., Wang, G., Zhang, Y., Deng, Z., & Chen, J. (2024). Internal representations of the canonical real-world distance of objects. *Journal of Vision*, *24*(2):14, 1–14, https://doi.org/10.1167/jov.24.2.14.

https://doi.org/10.1167/jov.24.2.14 Received May 4, 2023; published February 27, 2024 ISSN 1534-7362 Copyright 2024 The Authors

In the real world, every object has its canonical distance from observers. For example, airplanes are usually far away from us, whereas eyeglasses are close to us. Do we have an internal representation of the canonical real-world distance of objects in our cognitive system? If we do, does the canonical distance influence the perceived size of an object? Here, we conducted two experiments to address these questions. In Experiment 1, we first asked participants to rate the canonical distance of objects. Participants gave consistent ratings to each object. Then, pairs of object images were presented one by one in a trial, and participants were asked to rate the distance of the second object (i.e., a priming paradigm). We found that the rating of the perceived distance of the target object was modulated by the canonical real-world distance of the prime. In Experiment 2, participants were asked to judge the perceived size of canonically near or far objects that were presented at the converging end (i.e., far location) or the opening end (i.e., near location) of a background image with converging lines. We found that regardless of the presentation location, participants perceived the canonically near object as smaller than the canonically far object even though their retinal and real-world sizes were matched. In all, our results suggest that we have an internal representation of the canonical real-world distance of objects, which affects the perceived distance of subsequent objects and the perceived size of the objects themselves.

Introduction

For simplicity, in the laboratory, we usually present images of a single object on the screen to study object recognition. As a result, we are able to present objects that are much farther and larger than the computer screen in the real world onto the screen in front of us. This makes it easier to study how object recognition is affected by the physical properties of the image of an object projected onto the retina, such as retinal size and luminance. However, the visual processing in the real world may be different (Chainay & Humphreys, 2001; Freud et al., 2018; Marini, Breeding, & Snow, 2019; Snow & Culham, 2021), and the perception of object images and the representations of them in the cortex may be deeply affected by the real-world properties of the object (Gerhard, Culham, & Schwarzer, 2016, 2021; Mustafar, De Luna, & Rainer, 2015; Snow et al., 2011). For example, a picture of the sun resulted in a pupillary constriction even though the luminance of the sun picture had no difference from the luminance of the background (Binda, Pereverzeva, & Murray, 2013). This suggests that the pupil size was automatically affected by the real-world luminance of the object.

Another group of evidence that highlights the role of real-world properties in object recognition is provided by Konkle and colleagues (Konkle, 2011; Konkle & Oliva, 2011, Konkle & Oliva, 2012a, Konkle & Oliva, 2012b; Long & Konkle, 2017; Long, Konkle, Cohen, & Alvarez, 2016; Long, Moher, Carey, & Konkle, 2019; Long, Yu, & Konkle, 2018). They showed that real-world size information was extracted automatically and affected the organization of the representation of object categories in the occipitotemporal cortex. For example, in a behavioral study, they adopted a Stroop-like paradigm in which the visual sizes of two objects were either congruent or incongruent with their real-world size (Konkle & Oliva, 2012a). They found that the reaction time was longer in incongruent conditions than in congruent conditions when participants were required to determine which object on the screen was visually bigger or smaller. In other words, real-world size is an automatic property of object representation.

In the real world, every object also has its canonical distance from observers. For example, airplanes are usually far away from us, whereas eyeglasses are close to us. Do we have an internal representation of the canonical real-world distance of objects in our cognitive system similar to the representation of real-world size? Here, the canonical real-world distance of an object refers to its typical distance from the observer in the real world and may vary with the individual observer. It is related to the spatial dimension of psychological distance but is unrelated to the temporal dimension (the present and the future) and social dimension (between the observer and other people) of psychological distance (Fiedler, Jung, Wänke, & Alexopoulos, 2012; Trope & Liberman, 2010). We are interested in whether or not canonical real-world distance is an important concept and the visual property of objects that influence the perception and recognition of objects.

Previous studies showed that the processing of pictures and words depends on whether they represent proximal or distal distances. Specifically, in one experiment, they presented one picture/word of a dog/chair either near the observer or in the distance and asked participants to classify the stimulus but ignore the physical distance. They found that the responses were faster to near objects than to far objects for pictures but were faster to far objects than to near objects for words, suggesting that distance information plays a role in modulating the speed of response (Amit, Algom, & Trope, 2009). Related effects of distance were also observed for memory (Amit et al., 2019).

Brain imaging studies also reported the representation of perceived distance in the parahippocampal place area (PPA), the transverse occipital sulcus (TOS) and the lateral occipital (LO) along the ventral visual stream (Amit, Mehoudar, Trope, & Yovel, 2012), and the retrosplenial complex (RSC) and the occipital place area (OPA) (Persichetti & Dilks, 2016). For real physical distances, Gallivan, Cavina-Pratesi, and Culham (2009) presented graspable objects at either reachable or unreachable locations and asked participants to reach or grasp the objects. They found that the superior parieto-occipital cortex (SPOC) was more activated for targets within reach than beyond.

Although all these studies suggest that there is a representation of distance, the distance information was induced by the background context (Amit et al., 2009; Amit et al., 2012; Amit et al., 2019; Persichetti & Dilks, 2016) or during hand actions (Gallivan et al., 2009). Therefore, it is still unclear whether or not we have an internal representation of the canonical real-world distance of objects themselves in our cognitive system, even when the object was presented in isolation without any background context. In other words, if we present an object in isolation, do we automatically extract the distance information? This is our first research question. If the canonical real-world distance of objects is represented automatically even when an object is presented in isolation without any background context providing distance cues, then it would suggest that the canonical real-world distance is inherent to our representation of these objects, similar to the real-world size of objects.

Distance not only provides information about how far away an object is located but also enables size constancy (Chen, Sperandio, & Goodale, 2018; Chen, Sperandio, Henry, & Goodale, 2019; Holway & Boring, 1941; Sperandio & Chouinard, 2015; Sperandio, Kaderali, Chouinard, Frey, & Goodale, 2013). When a car is moving away from us, the image it projects on our retina gets smaller and smaller, but we perceive the car of constant size, which demonstrates how we compensate for the decrease in retinal image size by the increase in viewing distances to compute the real-world size of an object (Holway & Boring, 1941; Sperandio & Chouinard, 2015). Our second research question is if we do have an internal automatic representation of the canonical real-world distance of an object, whether or not the canonical distance also influences the perception of object size.

In this study, we conducted two behavioral experiments to address these questions. In both experiments, we first asked participants to rate the canonical distance of objects to obtain the canonical real-world distance of each object for each participant. In Experiment 1, a priming paradigm was used. Specifically, pairs of object images were presented one by one on a trial, and participants were asked to rate the distance of the second image (i.e., target). We tested whether the distance of the target object would be affected by the canonical real-world distance of the first one (i.e., prime). In Experiment 2, objects were presented on a picture with converging lines that provide pictorial distance cues (i.e., Ponzo illusion). The converging end indicates a far location, whereas the opening end indicates a near location. Canonically near or far objects could be presented at the near or far locations in the background. Participants were asked to rate the size of each object at each location. We investigated whether or not the canonical distance would cooperate with the pictorial cues provided by the converging lines to affect the perceived size of objects.

General methods

Participants

Thirty-nine college students (ages ranging from 18 to 23, $M = 20.74$, $SD = 1.86$; 11 males and 28 females) participated in Experiment 1. Forty-three college students (ages ranging from 18 to 27, $M = 20.86$, $SD = 3.85$; 38 females and 5 males) participated in Experiment 2. All participants had normal or corrected-to-normal vision. They were naive to the purpose of the experiments, and all gave written informed consent. The study was approved by the Human Research Ethics Board at South China Normal University, and the methods were in accordance with the guidelines established in the Declaration of Helsinki.

Apparatus

The stimulus was presented to participants on an LCD monitor (Hewlett Packard; resolution, 1,920 × 1,020) using PsychToolbox 3 (Brainard, 1997; Pelli, 1997) embedded in MATLAB (The MathWorks, Natick, MA, USA; https://ww2.mathworks.cn/). The viewing distance was 57 cm.

Experiment 1

In Experiment 1, we adopted a priming paradigm to investigate whether there is a representation of the canonical distance of objects themselves. If we have the representation of canonical distance, then the perception of distance and the reaction time for distance judgment of the target image would be affected by the canonical distance of the prime image.

Stimuli

Ten images were matched in their retinal and real-world size. There were five sizes that made to 50 color images of real-world objects in total (Figure 1A). The images of objects were selected on websites, especially online shopping platforms. The size of the

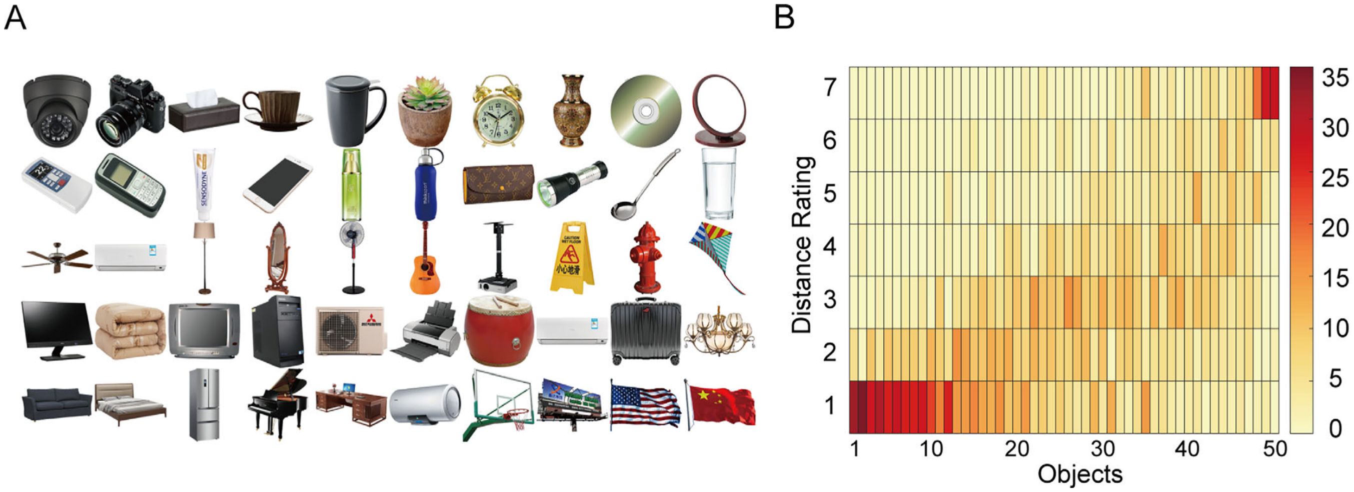

Figure 1. Stimuli and canonical distance rating results of Experiment 1. (**A**) The 50 real-world objects that were used in Experiment 1. Each row shows the group of 10 objects that were matched in real-world size and shape but varied in canonical distance. The prime and target on each trial were selected from each row. (**B**) The rating results. Participants were required to rate the canonical distance of each object from 1 to 7. On the horizontal axis, stimuli were ordered according to the rating scores averaged across all participants (note: not the same order as A). The vertical axis is the rating of distance. The color bar shows the number of people who rated specific scores for each object.

images was 156 × 156 pixels (i.e., 1.3° × 1.3°). Following Konkle and Oliva (2011), for each image of an object, the real-world size of the object was measured as the diagonal of the three-dimensional (3D) bounding box (height × width × depth) of a corresponding real object, or the dimensions were found on the Internet.

Design and procedure

Before the experiment, participants were asked to rate the canonical distance of the 50 objects, with “1” representing the nearest and “7” representing the farthest, and this allowed us to determine the canonical distance of the objects a posteriori for each participant. Canonical distance refers to the distance an object is usually from us in the real world. For example, a kite is farther than a pair of glasses, and a plane is farther than a car from us. Participants were also asked to ensure that the scores should not be influenced by their temporal or social distance (Amit et al., 2009; Fiedler et al., 2012; Trope & Liberman, 2010). For example, an old-fashioned Nokia cellphone should be rated the same as an Apple cellphone.

The experiment used a priming paradigm (Figure2A). Each trial began with a blank screen with a fixation point of 300 ms. Then, the prime was presented in the center of the screen for 300 ms, followed by a target for 300 ms. On each trial, the prime and target were selected from the 10 images that were matched on their real-world size. Our previous research shows that global shape (elongated or stubby) is an important feature of objects (Chen, Snow, Culham, & Goodale, 2018). Although it is unclear whether or not the perceived size is affected by the global shape of objects, we also matched the global shape of the prime and target just to be safe.

Participants had two tasks. First, they were asked if the target was near or far in the real world by pressing Key 1 or Key 2 (counterbalanced across participants). Their reaction time, but not their specific answer, was used in the data analysis. Then, they were asked to rate the canonical distance of the target from 1 to 7 without a time limit for a second time. This rating was compared with the rating before the experiment to measure the influence of the prime on the distance perception of the target. The trial would not proceed to the next one until the participants’ responses were recorded.

There were 450 trials in total. These trials were separated into 10 blocks. The order of trials was counterbalanced across participants. It took about 25 min to finish the whole experiment.

Data analysis

First, to evaluate whether or not participants have a consistent perception of canonical distance, the interrater consistency of the rating of canonical distance before the priming experiment was evaluated using Cronbach’s alpha (Cronbach, 1951), Krippendorf’s alpha, and intraclass correlation coefficient (ICC). Cronbach’s alpha was commonly used to evaluate the internal consistency across items or raters of a survey

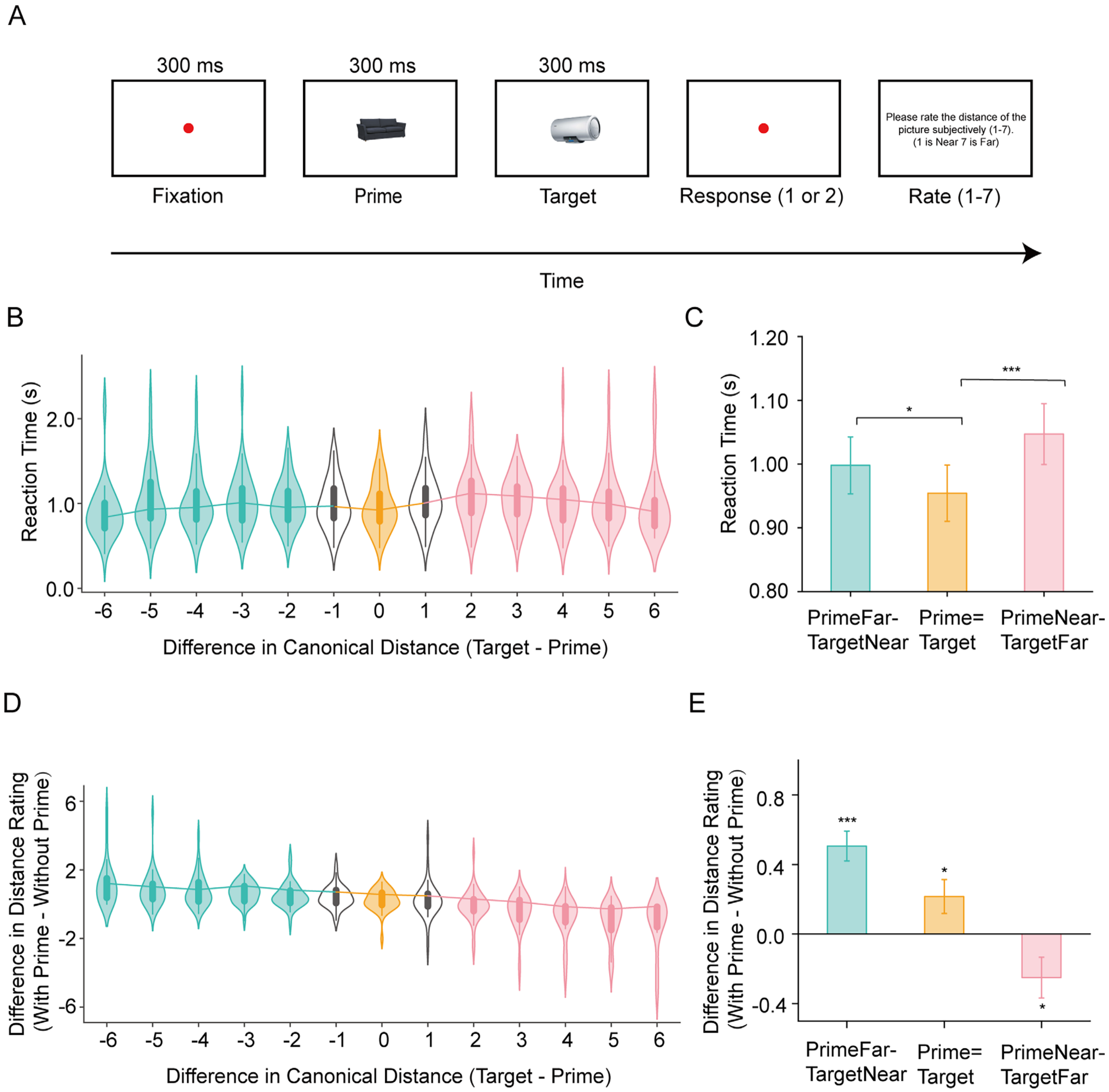

Figure 2. Protocol and results of Experiment 1. (**A**) Each trial began with a blank screen with a fixation point for 300 ms. Then the prime was presented in the center of the screen for 300 ms, followed by a target for 300 ms. Participants were asked to judge whether the target object was near or far from us in the real world by pressing Key 1 or 2 as soon as possible. After the key press, an instruction appeared on the screen, and participants were asked to rate the distance of the target following the instruction. (**B**) Distribution of the reaction time for all trials. The horizontal axis shows the difference in canonical distance rating between the prime and the target. The vertical axis shows the reaction time of Key press. (**C**) Reaction time results when the difference in the canonical distance between the target and the prime was smaller than –1 (i.e., PrimeFar-TargetNear, corresponds to green violins in panel B), was 0 (i.e., Prime=Target, corresponds to yellow violins in panel B), and was larger than 1 (i.e., PrimeNear-TargetFar, corresponds to pink violins in panel B). (**D**, **E**) Similar to B and C, respectively, but showing the results of distance rating (i.e., the difference in distance rating of the target with and without priming). The error bars show $1 \pm SEM$. $^{*}p_{\mathrm{corr}} < 0.05$, $^{**}p_{\mathrm{corr}} < 0.01$, $^{***}p_{\mathrm{corr}} < 0.001$. *P* values were corrected with Holm correction.

(Eser & Aksu, 2022). A Cronbach's alpha of more than 0.5 is usually considered acceptable, and 0.7 or more is considered good. Here, we used it to evaluate the consistency of rating across participants for all the stimuli (DeVellis, 2005). It should be noted that there are numerous indices of interrater reliability, and experts disagree on which ones are legitimate or more appropriate. Two recent research articles compared the different indices of interrater reliability (Eser & Aksu, 2022; Zhao, Feng, Ao, & Liu, 2022), both of which suggest that Krippendorf's alpha underestimated the reliability. Therefore, we reported all three indices to provide a comprehensive overview of the interrater consistency.

Then, to test how the relationship between the canonical distance of the prime and target would affect the reaction time and perceived distance of the target, for each trial, we first calculated the difference in canonical distance between the prime and the target. Then, the distribution of reaction time and distance rating was calculated for each difference. We used R (R Core Team, 2013) and lme4 (Bates, Mächler, Bolker, & Walker, 2015) to perform linear mixed-modeling (LMM) analyses. Specifically, an LMM with the subject and item as random factors and the absolute rating difference in canonical distance between the prime and target as a fixed factor (a continuous variable) was performed to examine the effect of canonical distance on the reaction time of the target. The linear mixed model was constructed as follows in R:

$$\text{Reaction Time} \sim \text{Absolute Difference in canonical distance (Target} - \text{Prime)} + (1 \mid \text{Subject}) + (1 \mid \text{Item})$$

Although the rating results were not perfectly normally distributed, previous literature suggested that LMM is quite robust to deviations from normality (Knief & Forstmeier, 2021; Schielzeth et al., 2020). Therefore, we also performed LMM on the rating results. Because preliminary analysis showed that the direction of difference (i.e., whether the prime was nearer or farther than the target) matters for the rating results, the original rating difference ranging from –6 to +6 was used as a fixed factor for the LMM analysis of the distance rating results. The linear mixed model was constructed as follows in R:

$$\text{Difference in rating (post-pre)} \sim \text{Difference in canonical distance (Target-Prime)} + (1|\text{Subject}) + (1|\text{Item})$$

In addition to LMM, we also performed analysis of variance (ANOVA) to confirm the main findings. Specifically, the trials were divided into three conditions for analysis: PrimeFar-TargetNear, Prime=Target, and PrimeNear-TargetFar. The "Near" and "Far" were a relative relationship between images. If the difference in the distance rating of two images was 0, we considered them the same, and the trial would be labeled "Prime=Target." Otherwise, they would be labeled as either "PrimeNear-TargetFar" or "PrimeFar-TargetNear." The dependent variables were the reaction time when participants made distance judgments and the difference in the distance rating of the target without prime (i.e., the rating before the experiment) and with prime. Trials in which the reaction time was out of 3 standard deviations were excluded.

Repeated ANOVAs were done to reveal the main effect of the prime–target relationship. Greenhouse–Geisser method was used to correct violations of sphericity. Post hoc paired *t* tests were performed to reveal the difference in reaction time between pairs of conditions. A one-sample *t* test was also conducted to test whether the change of the rating of distance after priming was different from 0. Holm correction was applied for multiple comparison corrections whenever needed. All the ANOVAs and interrater consistency indices in this study were calculated with JSAP (Love et al., 2019; Rizopoulos, 2007).

Results

Before the priming experiment, we explicitly asked participants to rate the canonical real-world distance of objects from 1 to 7. Figure 1B shows the rating results of the canonical distance of the 50 images. The stimuli on the horizontal axis were ranked according to their mean distance rating. The color bar shows the number of participants who rated specific scores for each object. The three interrater consistency indices, Cronbach's alpha, Krippendorf's alpha, and ICC, were 0.91, 0.38, and 0.452, respectively. Importantly, there were some images where ratings were notably consistent, although the ratings for middle distances varies. For example, 80.3% of participants rated "1" for the nearest three images, and 63.2% rated "7" for the farthest three images. All these results suggest that there is a consistent representation of canonical distances for objects across participants.

In the priming experiment, participants were asked to judge the distance and then rate the distance of the target but not the prime. Therefore, the priming experiment examined whether the canonical real-world distance of the prime would implicitly influence the distance perception of the target. Note again that the retinal size and the real-world size of the prime and target were matched.

First, we calculated the difference in canonical distance between the prime and target and showed the distribution of the reaction time (Figure 2B) and the distribution of the difference in canonical distance rating with and without the prime (Figure 2D) by the difference in canonical distance between the prime and target of all trials. Because Figure 2B showed a symmetrical relationship between the effects of negative and positive differences on reaction time, we used the *absolute* difference in canonical distance as an independent variable to examine the relationship between the absolute difference in canonical distance and the reaction time to the distance of the target. An LMM showed that the main effect of the absolute difference in canonical distance between target and prime on reaction time was significant (estimate $= 0.005$, $SE = 0.002$, $t = 2.479$, $p = 0.013$). The R^2, which reflected the effect size, was 0.33. (The estimate looks small because we used "second" as the unit of reaction time in our analysis. If we use milliseconds as the unit, the beta becomes 5.134.) The LMM result suggests that an increase of 1 unit in difference in canonical distance (Prime – Target) between prime and target decreased the reaction time of the target by 0.005 seconds (i.e., 5 ms).

However, because reaction time (RT) data are not distributed normally, as considered in many studies (De Heering, Collignon, & Kolinsky, 2018; De Heering & Kolinsky, 2019; Kolinsky & Fernandes, 2014; Pegado et al., 2014), we also transformed the RT data into log-transforms following the Box–Cox transformation procedure (Box & Cox, 1964). Results after transformation also showed that the main effect of absolute difference in ratings was significant (estimate $= 0.006$, $SE = 0.002$, $t = 3.416$, $p < 0.001$). The R^2, which reflected the effect size, was 0.40.

In fact, RTs seemed to be the lowest when prime and target were most different (Figure 2B). This could suggest that the visual system simply discards the prime for large differences in canonical distance. When the difference in distance is moderate, the prime is taken into account and we see a priming effect (i.e., faster response).

We also performed the classic ANOVA to examine further the effect of the difference in canonical distance between the prime and target on the reaction time data. To this end, trials were grouped into three categories for each individual based on their rating results (see Methods for details). We analyzed whether the reaction time for the distance report was affected by the relationship between the distances of the prime and the target (i.e., PrimeFar-TargetNear, Prime=Target, and PrimeNear-TargetFar; Figure 2C). Results of repeated-measures ANOVA showed that the main effect of the distance relationship was significant, $F(2, 76) = 14.313$, $p < 0.001$, $\eta^2 = 0.274$. Post hoc comparisons (Figure 2C) with Holm corrections revealed that the difference in reaction time between PrimeFar-TargetNear and Prime=Target trials was significant ($t(38) = 2.341$, $p_{\text{corr}} = 0.022$), and so was the difference in reaction time between PrimeNear-TargetFar and Prime=Target trials ($t(38) = 5.337$, $p_{\text{corr}} < 0.001$). Generally speaking, the reaction time was faster when the distance of the prime and the target was similar than when they were dissimilar (note, however, the reaction time was actually the shortest when they were most dissimilar; see Figure 2B), which is in line with the typical priming effect and suggests that the relationship between the canonical distances of the prime and target object does affect the speed of the distance judgment of the target.

Similar to reaction time, we also performed LMM on the results of distance rating to examine whether the difference in the canonical distance between the prime and the target would affect the perceived distance of the target. A positive difference in canonical distance rating indicates that participants perceived the object farther away after priming than it was without priming, whereas a negative difference indicates that participants perceived the object nearer after priming than it was without priming.

LMM showed that the effect of the difference in canonical distance was significant (estimate $= -0.189$, $t = -40.203$, $p < 0.001$), suggesting a significant effect of the prime on the distance rating of the target. The R^2, which reflected the effect size, was 0.24. The LMM result suggests that an increase of 1 unit in difference in canonical distance (Prime – Target) between prime and target decreased the distance rating of the target by 0.189.

A repeated-measures ANOVA showed that the main effect of the distance relationship is significant, $F(2, 38) = 35.425$, $p < 0.001$, $\eta^2 = 0.482$. Posthoc test showed that compared to Prime=Target, the far prime made the near target be perceived farther than it was without priming (PrimeFar-TargetNear vs. Prime=Target: $t(38) = 3.198$, $p_{\text{corr}} = 0.006$). In contrast, the near prime made the far target be perceived nearer than it was without priming (PrimeNear-TargetFar vs. Prime=Target: $t(38) = -5.144$, $p_{\text{corr}} < 0.001$). We performed the one-sample t test in the next step. The results again showed that the far prime made the near target be perceived farther than it was without priming (PrimeFar-TargetNear: $t(38) = 5.922$, $p_{\text{corr}} < 0.001$) (Figure 2E). In contrast, the near prime made the far target be perceived nearer than it was without priming (PrimeNear-TargetFar: $t(38) = -2.135$, $p_{\text{corr}} = 0.039$). These findings show that the prime pulled the target to the canonical distance of the prime itself, suggesting that the perceived distance of objects can be implicitly influenced by the canonical distance of the object presented before. Surprisingly, even when the canonical distance of the prime and target was the same, the target was also perceived farther with the prime than it was without the prime (Prime=Target; $t(38) = 2.216$, $p_{\text{corr}} = 0.033$). We explore this in the Discussion.

Experiment 2

In Experiment 1, we showed that participants gave consistent distance ratings of objects even though the objects were just images presented on screen. In addition, the perceived distance of a target object was significantly affected by the canonical real-world distance of the object presented before, which suggests the existence of an implicit representation of the canonical distance of an object even when the object was a picture on the screen. In this experiment, we tested whether or not the canonical distance would also affect the perceived size of objects.

To this end, we selected pairs of objects that had different canonical real-world distances but matched in real-world size and global shape (elongated or stubby) as stimuli. These stimuli were presented on a background with converging lines, which provided pictorial distance cues and made people perceive the object at the converging end as larger than the same object presented at the opening end, generating a size illusion called the Ponzo illusion (Fisher, 1968; Sperandio & Chouinard, 2015; Yildiz, Sperandio, Kettle, & Chouinard, 2022). Participants were asked to rate the perceived size of the object. With this background, we expected to see whether or not there was an effect of canonical real-world distance on perceived size and whether the canonical real-world distance and pictorial distance cues exert consistent influence on size perception.

Stimuli

Because the purpose of the experiment was to examine the effect of canonical distance on size perception, the experimenter did preliminary selection and matching on objects so that each object had a paired object that was matched on the real-world size (see Methods in Experiment 1) and global shape but was canonically nearer or canonically farther than it. Thirty-four color images of real-world objects were selected (Figure 3A). (Note that the data analysis was based on the distance rating of each participant. If a participant rated the pair of images with the same canonical distance, the pair would be excluded from analysis.)

The images of objects were selected on websites, especially online shopping websites. The size of the stimulus was 156 × 156 pixels (i.e., 1.3° × 1.3°). Images of objects were presented on a background with converging lines to provide distance cues.

Design and procedure

Again, before the experiment, participants were asked to fill in a questionnaire to rate the canonical distance of 34 objects, with "1" indicating the nearest and a rank of "7" indicating the farthest. This allowed us to determine the canonical distance of the objects for each participant. This rating result was then used in the analysis. Note that the experimenter selected pairs of images for presentation, but the analysis was based on the rating of the participants.

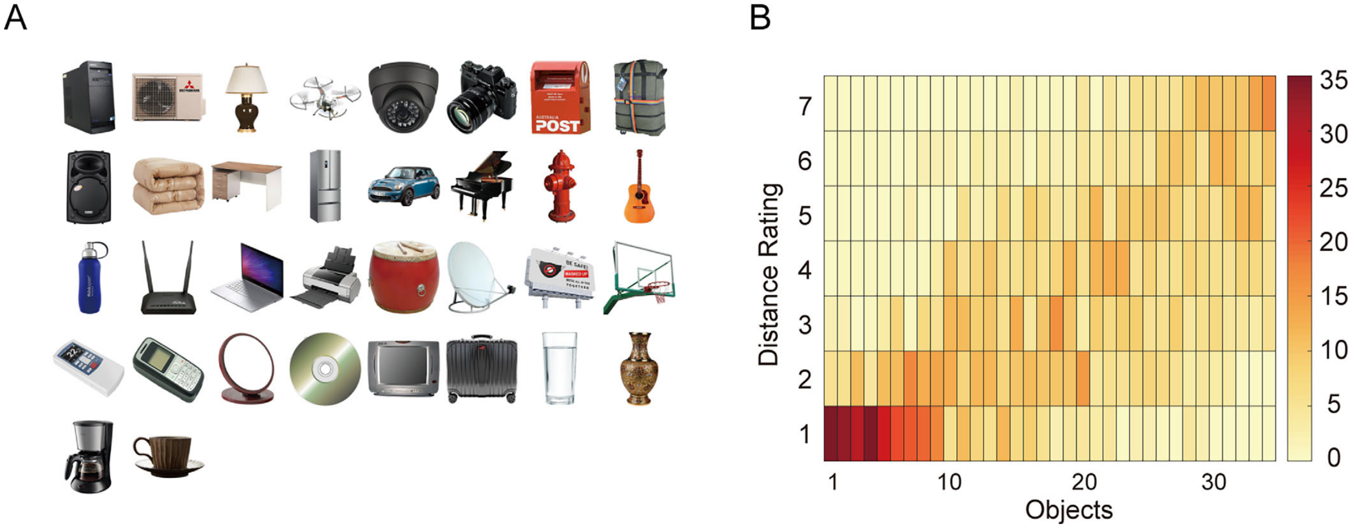

Figure 3. Stimuli and distance rating results of Experiment 2. (**A**) Thirty-four real-world objects were used as stimuli. (**B**) Participants were required to rate all objects from 1 to 7 based on their canonical distance. Stimuli on the horizontal axis were ranked according to their mean value. The vertical axis is the rating of canonical distance. The color bar shows the number of people who give a specific score of rating.

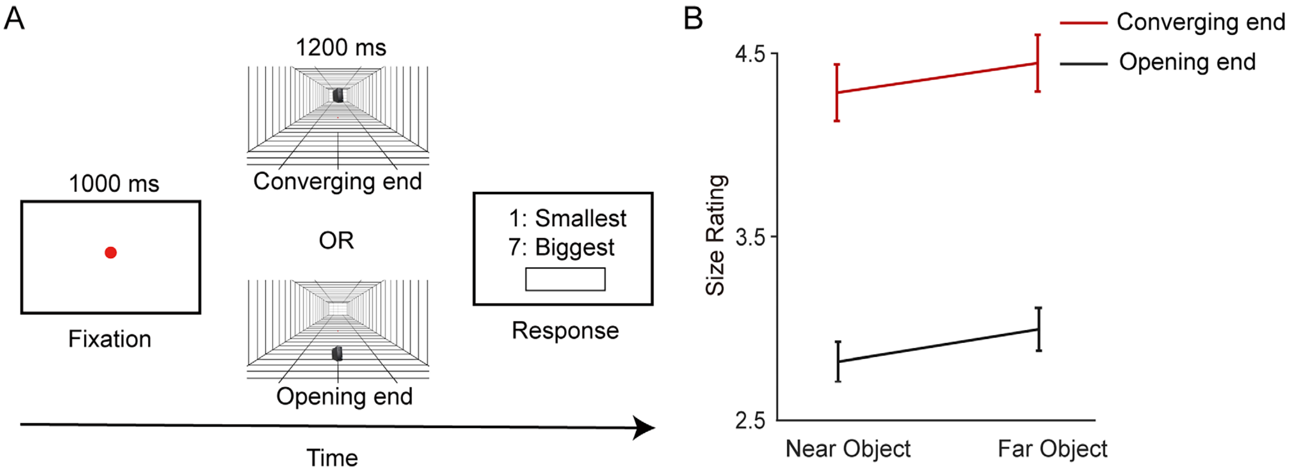

Figure 4. Protocol and Results of Experiment 2. (**A**) Each trial began with a blank screen with a fixation point for 1,000 ms. Then an object was presented at either the converging end or the opening end for 1,200 ms. Participants were asked to rate the perceived size of the object from 1 to 7 with no time limitation. The object could be presented at the converging end or the opening end. (**B**) Results of size ratings for canonically near and far objects presented at the far location (i.e., the converging end) and the near location (i.e., the opening end) on the background with converging lines. The red color indicates near or far objects presented at the converging end, while the black one indicates near or far objects presented at the opening end. The error bars show 1 ± *SEM*.

During the experiment, each trial began with a blank screen with a fixation point on the center of the screen for 1,000 ms. Then, an object was presented either at the converging end or the opening end of the converging lines on the background image for 1,200 ms (Figure 4A). The converging lines provided distance cues, with the converging end indicating far distance and the opening end indicating near distance. Each object could be presented either on the converging end or the opening end of the background image. Participants were required to rate the size of the object from 1 to 7, with "1" representing the smallest and "7" representing the biggest. It should be noted that although the images were paired in analysis, the rating was performed for each image one by one. No time limitation was imposed through this protocol.

Data analysis

Again, Cronbach's alpha, Krippendorf's alpha, and ICC were used to evaluate whether or not participants have a consistent perception of canonical distance.

To compare the perceived size of canonically near and canonically far objects in the background, pairs of objects matched on their real-world size (see Methods in Experiment 1) and their global shape (elongated or stubby) but different in canonical distance based on their own rating before the experiment were put into analyses. Again, both LMM and ANOVA were performed.

First, we performed an LMM, with subject and item as random factors, and the canonical distance (canonically near or far) and presentation location (converging end vs. opening end) as fixed factors with comprehensive considerations of their interactions, to analyze the size rating data. The model specification in R is as follows:

$$\text{Size Rating} \sim \text{Presentation location} * \text{ Canonical Distance} + (1 \mid \text{Subject}) + (1 \mid \text{Item})$$

Second, for ANOVA, trials were grouped into four conditions: canonically near or far objects presented at the converging or the opening end for each participant. Repeated-measures ANOVA with canonical distance (near vs. far) and presentation location (converging vs. opening ends) as within-subject factors were carried out to reveal the main effects and interactions. Typically, the same objects presented at the converging end (i.e., far location) would be perceived as larger than those in the open end (i.e., near location), which is called the Ponzo illusion. Here, we focused on the impact of the canonical distance (near vs. far) of the object on the perceived size. A significant main effect of canonical distance or significant interactions between canonical distance and Presentation location would suggest that the difference in the perceived size is modulated by the canonical distance of the object.

Results

First, we analyzed each participant's rating data of each object (Figure 3B). The color bar shows the number of participants who give a specific rating for each object. The three interrater consistency indices, Cronbach's Alpha, Krippendorf's alpha, and ICC, were 0.8, 0.291, and 0.313, respectively. Again, there are some images where ratings are very consistent (very near and very far). For example, 70.5% of participants rated "1" for the nearest three images, and 41.9% rated "7" for the farthest three images. All these results suggest that participants gave consistent ratings to each object, which is consistent with the results of Experiment 1.

LMM with both the item and subject as random factors showed that the effect of canonical distance (estimate $= 0.078$, $t = 1.136$, $p = 0.256$) was not significant, but the presentation location (estimate $= 1.450$, $t = 25.545$, $p < 0.001$) was significant. The interaction between canonical distance and presentation location was not significant either (estimate $= -0.020$, $t = -0.245$, $p = 0.807$). However, the LMM with only the subject as a random factor revealed a significant effect of both canonical distance (estimate $= 0.183$, $t = 2.821$, $p = 0.005$) and presentation location (estimate $= 1.450$, $t = 22.731$, $p < 0.001$), suggesting that effect of canonical distance on perceived size may rely on the items to some extent.

Repeated-measures ANOVA with canonical distance (near vs. far) and presentation location (converging vs. opening ends) also showed that the main effects of canonical distance ($F(1, 42) = 19.384$, $p < 0.001$, $\eta^2 = 0.009$) and presentation location ($F(1, 42) = 104.914$, $p < 0.001$, $\eta^2 = 0.688$) were significant. The perceived size was larger at the converging end than at the opening end, which is in line with the Ponzo illusion (Figure 4B). The canonically near object was perceived smaller than the canonically far object at both converging and opening ends, even though their real-world size was matched. In addition, the interaction between canonical distance and presentation location was not significant, $F(1, 42) = 0.107$, $p = 0.746$, $\eta^2 = 1.785\text{e}^{-5}$. Therefore, the ANOVA results suggest that canonical distance modulates the perceived size of objects.

One may ask whether or not the effect of canonical distance on the perceived size of objects depends on the real-world size of the objects. To test this, we separated trials into real-world small and real-world large groups and performed ANOVA with real-world size, canonical distance, and presentation location as factors. The repeated ANOVA revealed a significant main effect of real-world size ($F(1, 42) = 281.388$, $p < 0.001$, $\eta^2 = 0.626$), but the interaction between canonical distance and real-world size was not significant ($F(1, 42) = 0.450$, $p = 0.506$, $\eta^2 = 5.98\text{e-}5$), suggesting that the effect of canonical distance on perceived size was not modulated by the real-world size of the objects.

Discussion

In the current study, we investigated whether or not there is an internal representation of canonical real-world distance for objects presented in isolation and whether the canonical distance would automatically affect the distance-related judgment of subsequent objects and the perceived size of itself. We found that participants generally gave consistent ratings of the canonical distance of objects. In addition, the perceived distance of objects can be pulled toward the canonical distance of the object presented before. Moreover, the canonical distance of objects has an impact on their size perception, with canonically near objects appearing relatively smaller compared to canonically far objects. Overall, these results suggest that canonical distance is an inherent feature of objects that automatically affects the distance perception of other objects and the size perception of itself.

It should be noted that to detangle the effect of canonical distance, the pairs of stimuli were matched in real-world size and retinal size. Therefore, our results could not be attributed to any difference in these two features. Previous behavioral studies have also investigated the representation of distance but with distance cues (Amit et al., 2009; Amit et al., 2012; Amit et al., 2019; Gallivan et al., 2009; Persichetti & Dilks, 2016; Quinlan & Culham, 2007). Here we found that even when the prime object was presented in isolation without any distance cues, the canonical distance of the prime object affected the distance perception of that of the following object, which suggests that similar to canonical size, the canonical distance is also an inherent property of object.

Unlike the common priming studies (Bravo & Nakayama, 1992; Carr, McCauley, Sperber, & Parmelee, 1982; Eimer & Schlaghecken, 1998; Tipper, 1985), not only the reaction time but also the distance rating was affected by the prime. A near prime made a far object be perceived nearer. A far prime made a near target be perceived further. In other words, the prime not only affected the reaction time but also pulled the target toward it. Therefore, our results cannot be simply considered a congruency effect. It seems that the prime object created a semantic context that modified the distance knowledge of the target (Neely, 1977; Posner & Snyder, 2004; Rosch, 1975). However, it remains unclear why the target was perceived farther than it was when the prime and target were around the same canonical distance. We speculated that this was a phenomenon related to repetition suppression. That is, when a stimulus is presented repeatedly, its activation in the brain is usually weaker than when the stimulus is presented once (Garrido et al., 2009; Valyear, Gallivan, McLean, & Culham, 2012). It has been shown that neurons selectively responded to the distance of objects

(Dobbins, Jeo, Fiser, & Allman, 1998), and overall, there was a preference for near viewing distance (Quinlan & Culham, 2007). Therefore, it is possible that repeatedly presenting objects of a certain distance makes the neurons selectively responding to this distance fatigued and consequently makes the distance perceived even farther. Anyway, the existence of the pulling effect of priming suggests that there is neural representation of canonical distance of objects.

Previous studies showed that the canonical size affected the organization of categories in the temporal-occipital cortex (Konkle & Oliva, 2012b). It is unclear whether or not the canonical distance would also affect the object representation anywhere in the cortex. Based on the previous studies on distance perception induced by various distance cues, the candidate areas could be the PPA (Amit et al., 2012), RSC (Persichetti & Dilks, 2016), LO (Amit et al., 2012), and OPA (Persichetti & Dilks, 2016) along the ventral stream and the SPOC (Gallivan et al., 2009) along the dorsal stream. Future research is needed to clarify the neural substrates of canonical distance.

An intriguing result of our study is that not only perceived distance but also perceived size was affected by the canonical distance of objects. According to Emmert's law (Emmert, 1881; Sperandio & Chouinard, 2015), the perceived size of objects depends on the retinal size and the distance of the object. The distance information usually refers to the distance cues of objects. Here we provide compelling evidence showing that the canonical real-world distance of the object itself also affects the size perception of the object, which extends the understanding of size-distance computation.

Why were canonically near objects perceived smaller than the canonically far objects? According to Emmert's law, the perceived size is scaled to the distance of objects (Emmert, 1881; Sperandio & Chouinard, 2015). Canonically near objects had shorter distances than canonically far objects. When the canonical real distance was taken into account, there is no surprise why canonically near objects were perceived as smaller even when their retinal and real-world size were matched.

Our Experiment 2 adopted a complicated design. That is, the stimuli were presented on a background with pictorial cues. With the pictorial background included, we showed that canonical distance affected size perception even when the pictorial cue was provided, and the influence of canonical distance on size perception was in the same direction as pictorial cues (the farther the larger).

Above, we showed that canonical real-world distance, as knowledge of the typical distance of objects, affects the perceived distance and size of objects. Indeed, distance is also critical information for visually guided actions, such as reaching and grasping. A previous study indicated that individuals can utilize their familiarity with an object's size to infer distance information to control reaching and grasping (Marotta & Goodale, 2001). Proprioceptive distance information also helps restore size constancy in grasping when vision is limited (Chen, Sperandio, et al., 2018). Therefore, can canonical real-world distance information be used for reaching and grasping? An apparent example is that people with average cognition ability never reach to grasp the moon. For nonextreme situations, it is still unclear whether people would reach to canonically near (e.g., a mug) and canonically far objects (i.e., a lamp bulb) that were matched in real-world size with the same reaching distance. Further research is required to test the effect of canonical distance on reaching and grasping.

We defined the canonical real-world distance of an object as its typical distance from the observer in the real world. One limitation of our study is that it is difficult to exclude the influence of social distance when judging canonical real-world distance. An airplane is canonically far for most people but may be close to a pilot who flies with it every day. A coffee machine may be reported as close to people who have access to it in everyday life compared with those who never used it. In other words, canonical distance depends on variance in everyday experiences we make with certain objects. Therefore, unlike real-world size that does not depend much on the observer, the canonical real-world distance is more likely a summary knowledge of the real-world distance of objects, which may vary depending on the individual observer.

Conclusions

Overall, our finding suggests that real-world distance is a critical feature of objects that is automatically and involuntarily activated and therefore interferes with processing real-world distance and size of objects. Future studies should take canonical distance into consideration and rule out the possible confound introduced by difference in canonical real-world distance in research on object recognition.

Keywords: object recognition, real-world distance, priming effect, ponzo illusion

Acknowledgments

Supported by two National Natural Science Foundation of China grants (No. 31970981 and No. 31800908) and the National Science and Technology Innovation 2030 Major Program (STI2030-Major Projects 2022ZD0204802) to JC.

Data availability: Our data and analyses are available at https://osf.io/8m4pn/.

Commercial relationships: none.

Corresponding authors: Zhiqing Deng and Juan Chen.
Emails: zhiqingdeng@m.scnu.edu.cn; juanchen@m.scnu.edu.cn.
Address: Center for the Study of Applied Psychology, Guangdong Key Laboratory of Mental Health and Cognitive Science, and the School of Psychology, South China Normal University, Guangzhou, China.

*YW and JG contributed equally to this work.

References

Amit, E., Algom, D., & Trope, Y. (2009). Distance-dependent processing of pictures and words. *Journal of Experimental Psychology: General, 138*(3), 400.

Amit, E., Mehoudar, E., Trope, Y., & Yovel, G. (2012). Do object-category selective regions in the ventral visual stream represent perceived distance information? *Brain and Cognition, 80*(2), 201–213.

Amit, E., Rim, S., Halbeisen, G., Priva, U. C., Stephan, E., & Trope, Y. (2019). Distance-dependent memory for pictures and words. *Journal of Memory and Language, 105*, 119–130.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48.

Binda, P., Pereverzeva, M., & Murray, S. O. (2013). Pupil constrictions to photographs of the sun. *Journal of Vision*, *13*(6), 8, https://doi.org/10.1167/13.6.8.

Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological), 26*(2), 211–243.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*(4), 433–436.

Bravo, M. J., & Nakayama, K. (1992). The role of attention in different visual-search tasks. *Perception & Psychophysics, 51*(5), 465–472.

Carr, T. H., McCauley, C., Sperber, R. D., & Parmelee, C. M. (1982). Words, pictures, and priming: On semantic activation, conscious identification, and the automaticity of information processing. *Journal of Experimental Psychology: Human Perception and Performance, 8*(6), 757–777.

Chainay, H., & Humphreys, G. W. (2001). The real-object advantage in agnosia: Evidence for a role of surface and depth information in object recognition. *Cognitive Neuropsychology, 18*(2), 175–191.

Chen, J., Snow, J. C., Culham, J. C., & Goodale, M. A. (2018). What role does "elongation" play in "tool-specific" activation and connectivity in the dorsal and ventral visual streams? *Cerebral Cortex, 28*(4), 1117–1131.

Chen, J., Sperandio, I., & Goodale, M. A. (2018). Proprioceptive distance cues restore perfect size constancy in grasping, but not perception, when vision is limited. *Current Biology, 28*(6), 927–932.e924.

Chen, J., Sperandio, I., Henry, M. J., & Goodale, M. A. (2019). Changing the real viewing distance reveals the temporal evolution of size constancy in visual cortex. *Current Biology, 29*(13), 2237–2243.e2234.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297–334.

De Heering, A., Collignon, O., & Kolinsky, R. (2018). Blind readers break mirror invariance as sighted do. *Cortex, 101*, 154–162.

De Heering, A., & Kolinsky, R. (2019). Braille readers break mirror invariance for both visual Braille and Latin letters. *Cognition, 189*, 55–59.

DeVellis, R. F. (2005). Inter-rater reliability. In K. Kempf-Leonard (Ed.), *Encyclopedia of social measurement* (pp. 317–322). New York, NY: Elsevier.

Dobbins, A. C., Jeo, R. M., Fiser, J., & Allman, J. M. (1998). Distance modulation of neural activity in the visual cortex. *Science, 281*(5376), 552–555.

Eimer, M., & Schlaghecken, F. (1998). Effects of masked stimuli on motor activation: Behavioral and electrophysiological evidence. *Journal of Experimental Psychology: Human Perception and Performance, 24*(6), 1737–1747.

Emmert, E. (1881). Grössenverhältnisse der Nachbilder. *Klinische Monatsblätter für Augenheilkunde, 19*, 443–450.

Eser, M. T., & Aksu, G. (2022). Comparison of the results of the generalizability theory with the inter-rater agreement coefficients. *International Journal of Curriculum and Instruction, 14*(2), 1629–1643.

Fiedler, K., Jung, J., Wänke, M., & Alexopoulos, T. (2012). On the relations between distinct aspects of psychological distance: An ecological basis of construal-level theory. *Journal of Experimental Social Psychology, 48*(5), 1014–1021.

Fisher, G. H. (1968). Illusions and size-constancy. *The American Journal of Psychology, 81*(1), 2–20.

Freud, E., Macdonald, S. N., Chen, J., Quinlan, D. J., Goodale, M. A., & Culham, J. C. (2018). Getting a grip on reality: Grasping movements directed to real objects and images rely on dissociable neural representations. *Cortex, 98*, 34–48.

Gallivan, J. P., Cavina-Pratesi, C., & Culham, J. C. (2009). Is that within reach? fMRI reveals that the human superior parieto-occipital cortex encodes objects reachable by the hand. *Journal of Neuroscience, 29*(14), 4381–4391.

Garrido, M. I., Kilner, J. M., Kiebel, S. J., Stephan, K. E., Baldeweg, T., & Friston, K. J. (2009). Repetition suppression and plasticity in the human brain. *NeuroImage, 48*(1), 269–279.

Gerhard, T. M., Culham, J. C., & Schwarzer, G. (2016). Distinct visual processing of real objects and pictures of those objects in 7- to 9-month-old infants. *Front Psychol, 7*, 827.

Gerhard, T. M., Culham, J. C., & Schwarzer, G. (2021). Manual exploration of objects is related to 7-month-old infants' visual preference for real objects. *Infant Behavior and Development, 62*, 101512.

Holway, A. H., & Boring, E. G. (1941). Determinants of apparent visual size with distance variant. *The American Journal of Psychology, 54*, 21–37.

Knief, U., & Forstmeier, W. (2021). Violating the normality assumption may be the lesser of two evils. *Behavior Research Methods, 53*(6), 2576–2590.

Kolinsky, R., & Fernandes, T. (2014). A cultural side effect: Learning to read interferes with identity processing of familiar objects. *Frontiers in Psychology, 5*, 1224.

Konkle, T. (2011). *The role of real-world size in object representation.* Cambridge: Massachusetts Institute of Technology.

Konkle, T., & Oliva, A. (2011). Canonical visual size for real-world objects. *Journal of Experimental Psychology: Human Perception and Performance, 37*(1), 23.

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.

Konkle, T., & Oliva, A. (2012b). A real-world size organization of object responses in occipitotemporal cortex. *Neuron, 74*(6), 1114–1124.

Long, B., & Konkle, T. (2017). A familiar-size Stroop effect in the absence of basic-level recognition. *Cognition, 168*, 234–242.

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.

Long, B., Moher, M., Carey, S., & Konkle, T. (2019). Real-world size is automatically encoded in preschoolers' object representations. *Journal of Experimental Psychology: Human Perception and Performance, 45*(7), 863.

Long, B., Yu, C.-P., & Konkle, T. (2018). Mid-level visual features underlie the high-level categorical organization of the ventral stream. *Proceedings of the National Academy of Sciences, 115*(38), E9015–E9024.

Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., Verhagen, J., . . . Wagenmakers, E.-J. (2019). JASP: Graphical statistical software for common statistical designs. *Journal of Statistical Software, 88*(1), 1–17.

Marini, F., Breeding, K. A., & Snow, J. C. (2019). Distinct visuo-motor brain dynamics for real-world objects versus planar images. *Neuroimage, 195*, 232–242.

Marotta, J. J., & Goodale, M. A. (2001). The role of familiar size in the control of grasping. *Journal of Cognitive Neuroscience, 13*(1), 8–17.

Mustafar, F., De Luna, P., & Rainer, G. (2015). Enhanced visual exploration for real objects compared to pictures during free viewing in the macaque monkey. *Behavioral Processes, 118*, 8–20.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General, 106*, 226–254.

Pegado, F., Nakamura, K., Braga, L. W., Ventura, P., Nunes Filho, G., Pallier, C., . . . Dehaene, S. (2014). Literacy breaks mirror invariance for visual stimuli: A behavioral study with adult illiterates. *Journal of Experimental Psychology: General, 143*(2), 887.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision, 10*(4), 437–442.

Persichetti, A. S., & Dilks, D. D. (2016). Perceived egocentric distance sensitivity and invariance across scene-selective cortex. *Cortex, 77*, 155–163.

Posner, M. I., & Snyder, C. R. R. (2004). *Attention and cognitive control.* New York, NY: Psychology Press.

Quinlan, D., & Culham, J. C. (2007). fMRI reveals a preference for near viewing in the human parieto-occipital cortex. *Neuroimage, 36*(1), 167–187.

R Core Team. (2013). *R: A language and environment for statistical computing*. Vienna, Austria: R Project for Statistical Computing.

Rizopoulos, D. (2007). ltm: An R package for latent variable modeling and item response analysis. *Journal of Statistical Software, 17*, 1–25.

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General, 104*, 192–233.

Schielzeth, H., Dingemanse, N. J., Nakagawa, S., Westneat, D. F., Allegue, H., Teplitsky, C., . . . Araya-Ajoy, Y. G. (2020). Robustness of linear mixed-effects models to violations of distributional assumptions. *Methods in Ecology and Evolution, 11*(9), 1141–1152.

Snow, J. C., & Culham, J. C. (2021). The treachery of images: How realism influences brain and behavior. *Trends in Cognitive Sciences, 25*(6), 506–519.

Snow, J. C., Pettypiece, C. E., McAdam, T. D., McLean, A. D., Stroman, P. W., Goodale, M. A., . . . Culham, J. C. (2011). Bringing the real world into the fMRI scanner: Repetition effects for pictures versus real objects. *Scientific Reports, 1*, 130.

Sperandio, I., & Chouinard, P. A. (2015). The mechanisms of size constancy. *Multisensory Research, 28*(3–4), 253–283.

Sperandio, I., Kaderali, S., Chouinard, P. A., Frey, J., & Goodale, M. A. (2013). Perceived size change induced by nonvisual signals in darkness: The relative contribution of vergence and proprioception. *Journal of Neuroscience, 33*(43), 16915–16923.

Tipper, S. P. (1985). The negative priming effect: Inhibitory priming by ignored objects. *Quarterly Journal of Experimental Psychology A, 37*(4), 571–590.

Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review, 117*(2), 440–463.

Valyear, K. F., Gallivan, J. P., McLean, D. A., & Culham, J. C. (2012). fMRI repetition suppression for familiar but not arbitrary actions with tools. *Journal of Neuroscience, 32*(12), 4247–4259.

Yildiz, G. Y., Sperandio, I., Kettle, C., & Chouinard, P. A. (2022). A review on various explanations of Ponzo-like illusions. *Psychonomic Bulletin & Review, 29*(2), 293–320.

Zhao, X., Feng, G. C., Ao, S. H., & Liu, P. L. (2022). Interrater reliability estimators tested against true interrater reliabilities. *BMC Medical Research Methodology, 22*(1), 232.