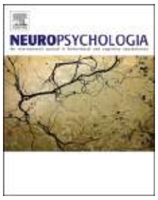
Neuropsychologia 196 (2024) 108838

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia



The contribution of semantic distance knowledge to size constancy in perception and grasping when visual cues are limited

Gexiu Wang [a], Chao Zheng [a], Xiaoqian Wu [a], Zhiqing Deng [a], Irene Sperandio [b], Melvyn A. Goodale [c], Juan Chen [a,d,*]

[a] 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, Guangdong Province, 510631, China
[b] Department of Psychology and Cognitive Science, University of Trento, Rovereto, TN, 38068, Italy
[c] Western Institute for Neuroscience and the Department of Psychology, The University of Western Ontario, London, ON, N6A 5C2, Canada
[d] Key Laboratory of Brain, Cognition and Education Sciences (South China Normal University), Ministry of Education, Guangzhou, Guangdong Province, 510631, China

ARTICLE INFO

Keywords:
Size
Distance
Size-distance scaling
Two-visual-systems hypothesis
Semantic distance knowledge
Perception
Reach-to-grasp movements

ABSTRACT

To achieve a stable perception of object size in spite of variations in viewing distance, our visual system needs to combine retinal image information and distance cues. Previous research has shown that, not only retinal cues, but also extraretinal sensory signals can provide reliable information about depth and that different neural networks (perception versus action) can exhibit preferences in the use of these different sources of information during size-distance computations. Semantic knowledge of distance, a purely cognitive signal, can also provide distance information. Do the perception and action systems show differences in their ability to use this information in calculating object size and distance? To address this question, we presented 'glow-in-the-dark' objects of different physical sizes at different real distances in a completely dark room. Participants viewed the objects monocularly through a 1-mm pinhole. They either estimated the size and distance of the objects or attempted to grasp them. Semantic knowledge was manipulated by providing an auditory cue about the actual distance of the object: "20 cm", "30 cm", and "40 cm". We found that semantic knowledge of distance contributed to some extent to size constancy operations during perceptual estimation and grasping, but size constancy was never fully restored. Importantly, the contribution of knowledge about distance to size constancy was equivalent between perception and action. Overall, our study reveals similarities and differences between the perception and action systems in the use of semantic distance knowledge and suggests that this cognitive signal is useful but not a reliable depth cue for size constancy under restricted viewing conditions.

1. Introduction

In everyday life, we rely on the integration of information from multiple sensory sources to establish a coherent and stable representation of the world. When information from one sensory system is limited, information from other sensory modalities will compensate for the loss of information. Research on multisensory integration suggests that the weighting of each sensory cue depends on the reliability of that cue (Alais and Burr, 2004; Ernst and Banks, 2002; Ernst and Bulthoff, 2004; Ohshiro et al., 2017; Rohe and Noppeney, 2018): the more reliable the cue (i.e., the lower the variance), the larger the weight associated to that cue in the multimodal integration. This principle has been widely accepted in the literature and it is believed to reflect how the brain optimally combines coherent multisensory information whilst resolving conflicted cues.

Importantly, in a recent study, we have shown that the integration of sensory signals also depends on the system that makes use of the sensory information (Chen et al., 2018). Specifically, we asked participants to view glow-in-the-dark spheres in complete darkness through a 1 mm hole so that the availability of monocular and binocular cues to distance was restricted. Participants were instructed to either manually estimate or to grasp spheres placed at different distances. Manual size estimation and grasping have been widely used before to examine the perception and action systems, respectively (Bulakowski et al., 2009; Chen et al.,

---

* Corresponding author. 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, Guangdong Province, 510631, China.
E-mail address: juanchen@m.scnu.edu.cn (J. Chen).

https://doi.org/10.1016/j.neuropsychologia.2024.108838
Received 31 July 2023; Received in revised form 4 January 2024; Accepted 21 February 2024

2015a; Chen et al., 2015b; Daprati and Gentilucci, 1997; Ganel et al., 2008; Ganel and Goodale, 2003; Hesse et al., 2012; Sun et al., 2021).

Typically, when participants view an object binocularly in a lit room, they exhibit perfect size constancy for both perception (Blakemore et al., 1972; Boring, 1940; Combe and Wexler, 2010; Holway and Boring, 1941; Sperandio and Chouinard, 2015; Sperandio et al., 2009) and action (Whitwell et al., 2020). That is, perceived size and maximum grip aperture remain constant regardless of variations in viewing distance thanks to a scaling mechanism which allows us to compensate for the decrease in retinal image size at greater viewing distances (for a review, see Sperandio and Chouinard, 2015). We found that under restricted viewing conditions, size constancy was disrupted and, as a consequence, participants' size estimations and grip apertures relied mainly on the retinal size of the object (Chen, et al., 2018). Nevertheless, when proprioceptive distance cues were provided by asking participants to hold with their non-grasping hand a pedestal upon which the sphere was placed, we observed that grasps followed size constancy rules: grip apertures were unaffected by viewing distance. In contrast, perceptual estimates of size were still mostly governed by retinal image size which decreased with an increase in viewing distance. This finding suggests that the two systems exhibit different profiles in their ability to incorporate proprioceptive distance information for the purposes of size constancy in perception and action (Kentridge, 2018).

In addition to sensory inputs, stored or acquired semantic knowledge can also provide critical information for cognition. For example, one can visually perceive the location of an object, and can also be told that an object is located at a certain distance. Unlike sensory information that can be combined together through multisensory integration during feedforward processing (Chai, et al., 2021; Foxe and Schroeder, 2005) the effect of knowledge on cognition is most likely implemented through top-down expectation and prediction (Broderick et al., 2019; Zhu et al., 2013). Whether or not the perception and action systems exhibit differences in their ability to use knowledge about distance remains an open question.

To address this issue in the present experiment, we tested participants in size estimation and grasping tasks under restricted-viewing condition, similar to our previous study (Chen, et al., 2018). To examine the contribution of distance knowledge on size constancy, we presented an auditory cue about object's distance through a speaker 2 s before participants viewed the glowing target. Size constancy depends on the reliability of distance information and the ability to integrate retinal size and distance information. To test whether or not the contribution to size constancy simply reflected the contribution to distance information we also quantified the reliability of the distance information with and without knowledge by asking participants to provide implicit and explicit measures of distance. On size estimation trials, following the manual estimation of object size, participants were explicitly asked to indicate the perceived distance of the object by pointing to its location on the table (i.e., estimated distance). On grasping trials, the reaching distance during the reach-to-grasp movement was considered as an implicit measure of distance processed by the visuomotor system (i.e., reaching distance). As such, estimated and reaching distance reflect the distance assessment in the perception and action systems, respectively (Loomis et al., 1992). It was expected that an analysis of size and distance computations invoked by the different tasks would reveal similarities and differences between the two visual systems in the processing of cognitive signals about distance when vision is limited.

2. Materials and methods

2.1. Participants

Twenty-four participants (12 males and 12 females; mean age: 20.75 ± 1.82) took part in the study. All participants had normal hearing and normal or corrected-to-normal vision. None of the participants reported a history of neurological or psychiatric disorders. All participants were naive as to the purposes of the study. The experiment was conducted following the ethical standards laid down in the 1964 Declaration of Helsinki and in accordance with the guidelines of the Research Ethics Board of the South China Normal University. All participants gave informed consent to participate in the study before the start of their experimental session and received Monetary compensation for their time.

2.2. Apparatus

Participants wore liquid crystal goggles (PLATO goggles; Translucent Technologies, Toronto, ON, Canada) to control vision of the object (Fig. 1A). Participants performed the task in an otherwise completely dark room. To further eliminate depth cues, a mask with a 1 mm-aperture at its center was inserted in the inner side of the goggles' right lens so that participants were able to see only the glow-in-the-dark target object through this 1-mm pinhole (Fig. 1B) when the right lens of the goggles was opened (i.e., restricted viewing). This manipulation is effective at removing all binocular, and most of the monocular depth cues. As a consequence, size constancy in both perception and action is typically disrupted under such restricted viewing conditions (Chen, et al., 2018; Chen et al., 2019; Holway and Boring, 1941; Maturana et al., 1972; Sperandio and Chouinard, 2015; Sperandio et al., 2009).

An OPTOTRAK system (Northern Digital, Waterloo, ON, Canada) with a sampling frequency of 200 Hz was used to record hand movements (Fig. 1A inset). Two infrared light emitting diodes (IREDs) were attached to the right corner of the thumbnail and the left corner of the index finger to record the 3D positions of the hand during grasping and manual estimation tasks.

The experiment was programmed in MATLAB (The Mathworks, Natick MA; https://ww2.mathworks.cn/) and the MOTOM toolbox (Derzsi and Volcic, 2018) was used to communicate with the OPTOTRAK.

2.3. Stimuli

Stimuli consisted of five 3D-printed hollow white spheres of different diameters (Fig. 1C). All the spheres were painted with glow-in-the-dark paint and they were recharged in a box by means of a strong light in between trials. The spheres were resting on top of a pedestal. Each sphere had a small base of a specific length attached to it. This was done to secure the spheres to the pedestal and ensure that the center of the sphere was always aligned with the participants' gaze (Fig. 1C).

The spheres of 2.5 cm (small) and 5 cm (large) diameter placed at 20 cm (near) and 40 cm (far) viewing distance were the target stimuli we used for analysis. The remaining three spheres of 1.25 cm, 3.75 cm, and 6.25 cm in diameter were used to increase variability in object size and to keep participants more engaged with the task. An additional viewing distance of 30 cm was also included. It should be noted that only the following four conditions were analyzed: small-near, small-far, large-near, and large-far. Crucially, the near-small and far-large stimuli subtended the same visual angle and, therefore, generated the same image size on the retina. In contrast, the small target at the far position and the large target at the near position generated the smallest and the largest retinal sizes, respectively (Fig. 1D).

2.4. Procedure and design

The aim of this study was twofold: 1) evaluate the contribution of distance knowledge to size constancy in perception and action; 2) evaluate the contribution of distance knowledge to explicit (estimated distance) and implicit (reaching distance during grasping) measures of distance.

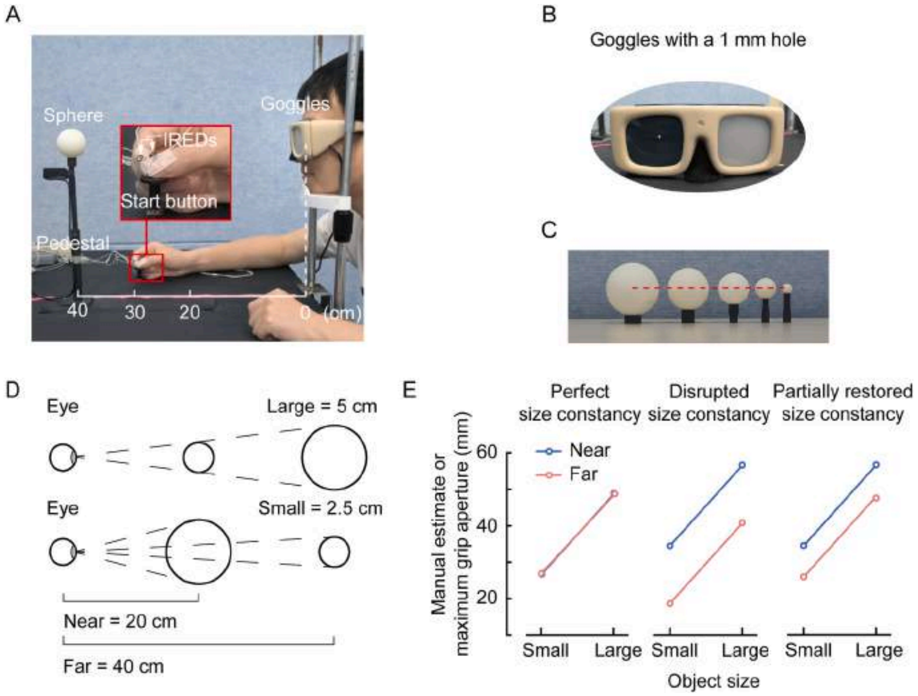


Fig. 1. Setup, stimuli and hypotheses. (A) The set-up of the experiment. Participants wore PLATO goggles throughout the experiment. The position of the fingers during grasping and size estimation were recorded by two infrared light emitting diodes (IREDs) attached to the right corner of the thumbnail and the left corner of the index. (**B**) PLATO goggles with a pinhole. The PLATO goggles could be turned on or off to control the visibility of the stimulus. A black paper patch with a 1 mm pinhole was attached to the right lens of the goggles so that participants could view the stimuli monocularly with their right eye through a 1-mm pinhole in complete darkness. (**C**) Stimuli. The stimuli were five 3D-printed spheres of 1.25 cm, 2.5 cm, 3.75 cm, 5 cm, and 6.25 cm in diameter. Each sphere was positioned on a pedestal in such a way that its center was always aligned with the participants' gaze. (**D**) To measure size constancy, the main experimental conditions included two object sizes (2.5 cm and 5 cm) and two viewing distances (20 cm and 40 cm). Additional sizes and distances were introduced to increase the unpredictability of conditions. (**E**) Hypotheses. If size constancy holds true, then participants' perceptual experience of object size as well as their grip apertures will be unaffected by viewing distance (i.e., size constancy in perception and action, respectively; red and blue lines perfectly overlap, i.e., perfect size constancy). If size constancy breaks down, then participants will perceive object size and scale their grip aperture according to the retinal image size of the objects. As a result, perceived size and maximum grip aperture will be larger at the near than at the far distance (i.e., disrupted size constancy). If semantic knowledge of distance can restore size constancy to some extent, then the gap between the near and far lines should be reduced (i.e., partially restored size constancy).

2.4.1. *The contribution of distance knowledge to size constancy*

To test the contribution of distance knowledge to size constancy in perception and action, participants were asked to complete two tasks in separate blocks: grasping and manual size estimation (Fig. 2). Manual size estimation and grasping have been widely used before to examine the perception and action systems, respectively (Bulakowski, et al., 2009; Chen et al., 2015a; Chen et al., 2015b; Daprati and Gentilucci, 1997; Ganel et al., 2008; Ganel and Goodale, 2003; Hesse et al., 2012; Sun et al., 2021).

The tasks were performed under two conditions: i) restricted viewing (i.e., monocular view, pinhole, dark room) without distance knowledge (i.e., *restricted-noKno*); ii) restricted viewing with distance knowledge (i.e., *restricted-withKno*). To provide distance knowledge, an auditory cue was delivered by means of a speaker placed on the table. The auditory cue consisted in a recorded male voice, saying the Chinese words for "20 cm", "30 cm", or "40 cm". Each audio file lasted for about 1 s.

During the experiment, participants sat in front of a black table with their chin on a chinrest. At the beginning of each trial, the goggles were closed and participants were asked to hold down a button with their thumb and index fingers pinched together. Meanwhile, the experimenter positioned a specified sphere on the pedestal at a specified distance, according to the conditions set for the trial. Next, the speaker played the audio message "20 cm", "30 cm", or "40 cm" to indicate the distance of the target object in the *restricted-withKno* condition, or stayed silent for 1 s in the *restricted-noKno* condition. Two seconds later, the right lens of the goggles was opened. Participants viewed the glowing object in complete darkness through the pinhole on the right lens of the goggles.

On manual size estimation trials, participants were asked to indicate the perceived size of the target sphere with the opening of their right thumb and index finger as accurately as possible (Fig. 2A). There was no time limit for this task, and they were allowed to adjust their estimation if they wanted to. When participants reported being satisfied with their size estimation, the OPTOTRAK would be triggered to record the position of their fingers for 500 ms (i.e., the record segment, solid line in Fig. 2B, the dashed part was to show the distance between the two fingers throughout of the trial). The average of the inter-finger distance recorded over the 125-ms interval was used as a perceptual judgment of object size (i.e., estimated size).

On grasping trials, participants were asked to grasp the sphere with their right thumb and index fingers naturally and accurately as soon as the goggles were opened (Fig. 2C). The positions of their fingers were recorded for 3 s from the opening of the right lens of the goggles (Fig. 2D). During grasping, the grip aperture became larger and larger and reached the maximum grip aperture (MGA) well before the fingers made contact with the object. The participants then released the sphere and brought their hand back to the start button to initiate the next trial.

2.4.2. *The contribution of distance knowledge to distance computations*

To evaluate the contribution of knowledge about distance to manual size estimation, participants were also asked to report the perceived location of the sphere by placing their index fingertip on the location of the target on the table right after they manually estimated the size of the target. Specifically, after the manual estimate was recorded, there was an auditory cue to remind participants to move their fingers back to the start button and then place their right index finger to where the object

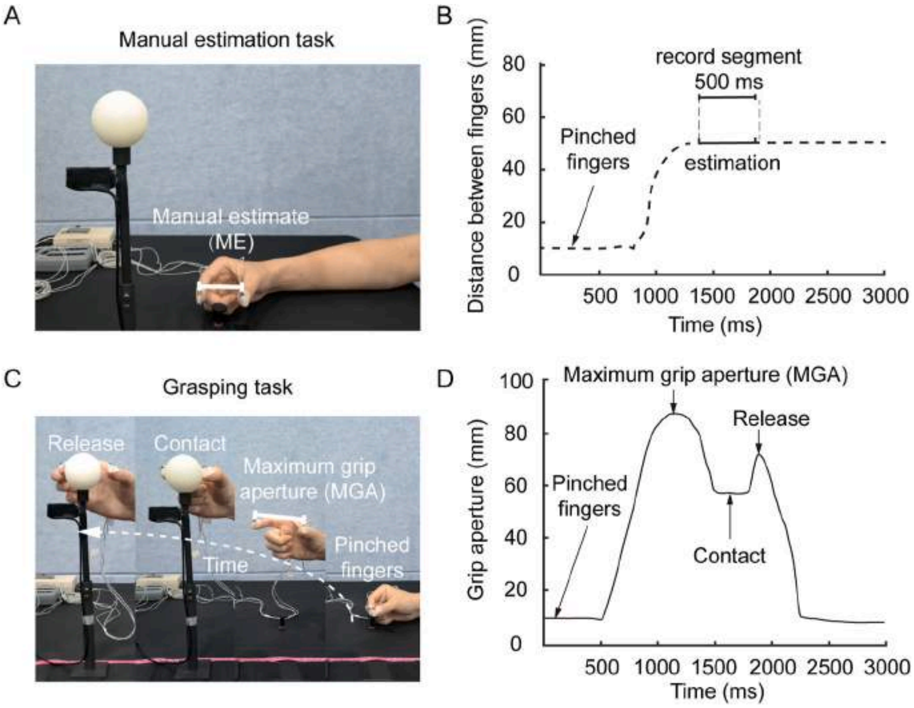


Fig. 2. Manual size estimation and grasping tasks. (**A**) Manual size estimation task. Participants were instructed to report the perceived size of the sphere as accurately as possible by opening their thumb and index finger a matching amount. (**B**) The Euclidean distance between the two fingers shown for 3000 ms during the manual estimation. The average of the inter-finger distance recorded over the 500 ms after participants reported being satisfied with their estimate was used as a perceptual judgment of object size (i.e., estimated size, solid line). (**C**) Grasping task. Participants reached out and picked up the sphere in a 'natural manner' with their thumb and index finger as quickly as possible after seeing the target object. (**D**) The Euclidean distance between the index finger and thumb while grasping. At the beginning of grasping trials, the fingers were pinched together. The fingers then began to open, reached maximum grip aperture (MGA), and then closed down on the object (Contact) and lifted it up, and finally put it down (Release). The maximum grip aperture (which always happens before participants make contact with the sphere) was used as the dependent variable for the grasping task to indicate the grip scaling of the hand when grasping the target object.

was. The distance between the tip of the index finger and the end of the table where participants' chest were contacting was recorded as an explicit measure of perceived distance. Participants were allowed to adjust their report and there was no time limit for the task. On grasping trials, regardless of whether or not participants reached towards the target successfully, the end point of the reach-to-grasp movement was considered as an implicit measure of distance.

For both estimation and grasping tasks, the goggles were closed as soon as participants released the button. In other words, participants performed the tasks without receiving visual feedback from their hands or the target (i.e., open loop). Hence, the maximum grip aperture mainly reflected the programming of grasping before the actual hand movement (Heath et al., 2005).

At the end of each estimation trial, the experimenter placed the sphere into the participants' hand so that they could receive the same haptic feedback as in the grasping task. On grasping trials, participants reached out but often failed to pick up the sphere (78.97% in the *restricted-noKno* condition, and 66.57% in the *restricted-withKno* condition). In this case, the experimenter put the sphere into the participants' hand to match haptic feedback across conditions. In doing so, we excluded any difference between tasks due to haptic feedback (Bozzacchi et al., 2014; Schenk, 2012).

The experiment consisted of four blocks, one block for each task (estimation and grasping) at each knowledge condition (*restricted-noKno* and *restricted-withKno*). In each block, each size (2.5 cm and 5 cm) and distance (20 cm and 40 cm) combination had 8 repetitions for a total of 32 trials to be included in the analysis. The other 3 spheres (1.25 cm, 3.75 cm and 6.25) were repeated twice at each of these two positions, and finally, all five spheres were presented once at the intermediate distance of 30 cm. These 11 trials were excluded from analysis. The order of the blocks as well as the conditions in each block was randomized.

2.5. *Data analysis*

2.5.1. *Analysis of size constancy in size estimation task and grasping task*

For the size estimation task, the average of the Euclidean distance between the index finger and thumb was recorded for 500 ms after participants reported being satisfied with their matching. This distance was used as the dependent measure of the perceived size of the target (manual estimates, ME). For the grasping task, during reach-to-grasp movement the participants' hand opened wider and wider as it approached the object and reached a peak aperture prior to target contact. The peak or maximum grip aperture (MGA) was used as the dependent measure for grasping. Because the MGA always occurs before participants make contact with the object and scales with the object size (Castiello, 2005), it can be used to indicate the extent to which participants rely on visual information to guide their actions.

First, to evaluate whether or not participants showed size constancy, repeated-measures ANOVAs with Distance (near vs. far) and Size (small and large) as within-subject factors were performed. A significant main effect of distance would indicate the disruption of size constancy.

Then, to evaluate the amount of size constancy that was disrupted due to the removal of visual cues and the extent to which it was restored by introducing distance knowledge, we calculated a size-constancy disruption index, which captures the differences in manual estimates (ME) or MGA between the near and far distances (Chen, et al., 2018). A size-constancy disruption index (DI) of 0 indicates no effect of viewing distance on ME and MGA, namely perfect size constancy. A positive DI indicates an effect of viewing distance on ME and MGA, namely a disruption in size constancy.

Because the functions describing the relationship between object size and ME and object size and MGA have different slopes, a 1 mm difference in ME may be different from a 1 mm difference in MGA (Chen et al., 2015a; Chen et al., 2015b; Chen et al., 2018; Westwood et al., 2002). As a result, it is not reasonable to directly perform ANOVA with task as a factor.

In order to directly compare the results between size estimation and grasping tasks, the DI was corrected by the average of the two "slopes" for two distances of corresponding experimental condition (i.e., restricted-noKno estimation; restricted-withKno estimation; restricted-noKno grasping; restricted-withKno grasping). Thus, the following calculations were performed on the data:

$$\mathrm{DI}_{\mathrm{estimation}} = (\mathrm{ME}_{\mathrm{near}} - \mathrm{ME}_{\mathrm{far}})_{\mathrm{Averaged\ Across\ Sizes}} / \mathrm{Slope}_{\mathrm{Averaged\ Across\ Distances}}$$

$$\mathrm{DI}_{\mathrm{grasping}} = (\mathrm{MGA}_{\mathrm{near}} - \mathrm{MGA}_{\mathrm{far}})_{\mathrm{Averaged\ Across\ Sizes}} / \mathrm{Slope}_{\mathrm{Averaged\ Across\ Distances.}}$$

The "slope" was $(\mathrm{ME}_{\mathrm{large}} - \mathrm{ME}_{\mathrm{small}})$/(Large-Small) for size estimation and $(\mathrm{MGA}_{\mathrm{large}} - \mathrm{MGA}_{\mathrm{small}})$/(Large-Small) for grasping. Given that (Large-Small) is constant in all cases, DI was defined as follows:

$$\mathrm{DI}_{\mathrm{estimation}} = (\mathrm{ME}_{\mathrm{near}} - \mathrm{ME}_{\mathrm{far}})_{\mathrm{Averaged\ Across\ Sizes}} / (\mathrm{ME}_{\mathrm{large}} - \mathrm{ME}_{\mathrm{small}})_{\mathrm{Averaged\ Across\ Distances}}$$

$$\mathrm{DI}_{\mathrm{grasping}} = (\mathrm{MGA}_{\mathrm{near}} - \mathrm{MGA}_{\mathrm{far}})_{\mathrm{Averaged\ Across\ Sizes}} / (\mathrm{MGA}_{\mathrm{large}} - \mathrm{MGA}_{\mathrm{small}})_{\mathrm{Averaged\ Across\ Distances}}$$

The same procedure was used in our previous study (Chen, et al., 2018).

The contribution of distance knowledge to size constancy in perception and grasping was quantified as the difference between the DI in the *restricted-noKno* condition and the DI in the *restricted-withKno* condition. Paired-samples *t*-test or one-sample *t*-test were used to verify whether the contribution of distance knowledge was significant or whether there was a difference in the contribution of knowledge between tasks.

2.5.2. Analysis of distance measures (estimated distance and reaching distance)

The analyses described so far were performed to assess the contribution of distance knowledge to size constancy in perception and action. Because the computation of object size relies on the processing of distance cues, we also examined the reliability of distance information with and without knowledge. To collect measures of perceived distance during the manual estimation task, participants explicitly report the distance by pointing to where the sphere was located as accurately as possible right after the size estimation and with no time limitation. For the grasping task, instead, distance was measured as the endpoint of reaching during reach-to-grasp movements. The distribution of estimated or reached distance was obtained for all experimental trials in all participants.

A two-sample Kolmogorov-Smirnov test was used to assess whether the frequency distribution of distance measures (i.e., estimated and reaching distance) were different with and without semantic distance knowledge. In addition, unsigned error between the actual target position and distance measures was also calculated for all conditions. Repeated measures ANOVAs with Task and Distance knowledge as within-subject factors were performed to reveal any main effects of task, distance knowledge or their interactions. The contribution of distance knowledge to distance was defined as the $\mathrm{Error}_{\mathrm{noKno}}$-$\mathrm{Error}_{\mathrm{withKno}}$. Paired-samples t-tests were used to indicate whether the contribution of distance knowledge differed across tasks.

It should be noted that the estimated distance and reaching distance were measured differently. First, the distance was estimated after size estimation on estimation trials. Therefore, there was a delay between viewing the object and estimating the distance. In contrast, the reaching movement immediately happened after participants viewed the object. Second, there was distance feedback on grasping trials when participants successfully grasped the object but no distance feedback on estimation trials. The influence of the delay and distance feedback must be considered when interpreting the results.

3. Results

3.1. Contribution of distance knowledge to size constancy

Fig. 3A shows the manual estimates (MEs) and maximum grip aperture (MGAs) in restricted-viewing condition with and without distance knowledge, respectively. Consistent with previous research (Chen, et al., 2018; Chen et al., 2019; Holway and Boring, 1941; Maturana et al., 1972; Sperandio et al., 2009), when participants performed the size estimation and grasping tasks in the restricted-viewing condition, size constancy was largely disrupted for both tasks (the main effect of distance was significant for both ME and MGA. ME: $F\,(1, 23) = 101.549$, $p < 0.001$, $\eta_p^2 = 0.815$; MGA: $F\,(1, 23) = 58.501$, $p < 0.001$, $\eta_p^2 = 0.718$) (Fig. 3A, left column).

The question then is whether or not the inclusion of semantic distance knowledge could restore size constancy at least to some extent. As it turned out, the main effect of distance in the restricted-withKno condition was also significant for both ME and MGA (ME: $F\,(1, 23) = 25.323$, $p < 0.001$, $\eta_p^2 = 0.524$; MGA: $F\,(1, 23) = 15.701$, $p < 0.001$, $\eta_p^2 = 0.406$), suggesting that size constancy was still disrupted even if distance knowledge was provided in the restricted-viewing condition (Fig. 3A, right column).

Nonetheless, visual inspection of Fig. 3A shows that for both ME and MGA the separation of the lines between near and far conditions got smaller when distance knowledge was added. To quantify this improvement in size constancy, we calculated the size-constancy DI as described earlier (see Data analysis) for the conditions with and without distance knowledge. In order to compare the results between ME and MGA, the DI was corrected by the "slope" of corresponding experimental treatment level (i.e., restricted-noKno estimation; restricted-withKno estimation; restricted-noKno grasping; restricted-withKno grasping). This correction was deemed necessary as MGA is usually characterized by a shallower slope than ME (Chen et al., 2015a; Chen et al., 2015b; Chen et al., 2018), as shown in Fig. 3A.

For both ME and MGA, the DI became significantly smaller when distance knowledge was provided (ME: paired samples *t*-test, $t\,(23) = 7.104$, $p < 0.001$, Cohen's $d = 1.450$, two-tailed criterion; MGA: $t\,(23) = 2.124$, $p = 0.045$, Cohen's $d = 0.434$, two-tailed criterion; Fig. 3B). Repeated ANOVAs with Knowledge and Task as within-subject factors revealed a significant main effect of Knowledge ($F\,(1, 23) = 21.491$, $p < 0.001$, $\eta_p^2 = 0.483$). Moreover, the non-significant main effect of Task ($F\,(1, 23) = 0.158$, $p = 0.695$, $\eta_p^2 = 0.007$) and the non-significant interactions between Knowledge and Task ($F\,(1, 23) = 2.778$, $p = 0.109$, $\eta_p^2 = 0.108$) suggest that the inclusion of Knowledge did not affect the two tasks differently.

Then we defined the contribution of distance knowledge to size constancy as the difference in DI between the restricted-withKno and restricted-noKno conditions to compare directly the contribution of distance knowledge to size constancy in the two tasks. We found that the contribution of distance knowledge was significantly larger than 0 for perception (i.e., ME; $t\,(23) = 7.104$, $p < 0.001$, Cohen's $d = 1.450$, two-tailed criterion) and grasping (i.e., MGA) ($t\,(23) = 2.124$, $p = 0.045$, Cohen's $d = 0.434$, two-tailed criterion) (Fig. 4). There was no significant difference in the contribution of semantic distance knowledge to the estimation (perception) and grasping (action) ($t\,(23) = 1.667$, $p = 0.109$, Cohen's $d = 0.340$, two-tailed criterion), which is consistent with the above ANOVA results.

3.2. Contribution of distance knowledge to distance assessment

Because size constancy relies on the reliability of distance information reflecting the integration of distance cues from various sources, we

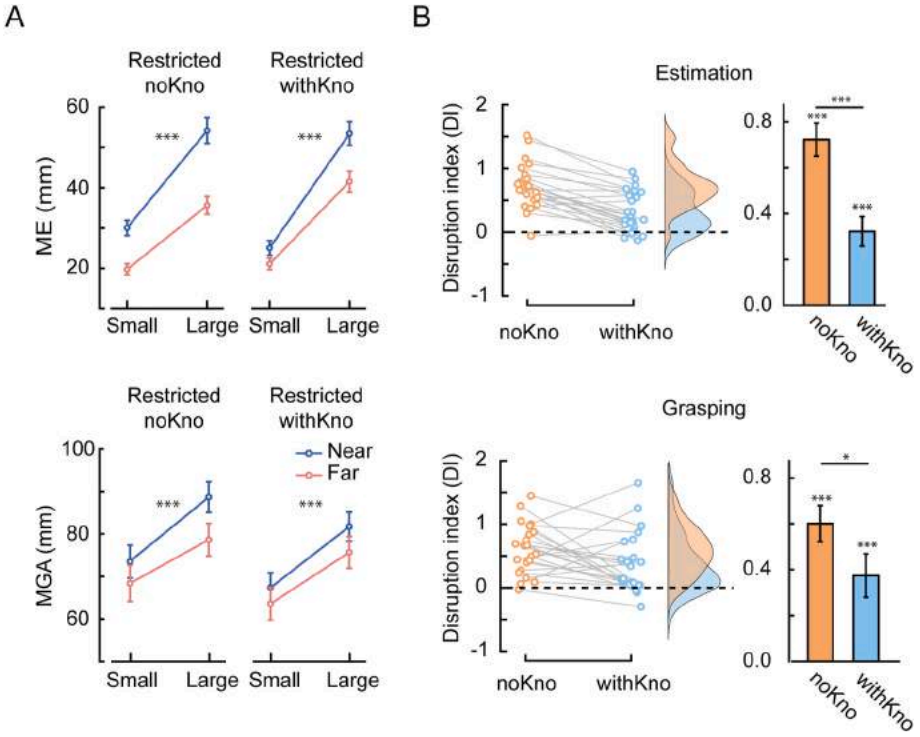


Fig. 3. Results for size constancy measures. (**A**) Manual size estimates (MEs) and maximum grip apertures (MGAs) as a function of object size (small vs. large), viewing distance (near vs. far) and knowledge conditions (restricted-noKno vs. restricted-withKno). (**B**) Disruption index (DI) in all conditions. The left part of the panel shows each participant's DI and the distribution of DIs, the right part of the panel shows the mean DI, with restricted-noKno in orange and restricted-withKno in blue. * indicates p < 0.05. *** indicates p < 0.001. Error bars represent the standard error of the mean.

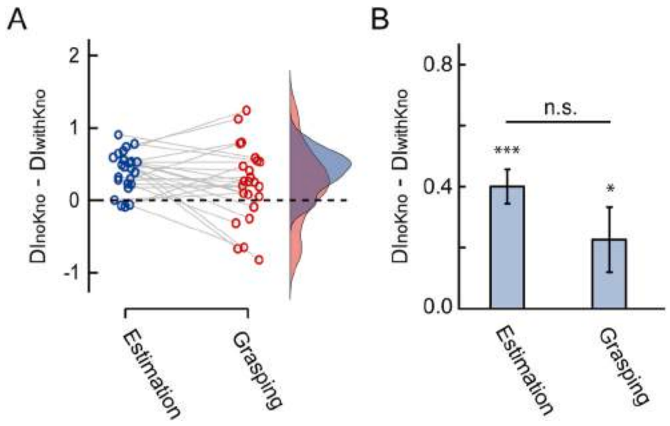


Fig. 4. Contribution of distance knowledge to size constancy. (**A**) Raincloud plots showing the contribution of knowledge for each participant, dark blue and dark red curves represent the contribution of knowledge in the estimation and grasping tasks, respectively. (**B**) The mean contribution of knowledge to estimation and grasping. *** indicates p < 0.001. * indicates p < 0.05. Error bars represent the standard error of the mean.

also analyzed the contribution of distance knowledge to implicit (reaching distance) and explicit (estimated distance) measures of distance. The frequency distribution of estimated distances and reaching distances for all the trials included in the analysis (i.e., spheres with a diameter of 2.5 cm or 5 cm placed at 20 cm or 40 cm) are plotted in Fig. 5A's top and bottom panels, respectively.

As can be seen in Fig. 5A (top panel), the estimated distance is more concentrated on the actual distance in the restricted-withKno conditions compared to the restricted-noKno conditions for both near (20 cm) or far (40 cm) distances. A two-sample Kolmogorov-Smirnov test confirmed that there was a significant difference in the frequency distribution of estimated distance between the restricted-noKno and restricted-withKno conditions (20 cm: $K = 0.3594$, $p < 0.001$; 40 cm: $K = 0.2503$, $p < 0.001$).

Similarly, the reaching distance recorded during grasping movements is also more concentrated on the actual location in restricted-withKno conditions than in the restricted-noKno conditions, as shown in Fig. 5A (bottom panel). A two-sample Kolmogorov-Smirnov test confirmed that there was a significant difference also in the frequency distribution of reaching distance between restricted-noKno conditions and restricted-withKno conditions (K–S test, 20 cm, $K = 0.4126$, $p < 0.001$; 40 cm, $K = 0.3545$, $p < 0.001$).

Interestingly, visual inspection of Fig. 5A also shows that the reaching distance is more accurate than the estimated distance in indicating the real distance of the target. That is, in the frequency distributions of reaching distance in the grasping trials in the restricted-noKno condition there are clear peaks at the actual distance of the target spheres for both the 20-cm and the 40-cm distances. This result is consistent with the finding that even when perceptual distance is disrupted, participants were able to walk to the actual position of the target accurately when their eyes were closed (Loomis, et al., 1992). However, this result could also be because the estimated distance was collected with a delay in estimation trials without distance feedback (see 2.5.2 for details).

To statistically compare the bias in estimated distance and reaching distance, we calculated the errors as the difference in centimeter between the measured and the actual distance in all conditions (Fig. 5B). A 2 × 2 repeated measures ANOVA was carried out on the errors with Knowledge condition and Task as main factors. The ANOVA revealed a significant main effect of Knowledge ($F\,(1, 23) = 154.387$, $p < 0.001$, $\eta_p^2 = 0.870$). For both estimated and reaching distance, the error in restricted-withKno condition was significantly smaller than that in the restricted-noKno conditions (i.e., distance estimation: $p < 0.001$; reaching distance during grasping: $p < 0.001$). The main effect of Task was not significant ($F\ (1,\ 23) = 4.046$, $p = 0.056$, $\eta_p^2 = 0.150$). In

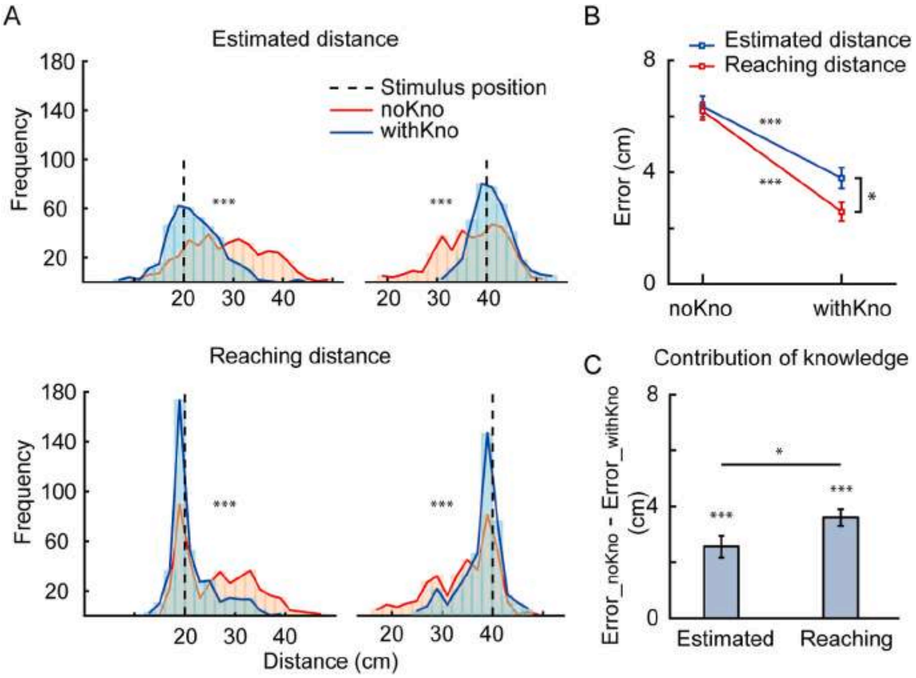


Fig. 5. Utilization of distance knowledge. (**A**) Frequency distributions of distance measures during manual estimation and reach-to-grasp movements for 20 cm and 40 cm distance conditions in the restricted-noKno and restricted-withKno conditions. (**B**) The mean error for both estimated and reaching distances. (C) Contribution of distance knowledge on the estimated distance and reaching distance. The contribution was calculated as the difference in error between noKno and withKno conditions. * indicates $p < 0.05$, *** indicates $p < 0.001$. Error bars represent within-subjects standard error of the mean.

addition, there was a significant interaction between Task and Knowledge (F (1, 23) = 4.678, p = 0.041, η_p^2 = 0.169). Post-hoc analysis showed that there was no significant difference in errors between estimated and reaching distance in the restricted-noKno conditions (t (23) = 0.424, p = 0.676, Cohen's d = 0.086, two-tailed criterion), whereas in the restricted-withKno condition the error for the estimated distance was significantly greater than the error for the reaching distance (t (23) = 2.672, p = 0.014, Cohen's d = 0.545, two-tailed criterion), suggesting differences between the perception and the action systems in their ability to use distance knowledge in distance measurement.

Finally, to compare the contribution of distance knowledge to estimated distance and reaching distance, we calculated the difference in errors between the noKno and withKno conditions (Fig. 5C). A paired-sample t-test showed that the contribution of knowledge was significantly larger in reaching than in estimated distance (t (23) = 2.163, p = 0.041, Cohen's d = 0.441, two-tailed criterion).

4. Discussion

In the current study, we examined the contribution of semantic information about distance to the restoration of size constancy in perceptual report and grasping under restricted viewing conditions. We also looked at how this cognitive signal contributed to explicit (estimated distance) and implicit (reaching distance) measures of distance. We found that although knowledge about distance, a purely cognitive signal, was able to support size-distance scaling operations to some extent, size constancy was never completely restored in either perception or action. This result suggests that, although cognitive-based knowledge about distance may modulate perceptual reports and grasping accuracy to some extent, it is not reliable enough to fully compensate for the lack of other distance cues. Importantly, semantic information about the distance of the target object contributed as much to size constancy in the perceptual task as it did in the grasping task, suggesting no differences between the two systems in the use of this high-level cognitive information for computing object size.

Our analysis of the implicit and explicit measures of distance revealed that reaching distance measured during grasping movements was more accurate (i.e., closer to the actual distance of the target object) than the perceived distance of the object reported by participants during the distance estimation task. This result is consistent with other reports showing that perceptual reports of distance are typically less accurate than motor responses across a range of different actions (Andre and Rogers, 2006; Bingham and Pagano, 1998; Loomis et al., 1992; Pagano and Bingham, 1998).

There is another way to look at the distance data, however. It could be the case that the estimated distance was less accurate than the reaching distance because the former was executed after the manual size estimation had taken place. Thus, the end point of the estimate reach would have depended more on memory – not so much a memory of the verbal instruction itself, but a memory of having viewed the target object in the context of that verbal instruction. There is considerable evidence in the literature that pointing to a remembered target location is less accurate than pointing to the target when it is visible, even in open loop (e.g., McIntyre et al., 1998). In contrast, reaching distance was measured during the reach-to-grasp movement at the moment the participants reached the target. But in any case, this concern was not an issue for our size-constancy results, because the distance information for size constancy in manual estimation would have been incorporated into the computation of target size at the moment the target became visible, just as it was for the grasping task.

It should also be noted, however, that participants experienced feedback of distance on the trials that they grasped the object successfully (although they failed to grasp objects on 78.97% of the trials in the *restricted-noKno* condition and 66.57% of trials in the *restricted-withKno* condition.) Although participants received feedback of the size of the object after each estimation trial, they did not receive feedback about its distance. Thus, it is possible that the visuomotor system could use the feedback about distance to refine the learned association between the verbal information about distance provided by the experimenter and the actual physical distance of the target object. But even if this were the case, there was no difference in the restoration of size constancy between the manual estimation and grasping conditions.

Indeed, it is likely that the verbal information about distance, whether presented in the manual estimation condition or in the grasping condition, invoked the same kind of conceptual information about the apparent distance of the target object. In other words, it seems likely that size constancy in grasping was making use of the same high-level information about distance as size constancy in perceptual report.

The fact that semantic information about distance contributed to some extent to, but did not fully restore, size constancy in perception and action, stands in stark contrast to our previous findings where we demonstrated under similar viewing conditions that proprioceptive distance signals from the non-dominant hand were able to completely re-establish size constancy in action but only partly in perception (Chen, et al., 2018). Proprioception, of course is much more frequently used than semantic distance information to control grasping. People often, for example, reach for objects held in their other hand when vision does not provide sufficient or reliable information – and this experience could be utilized to compute the real size of visible objects when visual cues to the distance of that object are not available or are degraded. We also showed that proprioception can make a modest contribution to restoring perceptual size constancy when visual information about distance is unavailable, but not nearly so strikingly as in the case of grasping (Chen, et al., 2018). But even though perceptual size constancy benefits only a little from proprioceptive information about distance, the actual perception of location through proprioceptive signals is still pretty impressive: studies using a position-matching task when eyes were closed or blindfolded have shown that the error of matching the location of one hand with the other hand was on average only about 1 cm (van Beers, Sittig and Denier van der Gon, 1998; Wang et al., 2022; Wilson et al., 2010). In the present study, however, the mean error in estimated and reaching distance with respect to the actual distance ranged between 3 and 4 cm, suggesting that extracting distance information from stored knowledge about the actual distance implied by a phrase such as "20 cm" is far less reliable than proprioceptive information. In short, without training, most people might not know exactly how far an object said to be 20, 30, or 40 cm away is actually located.

Moreover, semantic information about distance is an indirect conceptual/cognitive signal whereas proprioception is a direct sensory signal. Whilst the integration of signals between different sensory modalities has a great survival value and occurred early in evolution (Stein et al., 2014), verbal information about distance has a more recent evolutionary history and its use in computing size constancy may require some sort of explicit cognitive modeling of the verbally indicated distance and the perception of the extent of a visible object in an otherwise restricted visual environment. Indeed, some might argue that the need for such computations is rarely encountered at all in peri-personal space.

As suggested earlier, it is likely that both perceptual (manual) estimation and grasping made use of the same high-level representation of distance in the restricted-Kno condition. It is not clear how this information interacted with networks in the dorsal stream to compute the real-world size of the target sphere (i.e., size constancy). It is possible that the dorsal stream was only minimally involved. Indeed, patients with bilateral dorsal-stream lesions have been shown to exhibit better control over their reaching and grasping movements when they are using perceptual memory of the size and/or location of the target object rather than real-time information about the size of the object (Milner et al., 2003). The stored perceptual memories of object size may be akin in some ways to the high-level cognitive information used by the neurologically intact participants in the current study. In any case, it is evident that the restoration of size constancy by high-level distance information is no better for grasping than it is for perception.

Finally, it should be noted that Linton (2020, 2021a, 2021b) has recently put forth a purely cognitive explanation for size and distance perception according to which size-distance scaling is solely the result of our subjective knowledge about changes in viewing distance. Our findings, however, do not fully support Linton's theory. We showed that semantic knowledge of distance, a purely cognitive signal, is not enough to restore perfect size constancy under restricted viewing conditions for either perceptual report or grasping. Future studies of the role of cognitive processes on size-distance computations would be wise to consider how this perceptual information interacts with the neural pathways engaging perception and the visual control of action.

CRediT authorship contribution statement

Gexiu Wang: Data curation, Formal analysis, Methodology, Validation, Visualization. **Chao Zheng:** Data curation, Methodology. **Xiaoqian Wu:** Data curation, Methodology. **Zhiqing Deng:** Conceptualization, Project administration, Supervision. **Irene Sperandio:** Supervision, Writing – original draft, Writing – review & editing. **Melvyn A. Goodale:** Supervision, Writing – original draft, Writing – review & editing. **Juan Chen:** Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare no competing financial interests.

Data availability

Data will be made available on request.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (No. 31970981 and No. 31800908) and the National Science and Technology Innovation 2030 Major Program (STI2030-Major Projects 2022ZD0204802).

References

Alais, D., Burr, D., 2004. The ventriloquist effect results from near-optimal bimodal integration. Curr. Biol. 14, 257–262.

Andre, J., Rogers, S., 2006. Using verbal and blind-walking distance estimates to investigate the two visual systems hypothesis. Percept. Psychophys. 68, 353–361.

Bingham, G.P., Pagano, C.C., 1998. The necessity of a perception-action approach to definite distance perception: monocular distance perception to guide reaching. J. Exp. Psychol. Hum. Percept. Perform. 24, 145–168.

Blakemore, C., Garner, E.T., Sweet, J.A., 1972. The site of size constancy. Perception 1, 111–119.

Boring, E.G., 1940. Size constancy and Emmert's law. Am. J. Psychol. 53, 293–295.

Bozzacchi, C., Volcic, R., Domini, F., 2014. Effect of visual and haptic feedback on grasping movements. J. Neurophysiol. 112, 3189–3196.

Broderick, M.P., Anderson, A.J., Lalor, E.C., 2019. Semantic context enhances the early auditory encoding of natural speech. J. Neurosci. 39, 7564–7575.

Bulakowski, P.F., Post, R.B., Whitney, D., 2009. Visuomotor crowding: the resolution of grasping in cluttered scenes. Front. Behav. Neurosci. 3.

Castiello, U., 2005. The neuroscience of grasping. Nat. Rev. Neurosci. 6, 726–736.

Chai, Y.H., Liu, T.N.T., Marrett, S., Li, L.Q., Khojandi, A., Handwerker, D.A., Alink, A., Muckli, L., Bandettini, P.A., 2021. Topographical and laminar distribution of audiovisual processing within human planum temporale. Prog. Neurobiol. 205.

Chen, J., Jayawardena, S., Goodale, M.A., 2015a. The effects of shape crowding on grasping. J. Vis. 15, 1–9.

Chen, J., Sperandio, I., Goodale, M.A., 2015b. Differences in the effects of crowding on size perception and grip scaling in densely cluttered 3-D scenes. Psychol. Sci. 26, 58–69.

Chen, J., Sperandio, I., Goodale, M.A., 2018. Proprioceptive distance cues restore perfect size constancy in grasping, but not perception, when vision is limited. Curr. Biol. 28, 1–6.

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. Curr. Biol. 29, 2237–2243.e2234.

Combe, E., Wexler, M., 2010. Observer movement and size constancy. Psychol. Sci. 21, 667–675.

Daprati, E., Gentilucci, M., 1997. Grasping an illusion. Neuropsychologia 35, 1577–1582.

Derzsi, Z., Volcic, R., 2018. MOTOM toolbox: MOtion tracking via optotrak and matlab. J. Neurosci. Methods 308, 129–134.

Ernst, M.O., Banks, M.S., 2002. Humans integrate visual and haptic information in a statistically optimal fashion. Nature 415, 429–433.

Ernst, M.O., Bulthoff, H.H., 2004. Merging the senses into a robust percept. Trends Cognit. Sci. 8, 162–169.

Foxe, J.J., Schroeder, C.E., 2005. The case for feedforward multisensory convergence during early cortical processing. Neuroreport 16, 419–423.

Ganel, T., Chajut, E., Algom, D., 2008. Visual coding for action violates fundamental psychophysical principles. Curr. Biol. 18, R599–R601.

Ganel, T., Goodale, M.A., 2003. Visual control of action but not perception requires analytical processing of object shape. Nature 426, 664–667.

Heath, M., Rival, C., Westwood, D.A., Neely, K., 2005. Time course analysis of closed- and open-loop grasping of the müller-lyer illusion. J. Mot. Behav. 37, 179–185.

Hesse, C., Ball, K., Schenk, T., 2012. Visuomotor performance based on peripheral vision is impaired in the visual form agnostic patient DF. Neuropsychologia 50, 90–97.

Holway, A.H., Boring, E.G., 1941. Determinants of apparent visual size with distance variant. Am. J. Psychol. 21–37.

Kentridge, R.W., 2018. Vision: non-illusory evidence for distinct visual pathways for perception and action. Curr. Biol. 28, R264–R266.

Linton, P., 2020. Does vision extract absolute distance from vergence? Atten. Percept. Psychophys. 82, 3176–3195.

Linton, P., 2021a. Does vergence affect perceived size? Vision 5, 33.

Linton, P., 2021b. V1 as an Egocentric Cognitive Map. Neuroscience of Consciousness, p. niab017, 2021.

Loomis, J.M., Da Silva, J.A., Fujita, N., Fukusima, S.S., 1992. Visual space perception and visually directed action. J. Exp. Psychol. Hum. Percept. Perform. 18, 906–921.

Maturana, H.R., Varela, F.G., Frenk, S.G., 1972. Size constancy and the problem of perceptual spaces. Cognition 1, 97–104.

McIntyre, J., Stratta, F., Lacquaniti, F., 1998. Short-term memory for reaching to visual targets: psychophysical evidence for body-centered reference frames. J. Neurosci. 18, 8423–8435.

Milner, A.D., Dijkerman, H.C., McIntosh, R.D., Rossetti, Y., Pisella, L., 2003. Delayed reaching and grasping in patients with optic ataxia. Prog. Brain Res. 142, 225–242.

Ohshiro, T., Angelaki, D.E., DeAngelis, G.C., 2017. A neural signature of divisive normalization at the level of multisensory integration in primate cortex. Neuron 95, 399–411.e398.

Pagano, C.C., Bingham, G.P., 1998. Comparing measures of monocular distance perception: verbal and reaching errors are not correlated. J. Exp. Psychol. Hum. Percept. Perform. 24, 1037–1051.

Rohe, T., Noppeney, U., 2018. Reliability-weighted integration of audiovisual signals can Be modulated by top-down attention. Eneuro 5. ENEURO.0315-0317.2018.

Schenk, T., 2012. No dissociation between perception and action in patient DF when haptic feedback is withdrawn. J. Neurosci. 32, 2013–2017.

Sperandio, I., Chouinard, P.A., 2015. The mechanisms of size constancy. Multisensory Res. 28, 253–283.

Sperandio, I., Savazzi, S., Gregory, R.L., Marzi, C.A., 2009. Visual reaction time and size constancy. Perception 38, 1601–1609.

Stein, B.E., Stanford, T.R., Rowland, B.A., 2014. Development of multisensory integration from the perspective of the individual neuron. Nat. Rev. Neurosci. 15, 520–535.

Sun, C., Chen, J., Chen, Y., Tang, R., 2021. The influence of induced emotions on distance and size perception and on the grip scaling during grasping. Front. Psychol. 12.

van Beers, R.J., Sittig, A.C., Denier van der Gon, J.J., 1998. The precision of proprioceptive position sense. Exp. Brain Res. 122, 367–377.

Wang, C., Gao, J., Deng, Z., Zhang, Y., Zheng, C., Liu, X., Sperandio, I., Chen, J., 2022. Extracurricular sports activities modify the proprioceptive map in children aged 5–8 years. Sci. Rep. 12, 9338.

Westwood, D., Danckert, J., Servos, P., Goodale, M., 2002. Grasping two-dimensional and three-dimensional objects in visual-form agnosia. Experimental brain research. Experimentelle Hirnforschung. Expérimentation cérébrale 144, 262–267.

Whitwell, R.L., Sperandio, I., Buckingham, G., Chouinard, P.A., Goodale, M.A., 2020. Grip constancy but not perceptual size constancy survives lesions of early visual cortex. Curr. Biol. 30, 3680–3686.e3685.

Wilson, E.T., Wong, J., Gribble, P.L., 2010. Mapping proprioception across a 2D horizontal workspace. PLoS One 5, e11851.

Zhu, Y., Qian, X., Yang, Y., Leng, Y., 2013. The influence of explicit conceptual knowledge on perception of physical motions: an ERP study. Neurosci. Lett. 541, 253–257.