

# PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B

BIOLOGICAL SCIENCES

## Looking at the Ebbinghaus illusion: Differences in neurocomputational requirements, not gaze-mediated attention, explain a classic perception-action dissociation.

Journal:	<i>Philosophical Transactions B</i>
Manuscript ID	RSTB-2021-0459.R1
Article Type:	Research
Date Submitted by the Author:	n/a
Complete List of Authors:	Whitwell, Rob; University of Western Ontario, Department of Psychology Garach, Mehul ; University of Toronto, Division of Orthopaedics, St. Michael's Hospital Goodale, Melvyn; University of Western Ontario, Department of Psychology Sperandio, Irene; University of Trento, Department of Psychology and Cognitive Science
Issue Code (this should have already been entered and appear below the blue box, but please contact the Editorial Office if it is not present):	3DVISION
Subject:	Neuroscience < BIOLOGY, Cognition < BIOLOGY
Keywords:	Visual illusions, visual perception, grasping, eye movements, gaze, two visual systems hypothesis

SCHOLARONE™  
Manuscripts

1  
2  
3 **Author-supplied statements**  
4

5 Relevant information will appear here if provided.  
6

7  
8 **Ethics**  
9

10 *Does your article include research that required ethical approval or permits?:*

11 Yes  
12

13 *Statement (if applicable):*

14 The study was approved by the Research Ethics Board at the University of Western Ontario in  
15 accordance with the 1964 Declaration of Helsinki.  
16

17  
18 **Data**  
19

20 *It is a condition of publication that data, code and materials supporting your paper are made publicly*  
21 *available. Does your paper present new data?:*

22 Yes  
23

24 *Statement (if applicable):*

25 Data for the reported experiment is currently available via the Open Science Framework (OSF):  
26 [https://osf.io/3ucvd/?view\\_only=5030d0ac14df4e55b11d228af467fe9b](https://osf.io/3ucvd/?view_only=5030d0ac14df4e55b11d228af467fe9b)  
27  
28

29 **Conflict of interest**  
30

31 I/We declare we have no competing interests  
32

33 *Statement (if applicable):*

34 CUST\_STATE\_CONFLICT :No data available.  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

**Author-supplied statements****Ethics Statement:**

The study was approved by the Research Ethics Board at the University of Western Ontario in accordance with the 1964 Declaration of Helsinki.

**Data Statement:**

Data for the reported experiment is currently available via the Open Science Framework (OSF): [https://osf.io/3ucvd/?view\\_only=5030d0ac14df4e55b11d228af467fe9b](https://osf.io/3ucvd/?view_only=5030d0ac14df4e55b11d228af467fe9b)

**Conflict of Interest:**

We report no potential conflicts of interest

**Looking at the Ebbinghaus illusion: Differences in neurocomputational requirements, not gaze-mediated attention, explain a classic perception-action dissociation**

Robert L. Whitwell<sup>1</sup>, Mehul A. Garach<sup>2</sup>, Melvyn A. Goodale<sup>1</sup>, and Irene Sperandio<sup>3\*</sup>.

<sup>1</sup>Department of Psychology, University of Western Ontario, London, Ontario, N6A 5C2, Canada.

<sup>2</sup>Division of Orthopaedics, St. Michael's Hospital, University of Toronto, Toronto, Canada

<sup>3</sup>Department of Psychology and Cognitive Science, University of Trento, Rovereto (TN), 38068, Italy.

\*Corresponding Author:

Dr. Irene Sperandio

Department of Psychology and Cognitive Science

University of Trento

Phone: ++39 0464 80 8181

Web: <https://irenesperandio.wixsite.com/website>

Email: [irene.sperandio@unitn.it](mailto:irene.sperandio@unitn.it)

ORCID Author Identifier:

Whitwell 0000-0003-1352-1529

Goodale 0000-0001-6748-0561

Sperandio 0000-0003-1775-2842

## Summary

Perceiving and grasping an object present an animal with different sets of computational problems. The solution in primates entails the specialization of separate neural networks for visual processing with different object representations. This explains why the Ebbinghaus illusion minimally affects the grasping hand's in-flight aperture, which normally scales with target size, even though the size of the target disc **remains** misperceived. An attractive alternative account, however, posits that grasps are refractory to the illusion because participants fixate on the target and fail to attend to the surrounding context. To test this account, we tracked both limb and gaze while participants made forced-choice judgments of relative disc size in the Ebbinghaus illusion or did so in combination with grasping or manually estimating **the size of** one of the discs. We replicated the classic dissociation: grasp aperture was refractory to the measured illusory effect on perceived size, while judgments and manual estimates of disc size were not. Importantly, the number of display-wide saccades per second and the percentage of total fixation time or fixations directed at the selected disc failed to explain the dissociation. Our findings support the contention that object perception and goal-directed action rely on distinct visual representations.

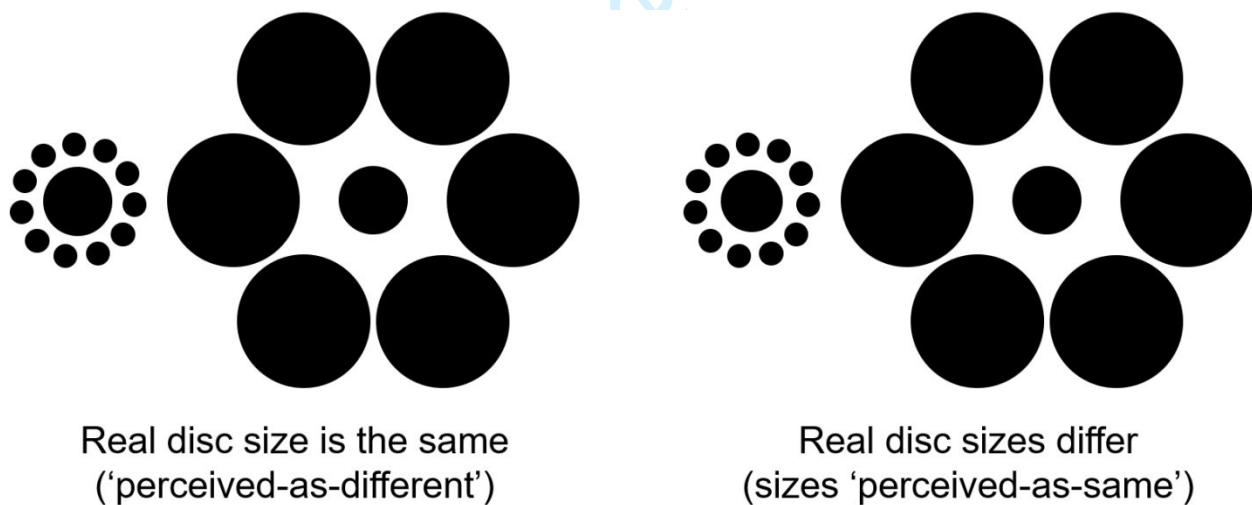
Keywords (4-6):

Visual illusions, visual perception, grasping, eye movements, gaze, two visual systems hypothesis

Words: 198

## Introduction

In a now classic grasping paradigm introduced in 1995, Aglioti and his colleagues [1] showed that the scaling of grasp aperture when reaching for a disc embedded in an Ebbinghaus display (see Fig. 1) was largely refractory to the disc's misperceived size. In their experiment, participants chose one or the other disc to reach for and pick up based on a rule; half of the participants are asked to select the disc on left side of the display if the sizes of the discs look the same and to select the disc on the right side of the display if the sizes of the discs look different; and the display-side contingency is reversed for the remaining half of the participants. An initial two-alternative forced-choice (2AFC) test session establishes the difference in real inner disc size required to make the sizes of two discs look the same. In this session, different-sized inner discs are paired in the display and the participant is asked to report whether they look the "same" or "different" in size [1]. Aglioti et al.'s [1] experiment showed, not surprisingly,



**Figure 1.** The Ebbinghaus display devised for the current study. Inner black-filled circles are surrounded by an annulus of small or large black-filled circles. Two versions of the display are illustrated, one in the left panel and one in the right, and each display is presented as a 'dual' configuration. Left panel: the traditional presentation in which the real sizes of the inner discs are the same but are misperceived as different. This is the critical feature of the display on 'perceived-as-different' trials. Right panel: a version in which the real sizes of the inner discs are different, leading to the impression that they are the same size. The difference in size required to achieve this varies from person to person, but it is scaled here to the most common difference in the study (2mm). A pair of discs whose real sizes were different to make them appear to be the same in size is the critical component of the 'perceived-as-same' display.

1  
2  
3 that disc choice, as expressed by which of the two discs was picked up, was based on illusory  
4 disc size. Yet at the same time, the hand's in-flight grasp aperture remained tuned to the  
5 target's real size.  
6  
7  
8  
9

10  
11 This dissociation was later replicated, using additional controls for visual and haptic feedback:  
12 the grasps were planned with vision available but executed without it [2]; and an additional task  
13 was introduced in which participants estimated the size of the selected disc by opening their  
14 thumb and index finger to create a matching-sized gap that was measured using the same  
15 motion-tracking techniques used for recording the grasps. Insofar as the estimate is a faithful  
16 'read-out' of the participant's visual experience of disc size, then the task permits a more direct  
17 measure of perceived size. These early studies support the view that object-vision for  
18 perception and actions rely on distinct neural representations in different networks of neural  
19 circuitry [3-7].  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30

31  
32 This interpretation has been contested by several groups for a host of different reasons (for a  
33 review, see [8]). Here, while controlling for a list of confounds (e.g., [9]), we address the earliest  
34 of these critiques [10-11], which was based on an asserted difference in the deployment of  
35 attention across the decision-making and movement execution phases of participants' response  
36 of the task. We refer to this as the attention account, but it relies more on gaze as a proxy for  
37 selective attention.  
38  
39  
40  
41  
42  
43  
44  
45  
46

47 The attention account highlights the 2AFC component, which by definition introduces a  
48 decision-making phase to the task before the reach is executed. According to this account,  
49 participants saccade between each side of the dual display during this phase, fixating on one  
50 inner disc or the other to determine whether their sizes look the same or different. Insofar as  
51 foveated (and perifoveal or central visual) areas of the display drive selective attention, the latter  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 will operate on both discs and their surrounding inducers, fostering primary and secondary  
4 sources of influence on the discs' perceived sizes: first, the effect of each inducer on the  
5 perceived size of its inner disc; and second, an enhancement of these single configuration  
6 effects via a size-contrast comparison of the two perceived sizes. The resulting effect on  
7 perceived disc size is stronger than the effect of each inducer alone [10,12].  
8  
9  
10  
11  
12  
13

14  
15  
16 Just as important to the attention argument is the execution phase of the grasp after one of the  
17 disc's has been selected. During the execution phase, while participant is reaching, gaze is  
18 assumed to fall preferentially on the target disc. This is believed to facilitate monitoring of the  
19 end of the reach when the hand contacts the disc and forms its grip around it (e.g., [13-16]).  
20 Thus, gaze should remain on the targeted disc-inducer configuration, to the exclusion of the  
21 other. This rationale extends to selective attention: the influence of the excluded disc and its  
22 inducer decays and so too does the size-contrast based enhancement of the illusion, leaving  
23 only the weaker single-configuration based effect to influence the hand's in-flight aperture.  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33

34  
35 For adherents to the attention-based account, the solution is a methodological one in which only  
36 one inducer and disc should be displayed per trial. This move is made in order to remove the  
37 enhanced, secondary effect of the dual display altogether, leaving the primary effect of the  
38 single configuration display intact. Following this move, these authors find that the effects of the  
39 illusion on perception and grasping are equivalent [10,11,17].  
40  
41  
42  
43  
44  
45  
46

47 An alternative method was later developed using dual-configured displays of the Muller-Lyer,  
48 Ponzo, and Wundt–Jastrow illusions, in which participants reach for (or estimate the sizes of)  
49 both illusory 3D targets using their left and right hands, with some authors replicating the  
50 perception-action dissociation [18-19] and others failing to do so when visual feedback is not  
51 available throughout the reach [20]. This design reasonably assumes that participants saccade  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 between both targets when reaching for them or estimating their sizes, and that comparisons  
4  
5 between the targets promote a synergistic, secondary, effect involving both inducers that would  
6  
7 register in both limbs if they were susceptible to it [20].  
8  
9

10  
11 Thus far, there is no controversy over whether selective attention influences the different effects  
12  
13 of the single vs. dual displays on visual perception. From the perspective of *scale attention* [21],  
14  
15 reducing the dual configuration to a single configuration removes the higher structural levels  
16  
17 (larger scale features) in the display that are normally available for selective attention to operate  
18  
19 on in service of scene- and ensemble-perception, which we hold as core functions of the visual  
20  
21 perceptual system. Without the higher-level structural features of the dual display, enhancement  
22  
23 from the secondary sources is not possible, leaving only the primary (inducer and its inner disc)  
24  
25 sources to influence perceived size.  
26  
27  
28  
29

30  
31 Nevertheless, as far as are aware, the attention account makes no claims about gaze or the  
32  
33 deployment of attention when participants manually estimate the size of the chosen disc. In  
34  
35 order to explain why grasp aperture resists the dual display while the manual estimates of disc  
36  
37 size do not, the attention account must group the manual estimation and *size-comparative*  
38  
39 judgements together. This account compels its advocates to make the straight-forward  
40  
41 prediction that gaze, *as a proxy for selective attention*, switches between the two configurations  
42  
43 more during manual estimation than it does during grasping – and by doing so maintains (or  
44  
45 iteratively re-kindles) the enhanced effect of the dual display.  
46  
47  
48

49  
50 Our position, speculative as it is, stands in contrast to that of the attention account. According to  
51  
52 our view, the manual estimation task requires participants to focus on one target disc to the  
53  
54 exclusion of the other in order to reliably reproduce the perceived size of the target disc. This  
55  
56 account predicts that gaze should behave similarly across the grasping and manual estimation  
57  
58  
59  
60



1  
2  
3 task, particularly during the response. Specifically, gaze will be largely directed at the target to  
4 facilitate extraction of its size in order to respond accordingly.  
5  
6  
7  
8

9 The current experiment was designed to test these ideas, including the assumptions about gaze  
10 patterns across different tasks. Briefly, we asked participants to make judgments of relative size  
11 across four different tasks, two of which assumed 2AFC format. As was done in Aglioti et al. [1],  
12 we first determined the difference in the discs' real sizes required for a perceived-as-same  
13 judgment for each participant using a verbal "same/different" 2AFC task. We referred to this  
14 difference as the *a priori* effect of the illusion, and it is based on the display in which the sizes of  
15 the discs look the same. Next, we tracked gaze in three additional tasks while using this  
16 different-sized pair of discs to set up two versions of the display: one in which the discs looked  
17 the same size (i.e., the 'perceived-as-same' display); and a second in which either two of the  
18 smaller discs or two of the larger ones were used so that the sizes of the discs looked different  
19 (i.e., the 'perceived-as-different' display). In the remaining three tasks, participants either made  
20 2AFC keypress judgments of the relative size of the discs, reached for the discs to pick them  
21 up, or manually estimated their sizes. In the latter two tasks, we tracked the hand using a  
22 motion-capture system. The 2AFC keypress judgments allowed us to observe baseline  
23 'decision-making' preparatory patterns of gaze.  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42

43 In line with the attention-based account during decision (*preparatory*) phase of the response, we  
44 expected gaze to operate similarly across the three tasks. During the *execution* phase,  
45 however, we expected gaze would cluster more, in terms of number and time, towards the  
46 chosen side of the display for both the grasps and manual estimates. Meanwhile, the effect of  
47 Ebbinghaus display was predicted to dissociate across the two tasks. Thus, our account  
48 emphasizes the behavioural endpoints of the task and the operations that support them. In  
49 short, it is what must be done with selected target information that matters most [4,22].  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## Materials and Methods

### *Participants*

Twenty-four adult participants (mean years of age was 21; 14 female) recruited from the student population at Western University took part in the study. All were naïve to the purpose of the experiment. Participants self-reported right-handed with normal- or corrected to normal vision, and written informed consent was obtained prior to testing. Furthermore, they received monetary compensation for their time. All procedures were approved by the Research Ethics Board of the University of Western Ontario and were carried out in accordance with the *Tri-Council Policy Statement Ethical Conduct for Research Involving Humans* [23].

### *Stimuli and Apparatus*

The workspace was comprised of a desk with an LCD monitor tilted ~10 degrees from horizontal and a chinrest. The monitor was tilted in order to provide a more comfortable view for the participant while approximating a perpendicular viewing angle for the head-mounted gaze-tracking equipment. For the experimental tasks, the Ebbinghaus display was placed on top of the LCD screen.

As shown in Fig. 1, the display consisted of a white board (38 cm wide x 22.5 cm in height) with two rings of matte black-filled circles, or *inducers*: one to the left of the display and one 13 cm to its right. Each inducer was centred at the midpoint of the display's height (~11.3 cm). For the display as oriented in Fig. 1, the centre of the inner disc on the left was 10.5 cm from the display's leftmost edge, whereas the centre of the inner disc on the right was 14.5 cm from the display's rightmost edge. The inducers on one side of the display comprised a ring of 11 black-filled circles 1-cm in diameter. Each of these small circles were spaced evenly apart from their immediately neighbour and centred 2.5 cm from the centre of the ring. The inducer on the other

1  
2  
3 side of the display comprised a ring of 6 black-filled circles each 5.8 cm in diameter. These large  
4 circles were spaced evenly apart from their immediately neighbours and centred 6 cm from the  
5 centre of the inducer. Real 3D discs were positioned at the centres of each inducer, creating an  
6 illusion of size. The discs were from two sets of five discs each 3-mm thick and painted matte  
7 black. The smallest disc used in each set was 3 cm in diameter while the largest was 3.5 cm.  
8  
9  
10  
11  
12

13  
14  
15 An OptoTRAK optoelectronic motion tracking system (NDI, Waterloo, Canada) was used to  
16 track hand movement at 100 Hz. An EyeLink II (SR Research Ltd., Mississauga, Canada)  
17 system was used to record the position of the right eye at a sampling rate of 250 Hz. A Logitech  
18 keypad was used in the keypress task to track participants' choice and response or *reaction*  
19 *time*, and a button-release device was used in the manual estimation and grasping tasks to  
20 track reaction time in these tasks. These were positioned on the table to the right of the display  
21 and were elevated to approximate the height of the display. One red LED light was used as a  
22 resting fixation point between trials, while a second LED light was used to illuminate the  
23 workspace to signal the start of the trial (see Supplementary Material Figure 1).  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36

### 37 *Experimental Design*

38  
39 Four tasks were administered. The first of these tasks was always verbal forced-choice ("same"  
40 or "different") judgments of relative disc size without movement recording equipment. [This was a](#)  
41 [semi-formal session used to determine the \*a priori\* effect of the illusion. In this session,](#)  
42 [participants were asked to say whether the disc sizes looked the same or different after each](#)  
43 [paired presentation. One of the discs was always 3cm in size, and the size of the other disc was](#)  
44 [varied in order to find a difference that elicited a reliable "same" response. The 3cm disc and its](#)  
45 [different-sized partner, the latter's size was specific for each participant, were presented in two](#)  
46 [ways in the remaining three tasks: First, for the perceived-as-same display, the physically larger](#)  
47 [disc was placed at the centre of the large-circles inducer, while the physically smaller disc was](#)  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 placed at the centre of the small-circles inducer; Second, for the perceived-as-different display,  
4 the discs were either the smaller or the larger partners of the pair.  
5  
6  
7  
8

9 Following the initial verbal 2AFC session, three tasks, 2AFC keypress, grasping, and manual  
10 estimation, were administered in a random order to each participant in a blocked-trial format.  
11 Eye-tracking was performed for all tasks, while hand movement tracking was performed in the  
12 grasping and manual estimation tasks. Half of the participants, selected at random, were  
13 provided with the following response rule: “if the sizes of the centre discs look the same, choose  
14 the disc on your left; if the sizes of the centre discs look different, choose the disc on your right”.  
15 For the remaining participants, the rule merely reversed the left/right contingency. In order to  
16 minimize rule confusion, each participant applied only one version of the two rules throughout  
17 the experiment. Furthermore, the display was flipped 180 degrees on half of the trials in a  
18 randomly interleaved format in order to minimize the possibility that participants would explicitly  
19 work-out the real sizes of the discs or, through motor learning or adaptation, calibrate their  
20 grasp aperture to their real sizes [24-27], (see the Supplementary Materials for additional  
21 details).  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38

39 In the grasping and manual estimation tasks, the participants expressed their judgement  
40 verbally: “same” for discs judged to look the same in size; and “different” for discs judged to look  
41 different in size. This was done to preserve the verbal component for these two tasks, and to  
42 make the disc choice for the manual estimation task explicit. In the grasping task, participants  
43 reached for the chosen disc to pick it up and put it back down. In the manual estimation task,  
44 participants indicated the perceived size of the chosen disc by opening their index finger and  
45 thumb to create a matching-sized gap, the manual estimate aperture. In both the grasping and  
46 manual estimation task, participants were encouraged to make their verbal judgment before  
47 initiating their hand movement. In general, participants adhered to this instruction. However,  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 without voice recording to make a formal evaluation, it is possible their brief “same”/“different”  
4  
5 verbal response overlapped with movement initiation on some trials.  
6  
7

8  
9 Each task comprised 24 trials. This number reflected a trade-off between sample stability and  
10  
11 the participant’s well-being, as we discovered during piloting the project that the head-mounted  
12  
13 eye-tracking unit was cumbersome for some participants, whose discomfort increased over  
14  
15 time. The 24 trials for the each of the keypress, grasping, and size-estimation tasks were  
16  
17 organized into 12 perceived-as-same trials and 12 perceived-as-different trials. On six of the 12  
18  
19 perceived-as-different trials, the smaller partner of the disc pair was presented in both inducers,  
20  
21 and on the remaining six trials, the larger partner was presented in both inducers. With board-  
22  
23 orientation flipping, the target was on the left for six of the 12 perceived-as-same trials and six of  
24  
25 the 12 perceived-as-different trials, and on the right for the remaining six trials of these two trial  
26  
27 types. The presentation order of these conditions was randomized.  
28  
29

### 30 31 32 *Procedure*

33  
34 The verbal judgment task was administered using a staircase method to determine the *a priori*  
35  
36 effect (see the Supplementary Materials for additional details). Participants were presented with  
37  
38 pairs of different-sized discs and asked to say whether they looked the same or different. The  
39  
40 experimenter began with no size difference and increased the size difference incrementally on  
41  
42 each trial by keeping one disc the same size and altering the size of the other. After a “same”  
43  
44 response was elicited, the experimenter started with the maximum difference in disc size (5mm)  
45  
46 and reduced it incrementally on each trial in a similar fashion as described above.  
47  
48  
49

50  
51 After the initial session, the Eye-Link II was calibrated using its native nine-point procedure  
52  
53 before the start of each of the three remaining tasks. For the three tasks, each trial began with  
54  
55 an auditory cue for participants to fixate on a red LED light located above the illusory display.  
56  
57  
58

1  
2  
3 Not long after, a second auditory cue coincided with the illumination of the workspace and cued  
4 participants to scan the display, make and announce their choice (in the case of the manual  
5 estimates and grasps) and then execute their response (see Supplementary Figure 1).  
6  
7  
8  
9

10  
11 For the grasping and manual estimation tasks, participants held down a button on a portable  
12 button box with their right index finger and thumb pinched together. Participants were instructed  
13 that the button was the hand's start and end location for these two tasks. The keypad for the  
14 keypress task was positioned at approximately the same location as the button box.  
15  
16  
17  
18  
19

20  
21  
22 In the grasping task, participants reached for their chosen disc, lifted it up using their index  
23 finger and thumb, and put it back down before returning to the home position. In the manual  
24 estimation task, participants spread their thumb and index finger apart to create a gap that  
25 matched the disc's perceived size. The participants were also instructed to refrain from reaching  
26 for the disc in this task. After they were finished with their estimate, they returned to the start  
27 button at which point the illuminator was switched on again, cueing the participants to reach for  
28 the previously chosen disc. This allowed the participants the same haptic feedback about the  
29 disc's real size that was available for the grasping task. After this, the participants returned their  
30 hand to the start button in preparation for the next trial. In the keypress task, participants  
31 indicated their disc choice by keypress on the Logitech keypad.  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44

#### 45 *Eye and hand tracking*

46  
47  
48

49 Eye movements were recorded for five seconds from the onset of each trial.

50  
51 The 3D positions of the right hand were recorded using an OPTOTRAK system (Northern  
52 Digital, Waterloo, ON, Canada) sampling the two IREDs on the participants hand at a rate of  
53  
54 100 Hz for a minimum of five seconds from the onset of each trial.  
55  
56  
57  
58  
59  
60

### *Data Pre-processing*

The limb kinematic data comprised of 3D positions of each IRED for each sample frame. This data was pre-processed and analysed in a standardized way to extract a number of kinematic landmarks (e.g., [28], see Supplementary Materials for additional details). Crucially, grasp aperture was computed as the distance between the index finger and thumb IREDs at each sample frame. The principal dependent measure for the grasps was peak grasp aperture (PGA), which was defined as the largest aperture achieved from the start of the movement to the end of the reach.

Like grasp aperture, the *estimate aperture* was computed as the difference between the index finger and thumb IREDs at each sample frame for the manual estimation task. The final estimate aperture (FEA) was determined as the point at which the estimate aperture's rate of change stabilized around zero, signalling that a relatively consistent estimate had been achieved (see Supplementary Materials for the definition and analysis of the peak estimate aperture, which did not meaningfully influence the results).

Trials in which participants gave an 'incorrect' judgment (e.g., a 'same' judgment when discs that should have been perceived as different were displayed) were removed from the analysis (8% of trials). This left a uniform sample of trials (92%) in which disc choice accorded with the intended effect of the illusion. Furthermore, all fixations shorter than 80ms in duration were discarded, as this corresponds to the minimum duration for higher level processing of visual information to occur (e.g., [29]). Fixations that fell outside the display area were also removed.

1  
2  
3 We applied cluster analysis to the gaze data (x and y values in units of screen-width and  
4 screen-height pixels, respectively) to identify, in a data-driven manner, four clusters  
5  
6 corresponding to the four positions of the target inner disc (across trials), arranged left to right,  
7  
8 and to assign each fixation to one of these four clusters. With the display oriented as depicted in  
9  
10 Fig. 1 (i.e., the small inducer on the left side), the clusters corresponding to the two  
11  
12 configurations were the leftmost and third cluster from the left. With the display flipped 180  
13  
14 degrees (i.e., the large inducer on the left side), the corresponding clusters were second from  
15  
16 the left and the rightmost cluster. Assigning fixations to each cluster allowed us to quantify the  
17  
18 following dependent variables: (1) The number of display-wise saccades per second, which was  
19  
20 defined as the number of instances in which temporally adjacent fixations left one cluster for  
21  
22 another where the second cluster skipped one or two adjacent clusters. For example, suppose  
23  
24 gaze falls at the leftmost cluster and then falls on either the third cluster from the left or the  
25  
26 rightmost one, then this transition would be counted as a display-wise saccade; (2) The number  
27  
28 of fixations on the target disc. Note that the target disc was the disc chosen in accordance with  
29  
30 the illusion; (3) the fixation time on the target disc; (4) the percentage of fixations on the target  
31  
32 disc (out of the total number of fixations); and (5) the percentage of fixation time spent on the  
33  
34 target disc (as a percentage of the total number of fixations).  
35  
36  
37  
38  
39  
40

41 For the fixation (gaze) data, the button release time (the reaction time, RT) was used to  
42  
43 separate the preparatory phase from the execution phase. For the grasps, estimates, and  
44  
45 keypress tasks, the preparatory phase began when the illuminating light was turned on. The  
46  
47 preparatory phase ended when the button was released for the grasps and estimates, or when  
48  
49 the keypress made for the keypress task. The execution phase for the grasps and manual  
50  
51 estimates began with the release of the start button. For the estimates, the end of the execution  
52  
53 phase was defined as the point in time at which the final estimate aperture was achieved. For  
54  
55  
56  
57  
58  
59  
60



1  
2  
3 the grasps, the end of the execution was defined as the point at which the reach ended. The  
4  
5 keypress task was considered to be comprised of preparatory time only.  
6  
7

### 8 9 *Statistical Analysis*

10  
11 For each combination of participant, task, and dependent measure, means were computed from  
12  
13 trials grouped into combinations of the target disc's size (small vs. large) and surrounding  
14  
15 inducer (small or large) and whether the display was setup to make the sizes of the inner discs  
16  
17 look the same or different (i.e., *display type*).  
18  
19

20  
21  
22 For the perceived-as-different display, the effect of the illusion was computed for the grasps and  
23  
24 manual estimates as the mean dependent measure when the target disc was surrounded by the  
25  
26 small inducer (i.e., when the disc was perceived as larger than its actual size) minus the mean  
27  
28 dependent measure when the target disc was surrounded by the large inducer (i.e., when the  
29  
30 disc was perceived as smaller than its actual size). Positive difference values therefore reflect  
31  
32 the expected effect of the illusion. Note that for the perceived-as-different trials, target disc size  
33  
34 varied independently of the surrounding inducer, and so the effect on each disc size in this  
35  
36 condition was collapsed for brevity and to boost sample stability. For completeness, t-tests for  
37  
38 an influence of disc size on the effect of the display in this condition were null for the grasps,  
39  
40  $t(18) = .34, p > .73$  and for the manual estimates,  $t(18) = .95, p > .35$ .  
41  
42  
43  
44

45 The effect of the perceived-as-same display requires one additional step, because the above-  
46  
47 mentioned difference formula would yield a value of zero for the estimates and the grasps; recall  
48  
49 that the sizes of the discs are perceived to be the same. Recall further that according to the null  
50  
51 hypothesis, both the estimates of disc size and the grasp aperture for the grasps are faithful  
52  
53 'readouts' of perceived size. The solution, for us, was to add the *a priori* effect (i.e., the  
54  
55 difference in disc size required to make the discs look the same size) to the result of the above-  
56  
57  
58  
59  
60

1  
2  
3 mentioned formula for both the grasps and the manual estimates (see the *Supplementary*  
4 *Material* for additional details).  
5  
6  
7

8  
9 To test the effects of the illusion, a repeated-measures ANOVA was performed on the  
10 participant means with the following factors: *task* (keypress, grasps, and manual estimates) and  
11 what we refer to as *display type* ('perceived-as-same' and 'perceived-as-different'  
12 arrangements). Type I error was set to 0.05 for each ANOVA. For each family of follow-up  
13 contrasts, we used the Holm step-down procedure [30] in combination with the moderately more  
14 powerful multiplicative Bonferroni inequality [31] (see also [32-33]). The Holm-Sidak procedure  
15 sequentially adjusts the per-contrast alpha rate for a family of tests while holding the type I error  
16 at 0.05 on a per-family basis; it is more powerful than the more commonly used Bonferroni  
17 procedure, yet maintains much of the latter's favoured simplicity.  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29

### 30 *Unadjusted and Adjusted Effects of the Illusion*

31  
32 Response sensitivity to differences in target size should be matched across tasks. When they  
33 are not, the difference can lead to spurious differences in effects across tasks when none exist.  
34 This holds for the effects of illusions across tasks [34]. A linear function is a good approximation  
35 of the relationship between peak grasp aperture (or the final estimate aperture) and target size.  
36 The crucial parameter is the coefficient relating differences in mean response to incremental  
37 changes in target size (i.e., the *slope*). The mean slopes indicate that the unadjusted effect of  
38 the illusion is overestimated for the manual estimates and underestimated for the grasps (*slope*  
39 *for the manual estimates*:  $M=1.15$ ,  $SEM=.104$ ,  $t(18)=11.12$ ,  $p<2\times 10^{-9}$ , *slope for the grasps*:  
40  $M=.84$ ,  $SEM=.102$ ,  $t(18)=8.45$ ,  $p<2\times 10^{-7}$ ). Moreover, the mean slopes differ significantly,  
41  $t(18)=2.44$ ,  $p<.03$ . Crucially, the mean illusory effect can be adjusted to compensate for this  
42 difference. This is done by dividing the mean unadjusted effect by the mean slope. However,  
43 this procedure neglects the variance of the slope and leaves the variance of the ratio  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 unspecified [35]. Fieller's theorem [36] provides exact confidence intervals for the ratio of two  
4 random variables, and a Taylor approximation has been shown to perform well when the slopes  
5 differ highly significantly from zero [37], as is the case here. The Taylor approximation is  
6 advantageous as it can be rearranged to provide an approximation for each participant, which is  
7 useful for performing statistical procedures on samples [38]. Thus, we adjusted the effects of the  
8 illusion on the grasps and manual estimates using this method, as recommended by Franz and  
9 colleagues (e.g., [37-38]). This adjustment did not affect the conclusions.

## 22 Results

### 23 *A Priori Effect Size and Choice Accuracy*

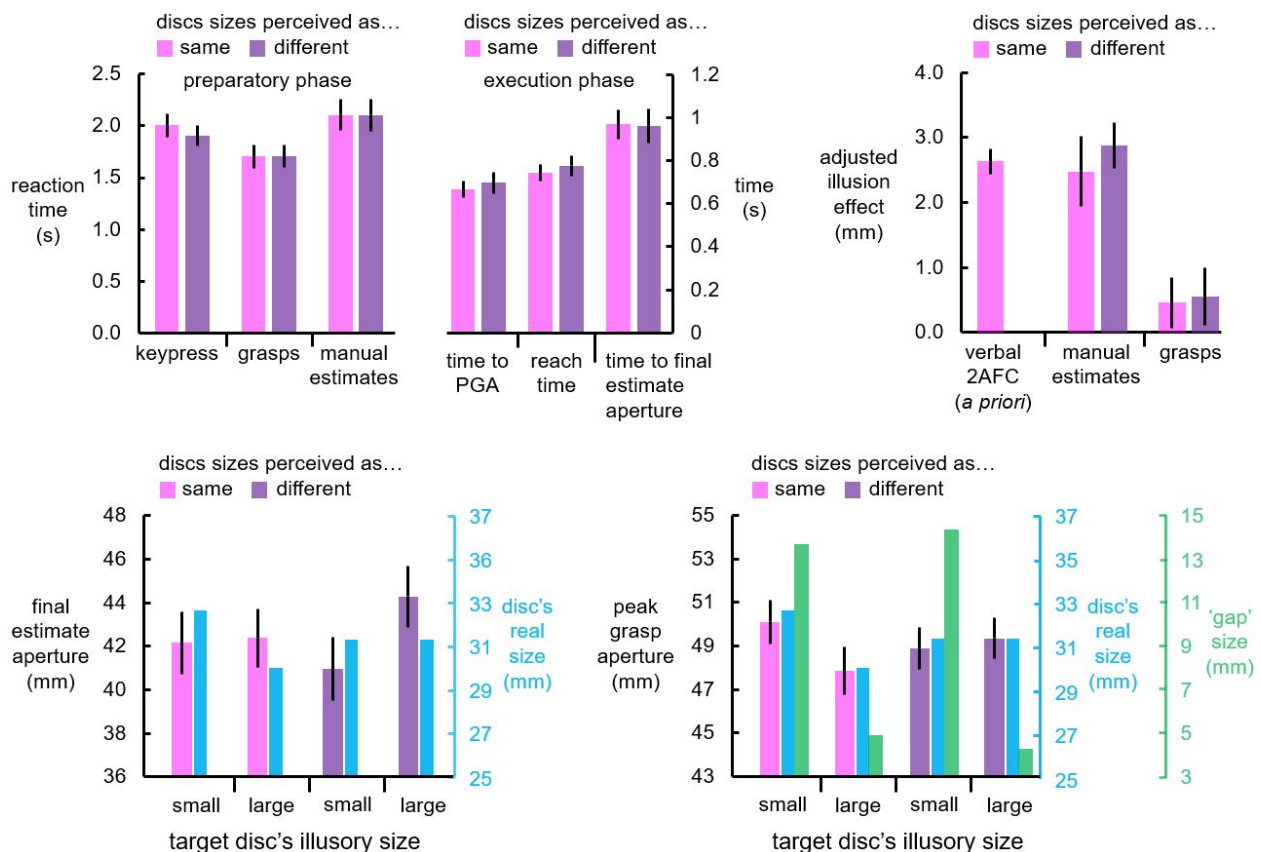
24 The mean difference in real size that resulted in reliable "same" size judgements was 2.6mm (1  
25 -4mm). Overall, on average, 92% of participants' choices matched the expected effect of the  
26 display. Furthermore, this level of performance was not modulated by task ( $F(2,34)=.28, p>.74$ ),  
27 the display type ( $F(1,18)=2.48, p>.13$ ), or the interaction effect,  $F(2,33) = 2.3, p>.12$ .

### 36 *Reaction Times (preparatory phase)*

37 The top left panel of Figure 2 ('preparatory phase') shows mean reaction time as a function of  
38 task and display type. Mean reaction time was influenced by task ( $F(2,32)=5.72, p<.01, \eta$   
39  $p^2=.24$ ), but not by the display type,  $F(1,18)=.35, p>.55$ . The interaction effect was null,  
40  $F(2,27)=2.47, p>.11$ . The grasps were faster to initiate than either the manual estimates  
41 ( $t(18)=3.44, p<.003$ ) or the 2AFC keypress responses ( $t(18)=2.58, p<.02$ ); the reaction time for  
42 the latter two did not differ statistically,  $t(18)=.93, p>.36$ .

### 53 *Grasp and Estimate Temporal Measures (execution phase)*

The top left panel of Figure 2 ('execution phase') shows the time to peak grip aperture (*PGA*) and reach completion time (*reach time*) for the grasps, and the time to final estimate aperture for the manual estimates. These measures varied with task ( $F(1,18)=63.56$ ,  $p<3\times 10^{-7}$ ,  $\eta_p^2=.78$ ) but not with display type ( $F(1,18)=.43$ ,  $p>.51$ ). The interaction was null,  $F(2,19)=.87$ ,  $p>.36$ . Pairwise comparisons were all significant (all  $ps<.007$ ), such that the time to peak grasp aperture occurred earliest, followed by the time to complete the reach, and then time to complete the estimate of target size.



**Figure 2.** Top left: Preparatory and execution phase temporal dependent measures as a function of task and display type. The grasps were initiated the fastest, and the mean initiation time for the estimates and the mean keypress time for the keypress task were statistically equivalent. With respect to the execution phase, the reach time for the grasps was faster than the final estimate aperture time. Top right: the *a priori* effect of the Ebbinghaus illusion, using the 'perceived-as-same' display (established in the initial verbal 2AFC task), and the adjusted effects of the Ebbinghaus illusion on the manual estimates (middle) and grasps (right) as a function of display type. The effect of the illusion on the estimates was comparable to the *a priori* effect, whereas the effect of the display on the grasps was null. Bottom panels: *Peak grasp aperture* for the grasps (left) and *final estimate aperture* for the manual estimates (right) as functions of the direction of the effect of the illusion ("small" – the inducer decreases the perceived size of the target; "large" – the inducer increases the perceived size of the disc) and the display type (i.e., the perceived-as-

1  
2  
3 same and perceived-as-different arrangements). The blue bars depict real disc size for the 'perceived-as-  
4 same' display, whereas they depict average disc size for the 'perceived-as-different' display. The average  
5 is used for real disc size (and for the illusion effect) because the two are independent of one another for  
6 these trials. Note that the blue bars possess their own ordinate axis to the right of each panel. The green  
7 bars in the bottom right panel reflect the mean distance between the outside edge of the target disc and  
8 the inner-most edges of tee surrounding inducer. The green bars possess their own ordinate axis to the  
9 right of this panel. Importantly, while all ordinate axes possess different minimum and maximum values,  
10 they share the same units and span the same distance (12 mm). The figure indicates that the illusion  
11 drives the estimates aperture but that real-disc size drives the peak grasp aperture. Error bars are SEMs.  
12

### 13 14 *Illusory Effects on Grasp Aperture and Final Estimate Aperture*

15  
16 The effect of the Ebbinghaus illusion established during the initial verbal 2AFC task, an effect  
17 we termed the *a priori* effect, is shown in the top-right panel of Figure 2. To the right of the *a*  
18 *priori* effect are the adjusted effects of the illusion on the final estimate aperture of the manual  
19 estimates (henceforth *manual estimates*) and on the peak grasp aperture (henceforth *grasps*),  
20 respectively (see Supplementary Figure 3 for both the unadjusted and adjusted illusory effects).  
21  
22 This panel bears three findings that are supported more formally below: (1) the similarity of size  
23 of the illusion effect in the two perceptual tasks: the verbal 2AFC and the manual estimates in  
24 the perceived-as-same display; (2) the difference between these perceptual effects and the null  
25 effect on the grasps for this display; and (3) the replication of these two points for the perceived-  
26 as-different display.  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38

39 The effect of the illusion on the manual estimates did not vary as a function of display type,  
40  $t(18)=.93, p>.36$ . An analogous test applied to the grasps was also null,  $t(18)=.16, p>.87$ . For  
41 the perceived-as-same display, the *a priori* effect of the illusion was significantly greater than the  
42 effect of the illusion on the grasps ( $t(18)=4.75, p<2 \times 10^{-4}$ ) but not significantly different from the  
43 effect of the illusion on the manual estimates,  $t(18)=.37, p>.71$ . Furthermore, the effect on the  
44 manual estimates was significantly greater than the effect on the grasps,  $t(18)=3.11, p<.006$ .  
45  
46 For the perceived-as-different display, the effect of the illusion on the manual estimates was  
47 significantly greater than the effect on the grasps,  $t(18)=5.42, p<4 \times 10^{-5}$ . Thus, both illusory  
48 arrangements consistently affected estimates of disc size but not the hand's grasp aperture.  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3  
4  
5 The bottom panels of Figure 2 reinforce the key takeaways from Figure 2's top-right panel. The  
6 bottom panels depict the mean final estimate aperture (bottom left) and the mean peak grasp  
7 aperture (bottom right) as functions of the direction of the effect of the illusion and the display  
8 type. Notably, regardless of the display type, the final estimate aperture tracks the perceived  
9 sizes of the discs; In contrast, the peak grasp aperture tracks the real sizes of the discs.  
10  
11  
12  
13  
14  
15  
16  
17

18 The bottom right panel of Figure 2 depicts mean peak grasp aperture as a function of the  
19 display type and the illusory size of the disc. In this panel, green bars depict the sizes of the  
20 gaps between the outside edge of the inner discs and the inside edge of the inducer. This panel  
21 suggests that peak grasp aperture is not influenced by the size of the gaps between the inner  
22 edges of the inducer and the outer edge of the inducer's inner disc. According to one account of  
23 the perception-action dissociation, the inducers are treated by the visuomotor system as  
24 obstacles the fingers must avoid when reaching for the target disc [39-40]. This avoidance  
25 translates into an influence on peak grasp aperture ([40-41], c.f., [26,42]). The original  
26 formulation of this idea suggested that the narrower the gap, the smaller the peak grasp  
27 aperture. Furthermore, because gap size is often smaller for the small inducers than for the  
28 large ones and because the small inducers increase the perceived size of the disc (whereas the  
29 large inducers decrease the perceived size of the disc but typically possess a larger gap), the  
30 obstacle avoidance system is thought to operate in opposition to the effect of the illusion,  
31 resulting in a reduced or null effect on grasps only. According to this account, the obstacle  
32 avoidance system masks a real effect of the illusion on the grasps.  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50

51 A formal test for influence of gap-size on grasp aperture was couched in terms of whether the  
52 reduction in gap size that occurs moving from the large inducers to the small ones affected peak  
53 grasp aperture *consistently* across the 'perceived-as-same' and 'perceived-as-different'  
54  
55  
56  
57  
58  
59  
60

displays. In line with the obstacle-avoidance account, peak grasp aperture was significantly smaller for the small gap than for the large one for the perceived-as-same display,  $t(18)=6.22$ ,  $p<8\times 10^{-6}$ . Crucially, however, peak grasp aperture did *not* vary as a function of the reduction in gap size for the perceived-as-different display,  $t(18)=-1.35$ ,  $p>0.19$ . Thus, the reduction in gap-size failed to reduce peak grasp aperture across both display types. Since the illusion did not influence peak grasp aperture across both display types, we are left with the real difference in size as the lone explanatory factor.

Overall, the analysis of peak grasp aperture shows that this measure tracks the real sizes of the discs, rather than the perceived size or the size of the gap between the outside of the disc and the inner edges of the surrounding annuli. The manual estimates, in stark contrast, track the perceived sizes of the discs.

### Fixations

Table 1 lists the five fixation-based measures, the ANOVA tests of the main effect of response type, and follow-up pair-wise comparisons for the preparatory time frame and then the preparatory and execution time frames combined.

**Table 1.** Fixation Measures, ANOVA tests performed on them ( $df_1/df_2=2/36$ , then Greenhouse-Geisser adjusted), and subsequent pair-wise comparisons ( $df=18$ ). Bolded entries indicate statistically significant tests.

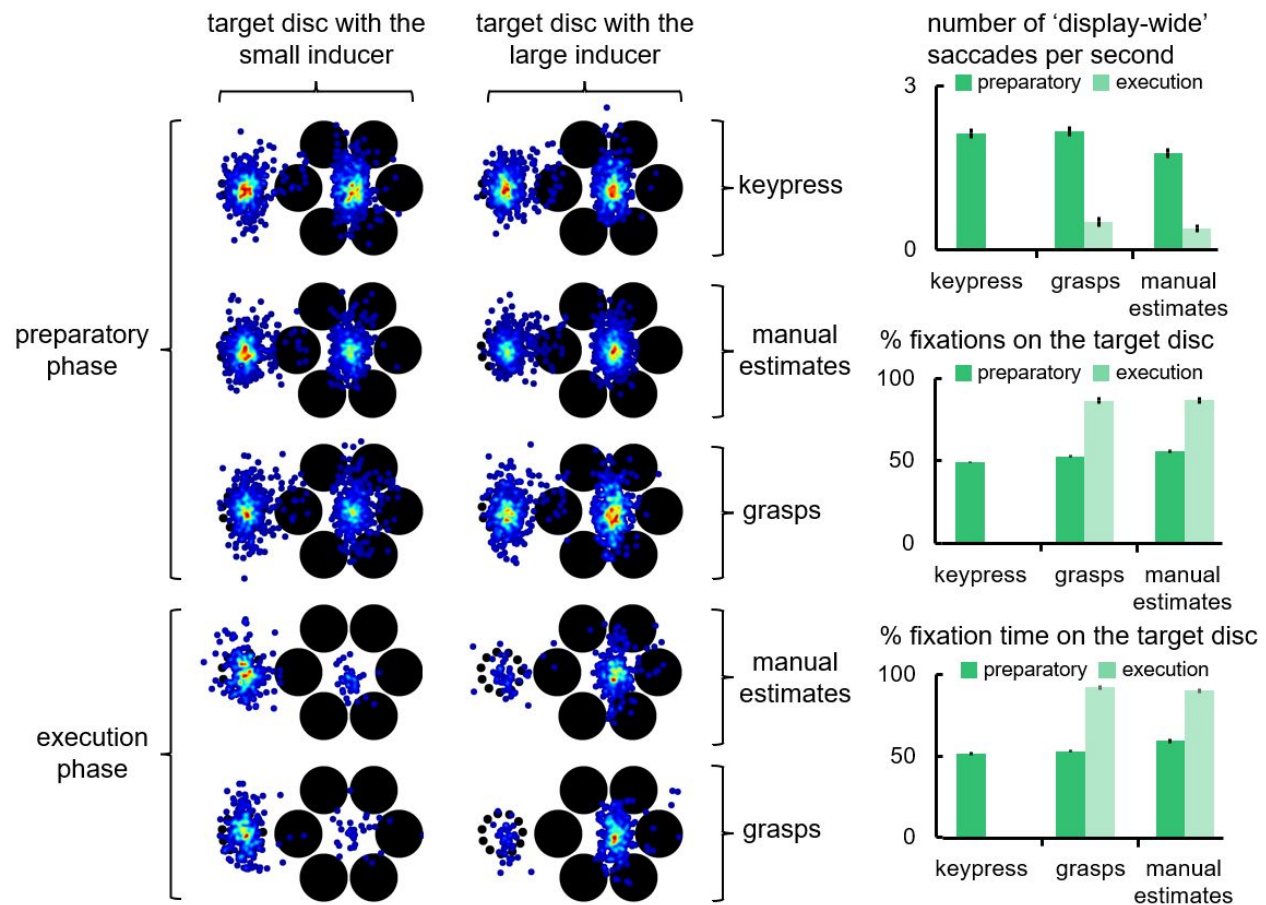
<u>Measure &amp; ANOVA statistic</u>	<u>Response Phase</u>	<u>Pair-wise t-test statistics</u>	
Number of display-wise saccades per second $F=16.83$ , $p<3\times 10^{-5}$ , $\eta_p^2=.48$	Preparatory	G > KP:	$t=.34$ , $p>.73$
		<b>KP &gt; M:</b>	<b><math>t=5.25</math>, <math>p&lt;6\times 10^{-5}</math></b>
		<b>G &gt; M:</b>	<b><math>t=5.42</math>, <math>p&lt;4\times 10^{-5}</math></b>
	Execution	G > M:	$t=1.42$ , $p>.17$
Number of fixations on the	Preparatory	<b>G &gt; KP:</b>	<b><math>t=2.14</math>, <math>p&lt;.05</math></b>

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

target disc $F=10.49, p<.002, \eta_p^2=.37$		<b>M &gt; KP:</b> <b>M &gt; G:</b>	<b><math>t=3.94, p&lt;.001</math></b> <b><math>t=2.85, p&lt;.02</math></b>
	Execution	G > M:	$t=0.4, p>.69$
Fixation time on the target disc $F=17.27, p<2 \times 10^{-5}, \eta_p^2=.49$	Preparatory	KP > G: <b>M &gt; KP:</b> <b>M &gt; G:</b>	$t=.52, p>.6$ <b><math>t=4.75, p&lt;2 \times 10^{-4}</math></b> <b><math>t=4.68, p&lt;2 \times 10^{-4}</math></b>
	Execution	<b>M &gt; G:</b>	<b><math>t=2.34, p&lt;.04</math></b>
Fixations on the target disc as % of total $F=21.99, p<1 \times 10^{-6}, \eta_p^2=.55$	Preparatory	<b>G &gt; KP:</b> <b>M &gt; KP:</b> <b>M &gt; G:</b>	<b><math>t=3.79, p&lt;.002</math></b> <b><math>t=6.4, p&lt;6 \times 10^{-6}</math></b> <b><math>t=3.06, p&lt;.007</math></b>
	Execution	M > G:	$t=.12, p>.9$
Fixation time on the target disc as % of total $F=24.89, p<1 \times 10^{-6}, \eta_p^2=.58$	Preparatory	G > KP: <b>M &gt; KP:</b> <b>M &gt; G:</b>	$t=1.61, p>.12$ <b><math>t=6.77, p&lt;3 \times 10^{-6}</math></b> <b><math>t=4.58, p&lt;3 \times 10^{-4}</math></b>
	Execution	G > M:	$t=1.49, p>.15$

The spatial distribution of fixations as a function of the task and the phase of the response, along with the mean fixation measures as a function of the task and the preparatory and execution phases of the response are depicted in Figure 3. Furthermore, because the reaction times varied as a function of the response type, and because the execution time differed between the grasps and manual estimates, Figure 3 was limited to the time-standardized fixation measures: the number of display-wise saccades per second; the percentage of fixations in the target area (of the total number of fixations); and the percentage of total fixation time spent in the target area (see Supplementary Figure 4 for the number of fixations on the target and the fixation time on the target).





**Figure 3.** The spatial distribution of fixations and the fixation measures as functions of task and the response phase (i.e., the preparatory and execution phases). Left: gaze maps show the spatial distribution of fixations parsed into trials in which the target disc was surrounded by the small or large inducer. Display orientation is collapsed such that fixations on those trials with the opposite board orientation to the one depicted were flipped across the central vertical axis of the full set of fixations. Top right: The number of 'display-wide' saccades (gaze shifts between disc positions two or more locations apart from one another) per second. Middle right: The percentage of total fixations that were directed at the target disc. Bottom right: The percentage of total fixation time that was directed at the target disc and inducer (right). During the preparatory phase, participants made more saccades between the discs, fixated on the target disc less, and spent less fixation time on the target, relative to the execution phase. Moreover, none of the fixation measures during the preparatory phase differentiate the keypress and manual estimation tasks from the grasping task. During the execution phase, the fixation measures fail to differentiate the grasps and manual estimates. Error bars are SEMs.

Overall, the fixation measures well-differentiated the tasks. However, for the time-controlled measures, the grasps and estimation tasks looked more similar than the keypress task. Notably, the attention account suggests that longer fixation time on the target weakens the illusion, because it reflects selective attention operating on the fixated target and its inducer to the

1  
2  
3 exclusion of the other disc and its inducer. By that same token, fixations distributed across both  
4 sides of the display should strengthen the illusion, as this pattern of gaze suggests selective  
5 attention is operating on both discs and their inducers. Given that the verbal 2AFC and manual  
6 estimation tasks yielded the largest illusion effects while the grasps remained unaffected, the  
7 attention account suggests that the fixations made during the two perceptual tasks should be  
8 distributed across both sides of the display while the fixations made when grasping should be  
9 anchored to the target. Crucially, we did not find this.  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19

20 Take, for example, the mean number of display-wise saccades per second. During the  
21 preparatory phase, this measure was lowest for the estimations. The attention- based account  
22 predicts that the keypress and estimates should exhibit higher mean number of display-wise  
23 saccades per second, while the grasps should exhibit the lowest mean number of display-wise  
24 saccades per second. Furthermore, the keypress and manual estimation means for the  
25 remaining preparatory phase time-standardized fixation measures were most dissimilar, with the  
26 mean values for the grasps typically falling in-between those of the keypress and estimation  
27 tasks (see Figure 3). Specifically, the mean fixation time on the target disc as a function of total  
28 fixation time was largest for the manual estimates, intermediate for the grasps, and least for  
29 keypress task (see Table 1); and the mean number of fixations on the target disc as a  
30 percentage of the total number of fixations was largest for the manual estimates, and least for  
31 the grasps and keypress tasks, the latter two were statistically indistinguishable (see Table 1).  
32 Thus, during the preparatory phase, the results indicate that gaze was more likely to shift  
33 between the discs and to engage the non-target disc during the keypress and grasping task  
34 than during the manual estimation task.  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53

54 The attention-based account assumes that gaze is directed preferentially to the target during  
55 grasp execution. We take no issue with this prediction. Moreover, the data support it. In fact,  
56  
57  
58  
59  
60

1  
2  
3 across all measures, gaze preference to the target increased significantly in the execution  
4 phase, relative to the preparatory phase, for the grasps and for the manual estimates (all  $p$ s <  
5  $2 \times 10^{-7}$ ). In each of the preparatory and execution phases, gaze time on the target was  
6 significantly greater for the manual estimates than for the grasps (see Table 1). This result  
7 stands in opposition to the attention account, given that the manual estimates, not peak grasp  
8 aperture, was affected by the illusion. Nevertheless, we find it more likely that this difference in  
9 time spent fixating the target merely stems from the increased preparatory and execution times  
10 for the estimates. This issue reinforces the importance of the measures of gaze that control for  
11 differences in preparatory and execution times. Indeed, the number of display-wide saccades  
12 per second, the number of fixations on the target as a percentage of total fixations, and the  
13 fixation time on the target as a percentage of total fixation time were each statistically  
14 indistinguishable between the grasps and manual estimates during the execution phase (see  
15 Table 1). Thus, the fixation analysis shows that gaze does not differentiate the two tasks in a  
16 way that can explain why the estimates faithfully expressed the misperceived sizes of the discs  
17 while the grasps expressed the real sizes of the discs.

## 37 Discussion

38 Contemporary theories of human vision hold as a given that the organization of primate visual  
39 system reflects the evolution of distinct networks of neural circuitry, each broadly suited for  
40 solving unique problems posed by different facets of adaptive behaviour (for reviews, see [5-  
41 6,43-44]). Visual perception entails the moment-to-moment construction of a mental model of  
42 what is 'out there' to serve as a kind of cognitive sandbox in which plans and goals can be  
43 formulated and behaviourally-relevant options derived, selected, and acted on. One function of  
44 visual perception is visual recognition which entails classifying and semantically elaborating on  
45 real-time visual sensory information with the assistance of stored representations. A core  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 challenge for visual perception is in achieving stable object constancies: that is, to filter out the  
4 viewpoint-dependency in the array to parse it and extract the core semantic relevancies to the  
5 organism. Solutions involve prioritizing information. The cognitive adjudication of numerous  
6 possibilities does not require the extraction of metric visual information. Put another way,  
7 selection for priority amongst competing possibilities requires relative comparisons, in which  
8 viewpoint dependencies and metric processing are discounted.  
9  
10  
11  
12  
13  
14  
15  
16  
17

18 When reaching for a goal object, however, viewpoint-dependent, metric information is critical.  
19 Here, the problem entails extracting the 3D object structure and representing its egocentric  
20 relationship to the viewer for executing the limb movement required for acquiring the goal  
21 object. The grasp points on the goal object must be suited for lifting and manipulating it, the  
22 hand's inflight grasp aperture must be large enough to accommodate the object while the arm  
23 movement must deliver the hand to it. As a consequence, neural circuitry has evolved for the  
24 visual control of object-directed actions that is distinct from that mediating the visual perception  
25 of objects [4,43]. Recent work has shown, for example, that grasp aperture and grasp angle are  
26 refractory to the mean size and orientation of an ensemble of distractors around a target object  
27 [45], and that under conditions of attentional crowding, grasp aperture remains tuned to target  
28 size while manual estimates of target size do not [46].  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42

43 Evidence from a number of different lines of work supports this duplex account of visual  
44 processing. One of the most controversial pieces of evidence comes from the apparent  
45 resistance of visually guided actions to familiar pictorial illusions, such as the Ebbinghaus and  
46 Ponzo illusion (for review, see [47]). It has also been argued, however, that the dissociation is  
47 not so much a difference in the neural circuitry mediating visual perception and the neural  
48 circuitry mediating visual control of action as it is a difference in the deployment of gaze (and  
49 selective attention) when judging an object's size vs. reaching out to pick it up. In the case of the  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 Ebbinghaus illusion, it has been postulated that fixations (selective attention) deployed across  
4 the full display enhance the contrast in apparent size between the two central discs that are  
5 already influenced by their surrounding annuli. When reaching out to pick up one of the central  
6 discs, however, fixations are restricted to that disc and its immediate vicinity, thereby reducing  
7 the illusory effect. In the current study, we systematically investigated the validity of this critique.  
8  
9  
10  
11  
12

13  
14  
15  
16 Using the original dual display paradigm, we replicated the finding that grasp aperture is  
17 refractory to the illusion even when one disc is chosen on the basis of differences in apparent  
18 size as well as the finding that manual estimates of disc size remain biased. Importantly, we  
19 tracked gaze while participants performed these tasks. Using gaze as a proxy for attention while  
20 participants performed three tasks: grasps, manual estimation, and a forced-choice keypress,  
21 we found that the patterns of fixations in terms of the number of display-wide saccades and the  
22 percentages of fixations and fixation time directed at the target disc and inducer, while  
23 differentiating the tasks did not do so in accordance with the effect of the illusion. Moreover,  
24 when the response execution phase was included, the fixations when grasping and when  
25 manually estimating disc size were statistically indistinguishable. In so far as gaze serves as a  
26 spatiotemporal marker of selective attention across the visual array, our findings suggest that  
27 selective attention to the target disc cannot explain the task-dependent effect of the display.  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42

43 One limitation of the current study is that we established an *a priori* effect using a  
44 “same”/“different” verbal 2AFC format and we did not measure the effect of the illusion on disc  
45 size during the “same”/“different” keypress version of the 2AFC task. Recall that gaze was  
46 tracked for the keypress version and not the verbal one. However, it is important to recall that  
47 the percentage of choices guided by the illusion was statistically equivalent across the keypress,  
48 grasps, and manual estimation tasks, and that the measured effect of the illusion on the  
49 estimates was no different than the *a priori* one, which was determined using “same/different”  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 verbal 2AFC. Thus, it seems reasonable to assume that the effect was operating just as strongly  
4 during the keypress task as it was during the verbal forced-choice and estimation tasks.  
5  
6  
7  
8

9 We conclude that key difference lies in the distinct problems that manually estimating the size of  
10 target and grasping the target pose for the visual system – and, as a consequence, the different  
11 computations that are required.  
12  
13  
14  
15  
16  
17  
18  
19

## 20 References

- 21 1. Aglioti S, DeSouza JF, Goodale, MA. 1995 Size-contrast illusions deceive the eye but not  
22 the hand. *Current biology* **5**(6), 679-685.
- 23 2. Haffenden AM, Goodale, MA. 1998 The effect of pictorial illusion on prehension and  
24 perception. *Journal of cognitive Neuroscience* **10**(1), 122-136.
- 25 3. de Haan HF, Cowey A. 2011 On the usefulness of ‘what’ and ‘where’ pathways in vision.  
26 *Trends in Cognitive Science* **15**(10), 460-466.
- 27 4. Goodale MA, Milner AD. 1992 Separate visual pathways for perception and action. *Trends*  
28 *in Neurosciences* **15**(1), 20–25.
- 29 5. Kravitz DJ, Saleem KS, Baker CI, Mishkin M. 2011 A new neural framework for  
30 visuospatial processing. *Nature Reviews Neuroscience* **12**, 217-230.
- 31 6. Kravitz DJ, Kadharbatcha SS, Baker CI, Ungerleider LG, Mishkin M. 2013 The ventral  
32 visual pathway: An expanded neural framework for the processing of object identity.  
33 *Trends in Cognitive Sciences* **17**, 26-49.
- 34 7. Rizzolatti G, Matelli M. 2003 Two different streams form the dorsal visual system: Anatomy  
35 and functions. *Experimental Brain Research* **153**, 146-157.
- 36 8. Smeets JBJ, Brenner E. 2006 10 years of illusions. *Journal of Experimental Psychology:*  
37 *Human Perception and Performance* **32**(6), 1501–1504.
- 38 9. Whitwell RL, Goodale MA, Merritt KE, Enns JT. 2018 The Sander parallelogram illusion  
39 dissociates action and perception despite control for the litany of past confounds. *Cortex*  
40 **98**, 163-176.
- 41 10. Franz VH, Gegenfurtner KR, Bühlhoff, HH, Fahle M. 2000 Grasping visual illusions: No  
42 evidence for a dissociation between perception and action. *Psychological Science* **11**, 20–  
43 25.
- 44 11. Pavani F, Boscagli I, Benvenuti F, Rabuffetti M, Farnè A. 1999 Are perception and action  
45 affected differently by the Titchener circles illusion? *Experimental Brain Research* **127**(1),  
46 95-101.
- 47 12. Foster RM, Franz VH. 2014 Superadditivity of the Ebbinghaus and Muller-Lyer illusions  
48 depends on the method of comparison used. *Perception* **42**, 787-796.
- 49 13. Desanghere L, Marotta JJ. 2011 “Graspability” of objects affects gaze patterns during  
50 perception and action tasks. *Experimental brain research* **212**(2), 177-187.
- 51 14. Desanghere L, Marotta JJ 2015 The influence of object shape and center of mass on  
52 grasp and gaze. *Frontiers in psychology* **6**, 1537.
- 53  
54  
55  
56  
57  
58  
59  
60

15. van Doorn H, van der Kamp J, de Wit M, Savelsbergh GJP. 2009 Another look at the Müller-Lyer illusion: Different gaze patterns in vision for action and perception, *Neuropsychologia* **47**, 804-812.
16. Voudouris D, Smeets JBJ, Fiehler K, Brenner E. 2018 Gaze when reaching to grasp a glass. *Journal of Vision* **18**(8), 16.
17. Kopiske KK, Bruno N, Hesse C, Schenk T, Franz VH. 2016 The functional subdivision of the visual brain: Is there a real illusion effect on action? A multi-lab replication study. *Cortex* **79**, 130–152.
18. Dewar MT, Carey DP. 2006 Visuomotor ‘immunity’ to perceptual illusion: A mismatch of attentional demands cannot explain the perception–action dissociation. *Neuropsychologia* **44**(8), 1501-1508.
19. Ozana A, Ganel T. 2020 A double dissociation between action and perception in bimanual grasping: evidence from the Ponzo and the Wundt–Jastrow illusions. *Scientific Reports* **10**, 14665.
20. Foster RM, Kleinholdermann U, Leifheit S, Franz VH. 2012 Does bimanual grasping of the Müller–Lyer illusion provide evidence for a functional segregation of dorsal and ventral streams? *Neuropsychologia* **50**(14), 3392–3402.
21. Whitwell RL, Streimer CL, Cant JS, Enns JT. 2021 The Ties that Bind: Agnosia, Neglect and Selective Attention to Visual Scale. *Current Neurology and Neuroscience Reports* **21**(54), 1-23.
22. Foley RT, Whitwell RL, Goodale MA. 2015 The two visual-systems hypothesis and the perspectival features of visual experience. *Consciousness and Cognition* **35**, 225-233.
23. Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada, & Social Sciences and Humanities Research Council of Canada, 2018
24. Cesanek E, Domini F. 2017 Error correction and spatial generalization in human grasp control. *Neuropsychologia* **106**, 112-122.
25. Cesanek E, Campagnoli C, Taylor JA, Domini F. 2018 Does visuomotor adaptation contribute to illusion-resistant grasping? *Psychonomic Bulletin Review* **25**, 827–845.
26. Kopiske KK, Cesanek E, Campagnoli C, Domini F. 2017 Adaptation effects in grasping the Müller-Lyer illusion. *Vision Research* **136**, 21-31.
27. Whitwell RL, Buckingham G, Enns JT, Chouinard PA, Goodale MA. 2016 Rapid decrement in the effects of the Ponzo display dissociates action and perception. *Psychonomic Bulletin Review* **23**, 1157–1163.
28. Whitwell RL, Goodale MA, Merritt KE, Enns JT. 2018 The Sander parallelogram illusion dissociates action and perception despite control for the litany of past confounds. *Cortex* **98**, 163-176.
29. Nuthmann A. 2017 Fixation durations in scene viewing: Modeling the effects of local image features, oculomotor parameters, and task. *Psychonomic Bulletin & Review* **24**(2), 370-392.
30. Holm S. 1979 A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics* **6**, 65-70.
31. Sidak Z. 1967 Rectangular confidence region for the means of multivariate normal distributions. *Journal of the American Statistical Association* **62**, 626-633.
32. Abdi H. 2010 Holm’s sequential Bonferroni procedure. *Encyclopedia of research design* **1**(8), 1-8.
33. Kirk RE. 1995 *Experimental Design: Procedures for the Behavioral Sciences*. 3rd Edn. Pacific Grove: Brooks Cole.

- 1
- 2
- 3
- 4 34. Franz VH, Fahle M, Bühlhoff HH, Gegenfurtner KR. 2001 Effects of visual illusions on grasping. *Journal of Experimental Psychology: Human Perception and Performance* **27**(5), 1124.
- 5
- 6 35. Franz VH, Scharnowski F, Gegenfurtner, KR. 2005 Illusion effects on grasping are temporally constant not dynamic. *Journal of Experimental Psychology: Human Perception and Performance* **31**, 1359–1378.
- 7
- 8
- 9 36. Fieller E. 1954 Some problems in interval estimation. *Journal of the Royal Statistical Society: Series B (Methodological)* **16**(2), 175-185.
- 10
- 11 37. Franz, VH. 2007 Ratios: A short guide to confidence limits and proper use. *arXiv preprint arXiv:0710.2024*.
- 12
- 13 38. Hesse C, Franz VH, Schenk T. 2016 Pointing and antipointing in Muller-Lyer figures: Why illusion effects need to be scaled. *Journal of Experimental Psychology: Human Perception and Performance* **42**(1), 90-102.
- 14
- 15 39. Haffenden AM, Goodale MA. 2000 Independent effects of pictorial displays on perception and action. *Vision Research* **40**, 1597-1607.
- 16
- 17 40. Haffenden AM, Schiff KC, Goodale MA. 2001 The dissociation between perception and action in the Ebbinghaus illusion: Nonillusory effects of pictorial cues on grasp. *Current Biology* **11**, 177–181.
- 18
- 19 41. Chen Z, Zhang N, Saunders JA. 2021 Dual processes underlie the effect of the Ebbinghaus illusion on control of grasping. *Journal of Experimental Psychology: Human Perception and Performance* **47**(11), 1472–1492.
- 20
- 21 42. Franz VH, Bühlhoff HH, Fahle M. 2003 Grasp effects of the Ebbinghaus illusion: obstacle avoidance is not the explanation. *Experimental Brain Research* **149**, 470–477.
- 22
- 23 43. Milner AD, Goodale MA. 2006 *The Visual Brain in Action*. New York, NY: Oxford University Press.
- 24
- 25 44. de Haan EH, Cowey A. 2011 On the usefulness of ‘what’ and ‘where’ pathways in vision. *Trends in cognitive sciences* **15**(10), 460-466.
- 26
- 27 45. Fan AW-Y, Guo LL, Frost A, Whitwell RL, Niemeier M, Cant JS. 2021 Grasping of Real-World Objects Is Not Biased by Ensemble Perception. *Frontiers in Psychology* **12**, 1158.
- 28
- 29 46. Chen J, Sperandio I, Goodale MA. 2015 Differences in the Effects of Crowding on Size Perception and Grip Scaling in Densely Cluttered 3-D Scenes. *Psychological Science* **26**(1), 58–69.
- 30
- 31 47. Goodale MA, Ganel T. 2015 Different modes of visual organization for perception and for action. *Oxford handbook of perceptual organization* **3**(1), 1-19.
- 32
- 33
- 34
- 35
- 36
- 37
- 38
- 39
- 40
- 41
- 42
- 43
- 44
- 45
- 46
- 47
- 48
- 49
- 50
- 51
- 52
- 53
- 54
- 55
- 56
- 57
- 58
- 59
- 60



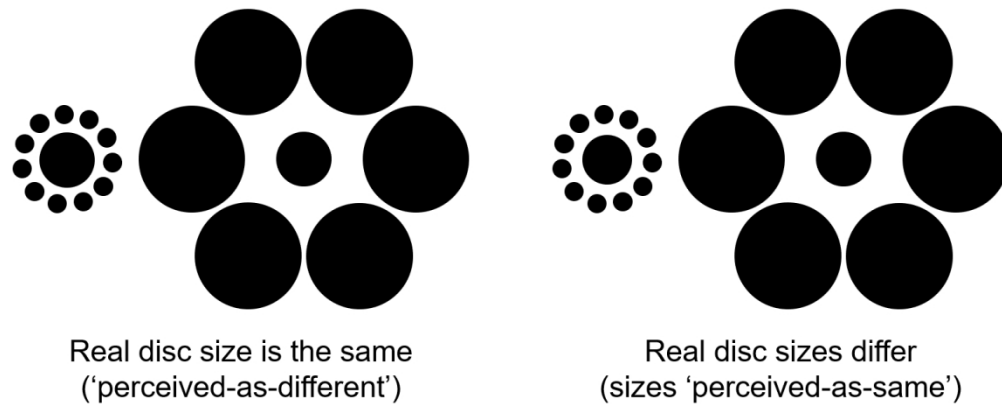


Figure 1. The Ebbinghaus display devised for the current study. Inner black-filled circles are surrounded by an annulus of small or large black-filled circles. Two versions of the display are illustrated, one in the left panel and one in the right, and each display is presented as a ‘dual’ configuration. Left panel: the traditional presentation in which the real sizes of the inner discs are the same but are misperceived as different. This is the critical feature of the display on ‘perceived-as-different’ trials. Right panel: a version in which the real sizes of the inner discs are different, leading to the impression that they are the same size. The difference in size required to achieve this varies from person to person, but it is scaled here to the most common difference in the study (2mm). A pair of discs whose real sizes were different to make them appear to be the same in size is the critical component of the ‘perceived-as-same’ display.

415x165mm (96 x 96 DPI)

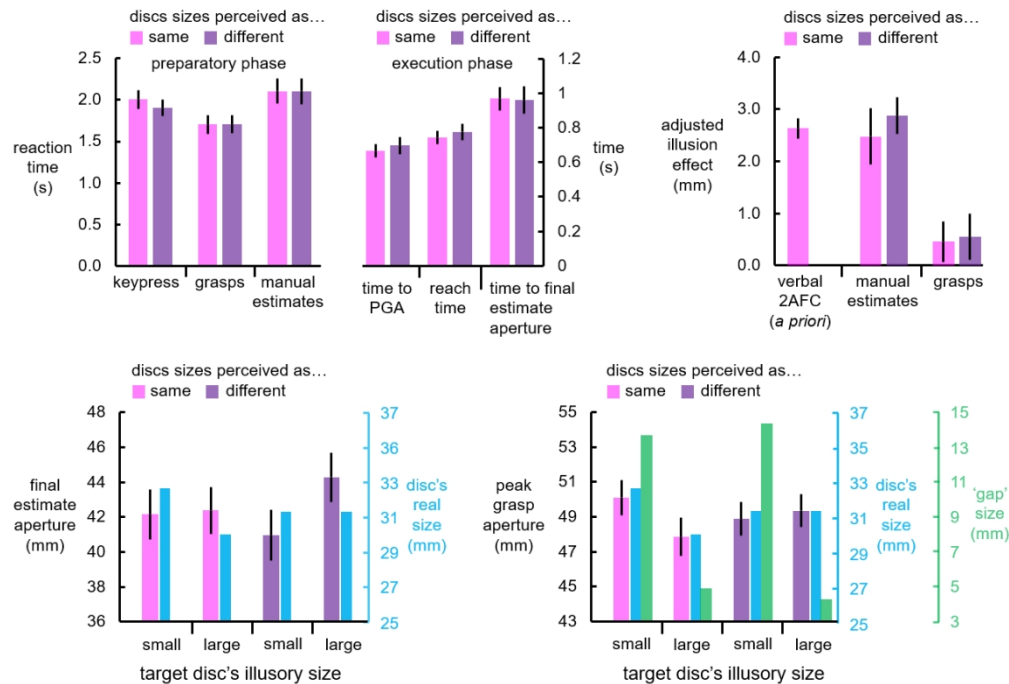


Figure 2. Top left: Preparatory and execution phase temporal dependent measures as a function of task and display type. The grasps were initiated the fastest, and the mean initiation time for the estimates and the mean keypress time for the keypress task were statistically equivalent. With respect to the execution phase, the reach time for the grasps was faster than the final estimate aperture time. Top right: the a priori effect of the Ebbinghaus illusion, using the 'perceived-as-same' display (established in the initial verbal 2AFC task), and the adjusted effects of the Ebbinghaus illusion on the manual estimates (middle) and grasps (right) as a function of display type. The effect of the illusion on the estimates was comparable to the a priori effect, whereas the effect of the display on the grasps was null. Bottom panels: Peak grasp aperture for the grasps (left) and final estimate aperture for the manual estimates (right) as functions of the direction of the effect of the illusion ("small" – the inducer decreases the perceived size of the target; "large" – the inducer increases the perceived size of the disc) and the display type (i.e., the perceived-as-same and perceived-as-different arrangements). The blue bars depict real disc size for the 'perceived-as-same' display, whereas they depict average disc size for the 'perceived-as-different' display. The average is used for real disc size (and for the illusion effect) because the two are independent of one another for these trials. Note that the blue bars possess their own ordinate axis to the right of each panel. The green bars in the bottom right panel reflect the mean distance between the outside edge of the target disc and the inner-most edges of tee surrounding inducer. The green bars possess their own ordinate axis to the right of this panel. Importantly, while all ordinate axes possess different minimum and maximum values, they share the same units and span the same distance (12 mm). The figure indicates that the illusion drives the estimates aperture but that real-disc size drives the peak grasp aperture. Error bars are SEMs.

338x229mm (96 x 96 DPI)

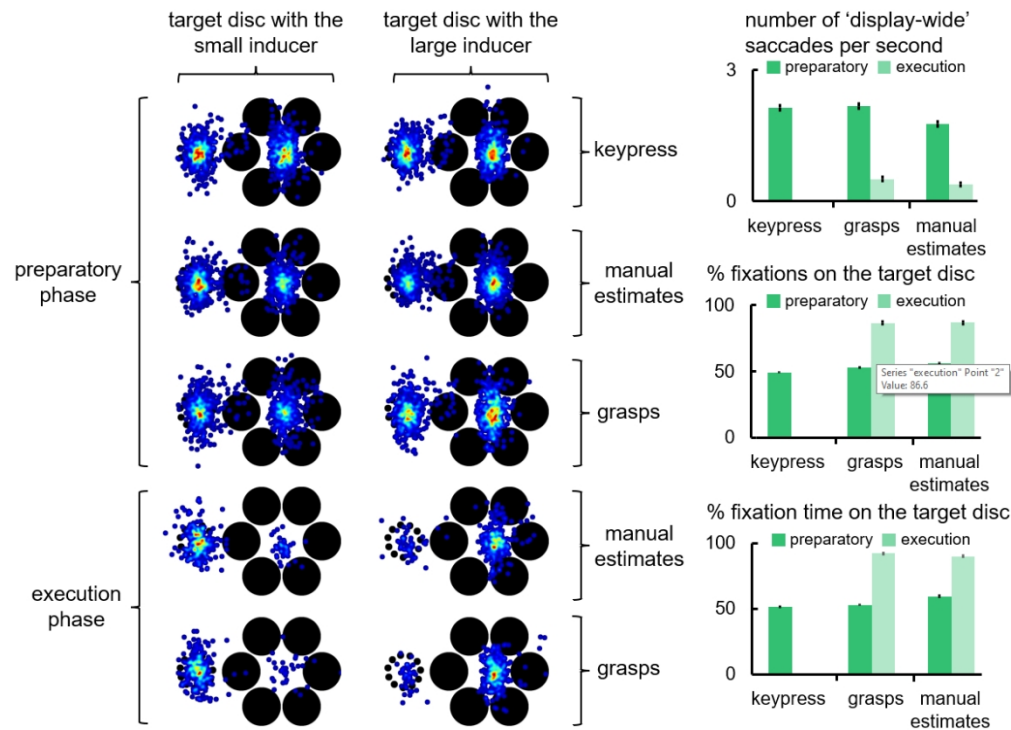


Figure 3. The spatial distribution of fixations and the fixation measures as functions of task and the response phase (i.e., the preparatory and execution phases). Left: gaze maps show the spatial distribution of fixations parsed into trials in which the target disc was surrounded by the small or large inducer. Display orientation is collapsed such that fixations on those trials with the opposite board orientation to the one depicted were flipped across the central vertical axis of the full set of fixations. Top right: The number of 'display-wide' saccades (gaze shifts between disc positions two or more locations apart from one another) per second. Middle right: The percentage of total fixations that were directed at the target disc. Bottom right: The percentage of total fixation time that was directed at the target disc and inducer (right). During the preparatory phase, participants made more saccades between the discs, fixated on the target disc less, and spent less fixation time on the target, relative to the execution phase. Moreover, none of the fixation measures during the preparatory phase differentiate the keypress and manual estimation tasks from the grasping task. During the execution phase, the fixation measures fail to differentiate the grasps and manual estimates. Error bars are SEMs.

311x224mm (96 x 96 DPI)