



OPEN ACCESS

EDITED BY
Jonathan W. Kelly,
Iowa State University, United States

REVIEWED BY
Gonzalo Garcia-Valle,
Leitat, Spain
David Harris,
University of Exeter, United Kingdom

*CORRESPONDENCE
Julia M. Juliano,
juliaang@usc.edu

SPECIALTY SECTION
This article was submitted to Virtual Reality and Human Behaviour, a section of the journal Frontiers in Virtual Reality

RECEIVED 19 April 2022
ACCEPTED 06 July 2022
PUBLISHED 08 August 2022

CITATION
Juliano JM, Phanord CS and Liew S-L (2022), Visual processing of actions directed towards three-dimensional objects in immersive virtual reality may involve holistic processing of object shape.
Front. Virtual Real. 3:923943.
doi: 10.3389/frvir.2022.923943

COPYRIGHT
© 2022 Juliano, Phanord and Liew. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Visual processing of actions directed towards three-dimensional objects in immersive virtual reality may involve holistic processing of object shape

Julia M. Juliano^{1,2*}, Coralie S. Phanord² and Sook-Lei Liew^{1,2,3,4}

¹Neuroscience Graduate Program, University of Southern California, Los Angeles, CA, United States, ²Neural Plasticity and Neurorehabilitation Laboratory, Chan Division of Occupational Science and Occupational Therapy, University of Southern California, Los Angeles, CA, United States, ³Division of Biokinesiology and Physical Therapy, University of Southern California, Los Angeles, CA, United States, ⁴Department of Neurology, USC Stevens Neuroimaging and Informatics Institute, Keck School of Medicine, University of Southern California, Los Angeles, CA, United States

Immersive virtual reality using a head-mounted display (HMD-VR) is increasing in use for motor learning and motor skill training. However, it remains unclear how visual information for action is processed in an HMD-VR environment. In the real world, actions towards three-dimensional (3D) objects are processed analytically and are immune to perceptual effects, such as processing object dimensions irrelevant to performing the action (i.e., holistic processing). However, actions towards two-dimensional (2D) objects are processed holistically and are susceptible to perceptual effects. In HMD-VR, distances are often underestimated, and the environment can appear flatter compared to the real world. Thus, actions towards virtual 3D objects in HMD-VR may be processed more like 2D objects and involve holistic processing, which is susceptible to perceptual effects. In an initial study, we used a Garner interference task to examine whether vision-for-action in HMD-VR is processed holistically and hypothesized that vision-for-action towards virtual 3D objects in HMD-VR would result in a Garner interference effect, suggesting holistic processing. We found Garner interference effects for reaction times to reach maximum grip aperture and to complete movement. These results show that visual processing of actions towards virtual 3D objects in HMD-VR may involve holistic processing of object shape. These findings demonstrate that visual information for action in HMD-VR is processed differently compared to real 3D objects and is susceptible to perceptual effects, which could affect motor skill training in HMD-VR.

KEYWORDS

immersive virtual reality, head-mounted display, virtual 3D grasping, vision-for-action, garner interference

1 Introduction

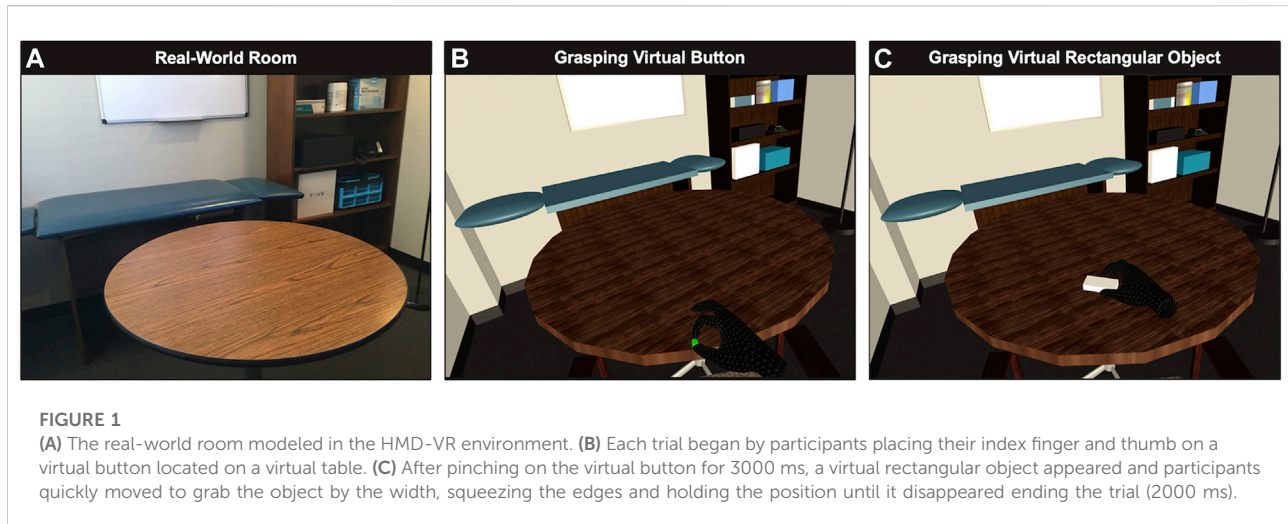
Immersive virtual reality using a head-mounted display (HMD-VR) is increasingly being used for motor learning purposes. However, it remains unclear whether visual processing for action in an HMD-VR environment is similar to processing that occurs in the real world. In the real world, visual processing for action (vision-for-action) and visual processing for perception (vision-for-perception) are thought to rely on two distinct but interacting cortical processing routes (Goodale and Milner, 1992; Goodale, 2011). Evidence from both lesion and imaging studies have demonstrated that vision-for-perception relies heavily on the ventral stream, projecting from the primary visual area (V1) to the inferior temporal cortex, while vision-for-action relies heavily on the dorsal stream, projecting from V1 to the posterior parietal cortex (Milner et al., 1991; Creem and Proffitt, 2001; Goodale et al., 2005; Goodale, 2014). The computations for vision-for-perception are also thought to be different compared to vision-for-action. Vision-for-perception is thought to rely on holistic processing, meaning that individual features and their spatial relations are perceived as a combined whole. However, when interacting with objects, the visual system only considers the individual features relevant to controlling the action and is not influenced by other features (Goodale, 2014). That is, rather than relying on holistic processing, vision-for-action is suggested to rely on analytical processing, where only the relevant features are considered without being influenced by other irrelevant information (Ganel and Goodale, 2003; Goodale, 2014).

Accumulating evidence for the perception–action distinction has been shown through behavioral experiments involving real three-dimensional (3D) objects (Goodale, 2011; Goodale, 2014). One such piece of experimental evidence can be observed through a psychophysical principle known as Weber's law. Weber's law states that the minimal difference between two objects that alters perceptual experience, called the just noticeable difference (JND), depends on the magnitude of the object (e.g., the larger the object, the larger the JND). Vision-for-perception for real 3D objects are shown to adhere to Weber's law and the JND increases linearly with object size (Ganel et al., 2008; Heath et al., 2017; Heath and Manzone, 2017). However, vision-for-action for real 3D objects does not adhere to Weber's law and the JND is unaffected by changes in object size (Ganel et al., 2008, Ganel et al., 2014; Ganel, 2015; Heath et al., 2017; Ayala et al., 2018). Evading the influence of Weber's law suggests that vision for action may be processed in an analytical fashion (Ganel, 2015; Ozana et al., 2018). Other experimental evidence can be observed through a selective attention paradigm known as the Garner interference, in which irrelevant information of an object interferes with the processing of relevant information. For example, if asked to either judge or grasp an object only by the width, relevant information would be the width and irrelevant information would be the length, and a Garner

interface occurs if the length interference with the processing of the width. A Garner interference effect is found when making speeded judgements (vision-for-perception) of real 3D objects, but not when performing speeded grasps (vision-for-action) towards real 3D objects (Ganel and Goodale, 2003; Ganel and Goodale, 2014). Garner interference and a lack thereof suggests holistic and analytical processing of a single dimension relative to the object's other dimensions, respectively. Together, this work demonstrates that analytical processing of vision-for-action for real 3D objects is not susceptible to perceptual effects.

In a technologically advancing world, it is common to interact with two-dimensional (2D) objects such as smartphones and tablets. As opposed to real 3D objects, accumulating evidence suggests that vision-for-action directed at 2D objects is susceptible to perceptual effects (Ganel et al., 2020). Unlike real 3D objects, action towards 2D objects has been shown to adhere to Weber's law, indicating a susceptibility to perceptual effects (Holmes and Heath, 2013; Ozana and Ganel, 2018; Ozana and Ganel, 2019a; Ozana and Ganel, 2019b; Ozana et al., 2020b). Action towards 2D objects has also been shown to produce a Garner interference effect, indicating holistic processing (Freud and Ganel, 2015; Ganel et al., 2020). Holistic processing suggests that irrelevant information may also be processed during visuomotor control. Together, this work shows that vision-for-action when movements are directed at 2D objects involves holistic processing and are susceptible to perceptual effects.

A particular technology increasing in use for motor learning purposes (e.g., motor rehabilitation and surgical training) is virtual reality using a head-mounted display (HMD-VR) (Huber et al., 2017; Levin, 2020). A driving factor for using HMD-VR for motor learning purposes includes the ability to replicate the real world, allowing for users to interact with virtual 3D objects in a fully adaptable and controlled environment. However, it is not clear whether vision-for-action directed towards virtual 3D objects is similar to real 3D objects. Several pieces of evidence suggest that vision-for-action in HMD-VR may shift users away from dorsal stream processing to more ventral stream processing (Harris et al., 2019). A first is that perceived space in HMD-VR can sometimes appear flat. While this underperception of distance is reported as being less in newer displays compared to older displays, this issue suggests that the HMD-VR environment may be processed more as 2D compared to the real world (Kelly et al., 2017). A second is recent findings that grasping virtual 3D objects adheres to Weber's law, and is therefore susceptible to perceptual effects (Ozana et al., 2018; Ozana et al., 2020a). Taken together, these findings suggest that vision-for-action for grasping virtual 3D objects in HMD-VR may be processed more like 2D objects and importantly, that actions in HMD-VR may involve holistic processing and be susceptible to perceptual effects. This has implications for motor skill training in HMD-VR as differences in underlying motor learning mechanisms between HMD-VR and the real world has been shown to result in a lack of contextual transfer (Harris et al., 2019; Juliano et al., 2021).



In this study, we examined whether grasping virtual 3D objects in HMD-VR produces a Garner interference effect. Participants completed a Garner interference task similar to the one used by [Ganel and Goodale \(2003\)](#). In this task, participants reach to 3D rectangular objects of varying widths and lengths, grasping the virtual object by the width. The task includes two different conditions: a Baseline condition where the width (relevant dimension) varies while the length (irrelevant dimension) remains constant, and a Filtering condition where both the width and length vary. Greater reaction times and response variability in the Filtering condition compared to the Baseline condition indicate a Garner interference effect ([Ganel and Goodale, 2003](#); [Ganel et al., 2008](#); [Ganel and Goodale, 2014](#)). We hypothesized that vision-for-action towards virtual 3D objects in HMD-VR would result in a Garner interference effect, suggesting holistic processing.

2 Materials and methods

2.1 Participants

Eighteen participants were recruited for the study (10 female/8 male; aged: $M = 26.7$, $SD = 4.8$). Eligibility criteria included right-handed individuals with no neurological impairments and normal or corrected-to-normal vision. Data was collected in-person during the COVID-19 pandemic and all participants wore surgical masks for the duration of the experiment. Written informed consent was electronically obtained from all participants to minimize in-person exposure time at the lab. The experimental protocol was approved by the USC Health Sciences Campus Institutional Review Board and performed in accordance with the 1964 Declaration of Helsinki.

2.2 Experimental apparatus

The task was programmed using the game engine development tool, Unity 3D (version 2019.4.11f1), and participants wore an Oculus Quest headset ($1,832 \times 1,920$ pixels per eye, 60 Hz) while completing the task. When wearing the headset, participants were immersed in an environment modeled after the real-world room they were located in and sat in a chair in front of a virtual table ([Figure 1A](#)). Participants also observed their right hand which was tracked by the built-in cameras located on the outside of the headset. The positions of the index finger, thumb, and wrist were recorded at each frame for the duration of the session. A 60 Hz sampling rate was used where the index finger and thumb positions were measured at the fingertips and the wrist position was measured at the wrist center. The Euclidean distance between the index finger and thumb was used to measure grip aperture and were transformed from Unity 3D coordinates to millimeters (mm).

2.3 Experimental design

The participant's virtual hand could interact with two virtual objects: a button ([Figure 1B](#)) and four rectangular objects ([Figure 1C](#)). The dimensions of the button were 10 mm in diameter and 12 mm in height. The dimensions of the four virtual rectangular objects were created from a factorial combination of two different widths (narrow: 30 mm, wide: 35.7 mm) with two different lengths (narrow: 63 mm, wide: 75 mm) ([Figure 2A](#)) (as in [Ganel and Goodale, 2003](#); [Freud and Ganel, 2015](#)). The heights of the rectangular objects were 15 mm. Participants did not experience haptic feedback when coming in contact with the virtual objects; however, they were instructed to treat all objects as if they were real. Moreover, participants were able to influence both virtual objects when

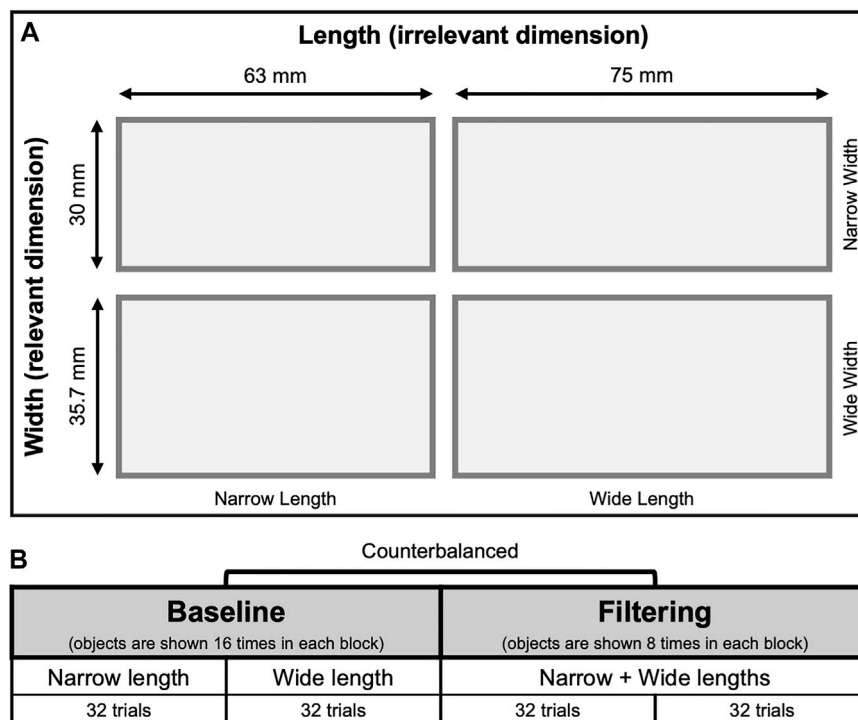


FIGURE 2
(A) The four virtual rectangular object dimensions were created from a factorial combination of two different widths (Narrow Width: 30 mm, Wide Width: 35.7 mm) and two different lengths (Narrow Length: 63 mm, Wide Length: 75 mm). Participants grasped objects along the width; therefore, the width is considered the relevant dimension while the length is considered the irrelevant dimension. **(B)** Participants completed baseline and filtering condition blocks, counterbalanced across participants. In the baseline condition, only the width varied across trials. In the filtering condition, both width and length varied across trials.

coming into contact with them through colliders within the Unity3D. Specifically, pinching the button would change the color to green and trigger a timer, but if either the index finger or thumb left the button before a rectangular object appeared (after 3,000 ms of pinching), the button would change to red, and the timer would restart until the button was pinched again. Similarly, grasping the rectangular object would trigger a timer, but if either the index finger or thumb left the object before it disappeared (after 2000 ms of grasping), the timer would restart until the object was grasped again. This was done through Unity3D physics engine which uses colliders to determine when the index finger and thumb interacted with either the button or rectangular objects and gave participants ability to influence the virtual objects.

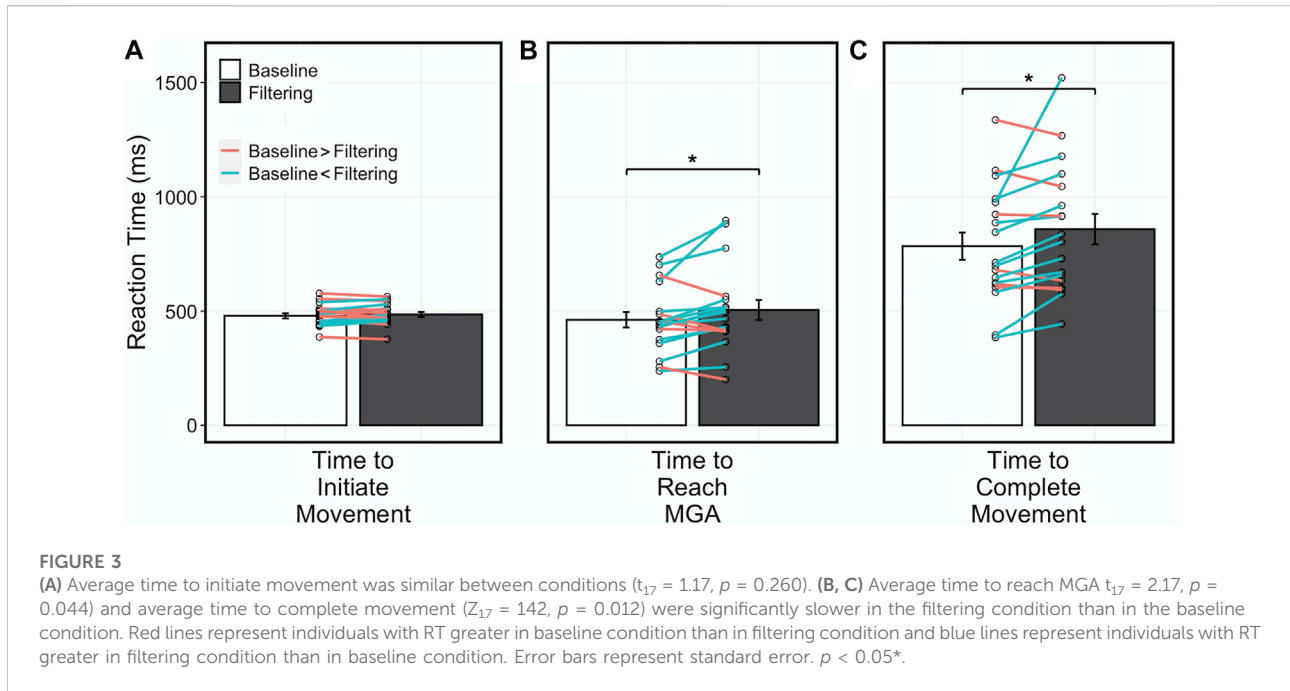
Similar to Freud and Ganel (2015), participants completed two baseline condition blocks and two filtering condition blocks, counterbalanced across participants (Figure 2B). Each block started with four practice trials which were excluded from the analysis. Then, participants completed 32 trials where, in each trial, participants reached for a virtual rectangular object. In the baseline blocks, rectangular objects randomly varied between trials only in the width dimension relevant for grasping, and participants completed one block with the two narrow length objects and one block with the two wide length objects. In the filtering blocks,

rectangular objects randomly varied between trials in both the relevant width dimension and the irrelevant length dimension and were divided equally between two blocks.

At the start of each trial, participants placed their index finger and thumb on the virtual button located on the edge of the virtual table (Figure 1B). Participants were instructed to pinch the walls of the button as if they were pinching a real button. After pinching the button for 3,000 ms, a virtual rectangular object appeared on the table approximately 30 cm away from the participant’s head. Participants were instructed to move as quickly as possible to grab the rectangular object by the width (Figure 1C). When grasping the rectangular object, participants were instructed to pretend that they were about to pick up the object by squeezing the edges, holding the position until it disappeared and the trial ended (2000 ms).

2.4 Movement analysis

All kinematic data was recorded by Unity 3D. The reaction time (RT) to initiate movement was measured as the time between when the rectangular object appeared and the time when both index finger and thumb left the button. The RT to complete movement was measured as the time between when



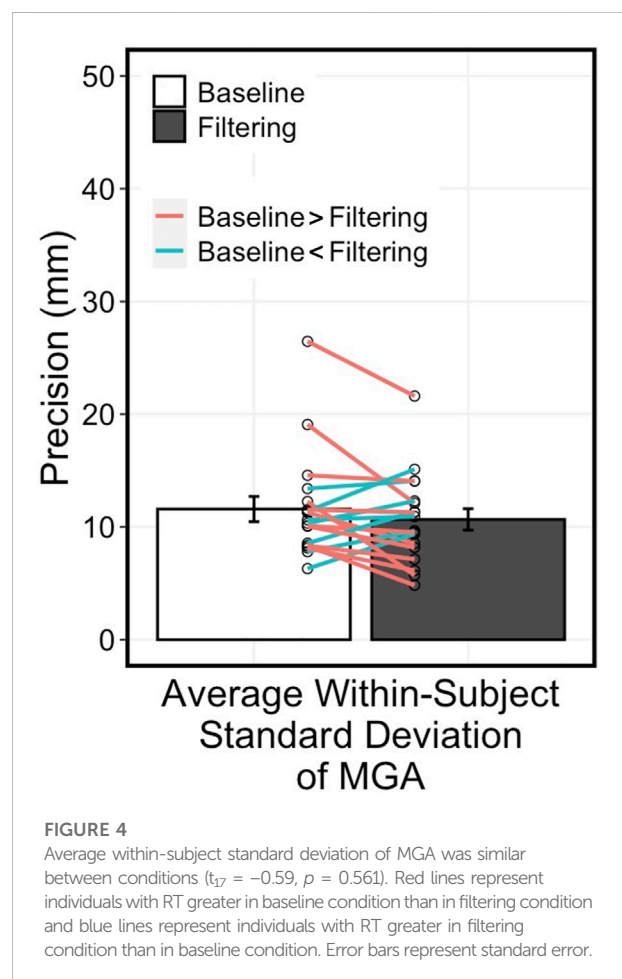
both index finger and thumb left the button and the time when both index finger and thumb collided with the rectangular object. RT to reach maximum grip aperture (MGA) was measured between when both index finger and thumb left the button and the time as the time when the distance between index finger and thumb was greatest during the movement trajectory.

We also calculated the response precision, measured by the within-subject standard deviation of MGA in baseline and filtering blocks (Ganel and Goodale, 2014; Freud and Ganel, 2015). To do this, the standard deviations for each of the four rectangular objects were computed and then averaged across participants for baseline and filtering blocks.

To validate that the spatial resolution recorded from the Oculus Quest was sufficient to accurately characterize subtle differences in movement trajectories in baseline and filtering conditions, we assessed whether MGAs were sensitive to width size in both conditions. We also calculated the grip aperture for both narrow and wide width rectangular objects across the movement trajectories in both conditions. Movements were segmented into 10 normalized time points from movement initiation (0%) to movement completion (100%), binned in 10% increments. Then, movement trajectories were averaged for narrow and wide width rectangular objects and collapsed across conditions (Freud and Ganel, 2015).

2.5 Statistical analysis

Statistical analyses were conducted using R (version 3.6.3). Participant RTs and MGAs greater than three times the standard



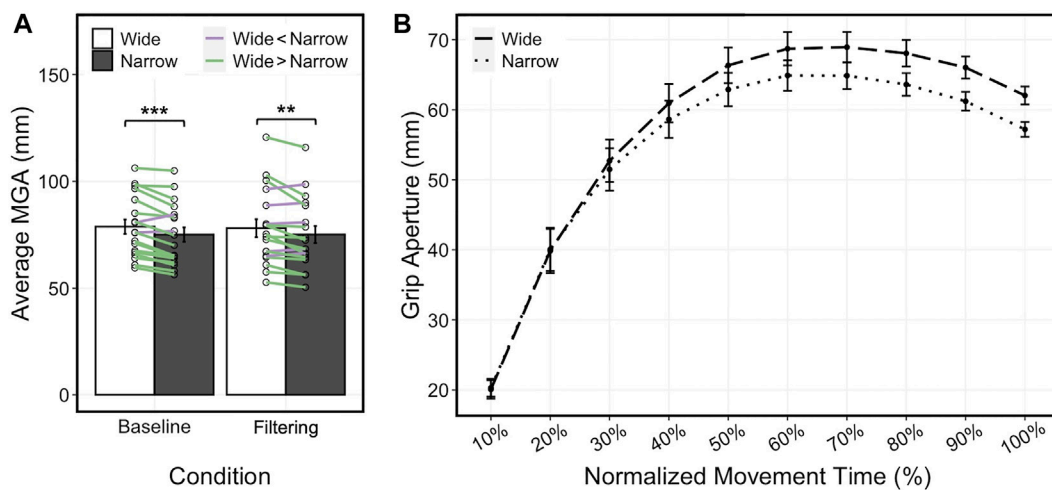


FIGURE 5

(A) Sensitivity to object width size was found for both baseline ($t_{17} = 4.65, p < 0.001$) and filtering ($t_{17} = 3.01, p = 0.008$) conditions, with larger MGAs for the wide width compared to the narrow width. Purple lines represent individuals with average MGA greater with narrow width compared to wide width and green lines represent individuals with average MGA greater with wide width compared to narrow width. (B) Average grip aperture across the movement trajectory. After normalizing movements into 10 time points, movement trajectories were averaged for narrow and wide width rectangular objects and collapsed across conditions. Maximum grip aperture reached a peak at around 70% of the movement time, resembling movement trajectories toward real 3D objects. Error bars represent standard error. $p < 0.01^{**}, p < 0.001^{***}$.

deviation of the mean were removed from the final analyses (2% of trials). To compare RTs and within-subject standard deviation of MGA between the baseline and filtering conditions, we first checked for normality of the differences between pairs using a Shapiro–Wilk test. A paired t-test was performed when data was normally distributed and a Wilcoxon Signed-Ranked test was performed when data was not normally distributed. All measures are reported as mean \pm standard deviations and significance was considered at $p < 0.05$.

3 Results

RTs to initiate movement, to reach MGA, and to complete movement were averaged for each participant for both baseline and filtering conditions (Figure 3). Average RT to reach MGA in the filtering condition was significantly slower than in the baseline condition (Figure 3B; $t_{17} = 2.17, p = 0.044, d = 0.26$; Baseline: 462 ± 143.4 ms; Filtering: 505 ± 185.9 ms). Similarly, average RT to complete movement in the filtering condition was significantly slower than in the baseline condition (Figure 3C; $Z_{17} = 142, p = 0.012, d = 0.28$; Baseline: 784 ± 254.8 ms; Filtering: 858 ± 282.2 ms). These results show that grasping virtual 3D objects in HMD-VR is susceptible to a Garner interference effect, and thus is subserved by a holistic representation. In contrast, average RTs to initiate movement were similar between the filtering and baseline conditions (Figure 3A; $t_{17} = 1.17, p = 0.260, d = 0.13$; Baseline: 479 ± 47.7 ms; Filtering: 485 ± 44.7 ms).

This result show that initiating movement to virtual 3D objects in HMD-VR may not be susceptible to a Garner interference effect.

Response precision was similar between the filtering condition and the baseline condition (Figure 4; $t_{17} = -0.59, p = 0.561, d = 0.21$; Baseline: 11.0 ± 4.4 mm; Filtering: 10.6 ± 4.0 mm). This result shows no evidence for variability-based Garner interference effects for grasping virtual 3D objects in HMD-VR.

A two-way repeated measures ANOVA was used to examine the main effects and interactions of width size (Narrow, Wide) and condition (Baseline, Filtering) on average MGA. Sensitivity to object width size was similar for baseline and filtering conditions (Figure 5A). There was a main effect of width size ($F_{1,17} = 28.63, p < 0.0001, \eta^2 = 0.63$) but no main effect of condition ($p = 0.851$) or interaction between width size and condition ($p = 0.585$). Post hoc analysis showed MGAs to be greater for the wide width compared to the narrow width for baseline ($t_{17} = 4.65, p < 0.001, d = 0.26$) and filtering ($t_{17} = 3.01, p = 0.008, d = 0.17$) conditions. These results are similar to previous studies with similar analyses of movement towards real 3D objects (e.g., Ganel et al., 2012; Freud et al., 2015) and show that the resolution of the Oculus Quest tracking system is sufficient to accurately characterize subtle differences in reaching and grasping during baseline and filtering conditions.

Movement towards 3D virtual objects also appeared to follow similar behavior to movement toward 3D real-world objects (as in Freud and Ganel, 2015). That is, grip aperture for narrow width objects was smaller than for wide width objects

(Figure 5B). A three-way repeated measures ANOVA was used to examine the main effects and interactions of width size (Narrow, Wide), condition (Baseline, Filtering), and normalized movement time on grip aperture. There were main effects of width size ($F_{1,17} = 41.49, p < 0.0001, \eta^2 = 0.71$) and movement time ($F_{9,153} = 84.53, p < 0.0001, \eta^2 = 0.83$), but not of condition ($p = 0.716$). There were significant two-way interactions between width size and movement time ($F_{9,153} = 25.09, p < 0.0001, \eta^2 = 0.60$), condition and movement time ($F_{9,153} = 2.37, p = 0.015, \eta^2 = 0.12$), but not between width size and condition ($p = 0.752$). There was also a significant three-way interaction between width size, condition, and movement time ($F_{9,153} = 3.95, p = 0.0002, \eta^2 = 0.19$). These results further show that recorded grip aperture sufficiently characterized the subtle differences in movement trajectories during baseline and filtering conditions between narrow and wide widths. Moreover, MGA was reached around 70% of the movement, similar to movement trajectories towards real 3D objects (e.g., Ganel et al., 2012). Taken together, while there were no direct comparisons made to grasping 3D objects in the real world, these results resemble that of movements previously reported in similarly designed studies (e.g., Freud and Ganel 2015).

4 Discussion

The purpose of this study was to examine whether vision-for-action directed toward virtual 3D objects in HMD-VR is processed holistically. We found that grasping virtual 3D objects in HMD-VR produced a Garner interference effect for reaction times to reach maximum grip aperture and to complete movement. This demonstrates that vision-for-action in HMD-VR may involve holistic processing during movement. However, both reaction times to initiate movement and the response precision of maximum grip aperture did not produce a Garner interference effect. The lack of significant Garner interference for movement initiation is not surprising as this also is observed for grasping directed at 2D objects and is possibly due to grasping in HMD-VR being hybrid in nature (Freud and Ganel, 2015). The lack of Garner interference for response precision could also be because grasping in HMD-VR is hybrid in nature. However, this could also be explained by a ceiling effect resulting from how finger position was measured, discussed in detail in the limitations. Overall, these findings from reaction times demonstrate that certain aspects of visual processing for action in HMD-VR is processed differently compared to real 3D objects and involve holistic processing which is susceptible to perceptual effects.

HMD-VR has been increasing in use for motor learning purposes, such as in surgical training (Mao et al., 2021), sports training (Bird, 2020), and motor rehabilitation (Levin, 2020).

Reasons for this increase include the ability for HMD-VR to replicate a real-world environment while simultaneously increasing control of the training environment. However, recent evidence suggests that motor learning in HMD-VR may differ from the real world (Levac et al., 2019). One difference is that the underlying motor learning mechanisms used in HMD-VR seem to differ from a conventional training environment. Specifically, HMD-VR has shown to recruit greater explicit, cognitive strategies during visuomotor adaptation compared to a conventional computer screen (Anglin et al., 2017). The results of the current study add further support that the underlying mechanisms driving movement in HMD-VR differ from the real world; but add that one of the reasons driving this difference is how visual information for action is processed in these types of virtual environments.

Converging evidence has also found cognitive load to be greater during motor learning tasks in HMD-VR compared to conventional training environments (Frederiksen et al., 2020; Juliano et al., 2021). However, it is unclear why cognitive load is higher in HMD-VR, compared to the real world, even when tasks and scenes are identical between the two (e.g., Anglin et al., 2017; Juliano and Liew, 2020). The results from the current study offer a potential explanation for why cognitive load is higher in HMD-VR during motor learning. Studies have found that a failure to filter out irrelevant information during movement can result in increased cognitive load (Thoma and Henson, 2011; Jost and Mayr, 2016). Holistic processing suggests that information unrelated to the task, such as irrelevant objects in the HMD-VR environment, may also be processed during visuomotor control. Therefore, if the holistic processing observed here results from a failure to filter out irrelevant information, then this could potentially explain increased cognitive load sometimes observed during motor learning in HMD-VR. Increased cognitive load in HMD-VR during motor learning has been shown to negatively affect long-term retention and context transfer (Juliano et al., 2021). Action towards 2D objects are also affected by irrelevant perceptual information and are shown to result in context-dependent learning (Ozana et al., 2018). Thus, if visuomotor control in HMD-VR is similarly affected by irrelevant perceptual information, then contextual transfer from HMD-VR could be limited, potentially explaining decreased context transfer of motor learning in HMD-VR to the real world (Levac et al., 2019; Juliano and Liew, 2020).

One potential explanation for a mechanistic shift to more ventral stream visuomotor control is the artificial presentations of distance cues used to perceive depth in an HMD-VR environment (Ganel and Goodale, 2003; Renner et al., 2013; Kelly, 2022; Kelly et al., 2022; refer to Harris et al., 2019 for a more detailed discussion). Briefly, depth perception occurs from incorporating both monocular cues (e.g., texture, shadows) and binocular cues (e.g., disparity, convergence). Monocular cues are thought to be primarily processed by the ventral stream and

binocular cues are thought to be primarily processed by the dorsal stream (Parker, 2007; Minini et al., 2010). Binocular cues are shown to be important for effective online control of hand movements in depth (Hu and Knill, 2011). While we provided depth cues in our HMD-VR environment (see Figure 1), depth perception relying on binocular cues has shown to result in greater inaccuracies and misestimations in HMD-VR compared to the real world (Jamiy and Marsh, 2019a; Jamiy and Marsh, 2019b; Ping et al., 2019; Hornsey and Hibbard, 2021). Therefore, HMD-VR may require more of a reliance on monocular cues in order to compensate for inaccurate binocular cues (Jamiy and Marsh, 2019a). A reliance on monocular cues during prehension has been shown to result in longer movement times (Mon-Williams and Dijkerman, 1999). Our results similarly show longer reaction times specifically during movement. Therefore, inaccurate binocular cues in HMD-VR could potentially explain why perceptual effects were isolated to the outcome measures related to movement.

Realism also has a role in depth perception and identifying correct distance estimations (Hibbard et al., 2017). Imaging studies have found that reaching towards and grasping real 3D objects and 2D objects rely on different neural representations. Specifically, the left anterior intraparietal sulcus, important in the control of voluntary attention and visually guided grasping, is thought to process object realness during planning and generate a forward model for visuomotor control of real 3D objects but not 2D images (Geng and Mangun, 2009; Freud et al., 2018). This finding emphasizes the role of the dorsal visual stream as being important in “vision for *real* actions upon *real* objects” (Freud et al., 2018). As opposed to 2D objects, our results did not show a Garner interference effect for precision of the response (Freud and Ganel, 2015). One potential explanation for this is that virtual 3D objects more resemble real 3D objects compared to 2D objects and therefore, allow for more precise and realistic movements than 2D objects. Future work should examine whether realism in HMD-VR is related to holistic processing.

There were multiple limitations to this study. The first was that in order to avoid any interference with objects in the real world and to generate an experience similar to that of a number of HMD-VR applications, our experimental design did not have participants come into contact with any of the objects in the virtual room. This absence of tactile feedback could also potentially explain the perceptual effects observed in this study. Future studies should examine whether this type of feedback could also avoid a Garner interference effect in HMD-VR-based grasping. Another limitation of this study was that the location where the index finger and thumb positions were measured was at the fingertips. As a result, the grip aperture appears artificially wide, which explains the large precision errors observed (i.e., >10 mm). While this

measurement has no effect on the reaction times results, it does affect the precision results and thus, the lack of Garner interference observed could partially be explained by a ceiling effect. Future studies should measure index finger and thumb positions at the finger pads to examine whether precision does result in a Garner interference effect in HMD-VR-based grasping. Another limitation of this study was the use of the built-in cameras located on the outside of the Oculus Quest headset to track hand movements. This was used as it reflected the data that generated the virtual hand observed by participants in the HMD-VR environment and previous evidence has found this type of hand tracking to be suitable for research (Holzwarth et al., 2021). As a validation for the spatial resolution, we showed that this type of tracking system was sufficient to accurately characterize subtle differences in movement trajectories in baseline and filtering conditions. However, as the sampling rate was lower compared to similarly designed studies, there may have been undetected effects that could have otherwise been observed with more superior motion tracking systems. Future studies should utilize devices with a higher sampling rate in addition to the Oculus Quest to compare and validate.

5 Conclusion

In this study, we used a Garner interference task to examine whether vision-for-action in HMD-VR is processed analytically or holistically. We show that grasping virtual 3D objects in HMD-VR is susceptible to a Garner interference effect during the movement trajectory, suggesting holistic processing during this time. This study offers an initial examination of vision-for-action in HMD-VR and shows that visual processing of action directed toward virtual 3D objects in HMD-VR may be susceptible to perceptual effects. Future studies should examine whether these perceptual effects could be avoided with the inclusion of haptic feedback as this may allow for more efficient analytical processing during movement and thus could potentially alleviate the perceptual effects observed in the current study. Future work should also examine the relationship between holistic processing and higher cognitive load as well as the effects of holistic processing on motor learning applications in HMD-VR.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <https://osf.io/qgmzx/>.

Ethics statement

The studies involving human participants were reviewed and approved by USC Health Sciences Campus Institutional Review Board. The participants provided their written informed consent to participate in this study.

Author contributions

JJ conceptualized study, designed experiment, collected data, analyzed data, and drafted manuscript. CP integrated the virtual hand with the experiment. SL advised on study conception, experimental design, and data analysis, and revised manuscript. All authors read and approved the final manuscript.

Funding

This study was funded by the National Center for Medical Rehabilitation Research (NIH K01HD091283) and National Institute of Neurological Disorders and Stroke (NIH R01NS115845).

References

- Anglin, J. M., Sugiyama, T., and Liew, S. L. (2017). Visuomotor adaptation in head-mounted virtual reality versus conventional training. *Sci. Rep.* 7, 45469. doi:10.1038/srep45469
- Ayala, N., Binsted, G., and Heath, M. (2018). Hand anthropometry and the limits of aperture separation determine the utility of Weber's law in grasping and manual estimation. *Exp. Brain Res.* 236, 2439–2446. doi:10.1007/s00221-018-5311-6
- Bird, J. M. (2020). The use of virtual reality head-mounted displays within applied sport psychology. *J. Sport Psychol. Action* 11, 115–128. doi:10.1080/21520704.2018.1563573
- Creem, S. H., and Proffitt, D. R. (2001). Defining the cortical visual systems: "What", "Where", and "How. *Acta Psychol. (Amst)*. 107, 43–68. doi:10.1016/S0001-6918(01)00021-X
- Frederiksen, J. G., Sørensen, S. M. D., Konge, L., Svendsen, M. B. S., Nobel-Jørgensen, M., Bjerrum, F., et al. (2020). Cognitive load and performance in immersive virtual reality versus conventional virtual reality simulation training of laparoscopic surgery: A randomized trial. *Surg. Endosc.* 34, 1244–1252. doi:10.1007/s00464-019-06887-8
- Freud, E., Aisenberg, D., Salzer, Y., Henik, A., and Ganel, T. (2015). Simon in action: The effect of spatial congruency on grasping trajectories. *Psychol. Res.* 79, 134–142. doi:10.1007/s00426-013-0533-5
- Freud, E., and Ganel, T. (2015). Visual control of action directed toward two-dimensional objects relies on holistic processing of object shape. *Psychon. Bull. Rev.* 22, 1377–1382. doi:10.3758/s13423-015-0803-x
- Freud, E., Macdonald, S. N., Chen, J., Quinlan, D. J., Goodale, M. A., Culham, J. C., et al. (2018). Getting a grip on reality: Grasping movements directed to real objects and images rely on dissociable neural representations. *Cortex* 98, 34–48. doi:10.1016/j.cortex.2017.02.020
- Ganel, T., Chajut, E., and Algom, D. (2008). Visual coding for action violates fundamental psychophysical principles. *Curr. Biol.* 18, R599–R601. doi:10.1016/j.cub.2008.04.052
- Ganel, T., Freud, E., Chajut, E., and Algom, D. (2012). Accurate visuomotor control below the perceptual threshold of size discrimination. *PLoS One* 7, e36253. doi:10.1371/journal.pone.0036253
- Ganel, T., Freud, E., and Meiran, N. (2014). Action is immune to the effects of Weber's law throughout the entire grasping trajectory. *J. Vis.* 14, 11. doi:10.1167/14.7.11

Acknowledgments

The authors would like to thank Nicolas Schweighofer for his suggestions and feedback on this work.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Ganel, T., and Goodale, M. A. (2014). Variability-based Garner interference for perceptual estimations but not for grasping. *Exp. Brain Res.* 232, 1751–1758. doi:10.1007/s00221-014-3867-3
- Ganel, T., and Goodale, M. A. (2003). Visual control of action but not perception requires analytical processing of object shape. *Nature* 426, 664–667. doi:10.1038/nature02156
- Ganel, T., Ozana, A., and Goodale, M. A. (2020). When perception intrudes on 2D grasping: Evidence from Garner interference. *Psychol. Res.* 84, 2138–2143. doi:10.1007/s00426-019-01216-z
- Ganel, T. (2015). Weber's law in grasping. *J. Vis.* 15, 18. doi:10.1167/15.8.18
- Geng, J. J., and Mangun, G. R. (2009). Anterior intraparietal sulcus is sensitive to bottom-up attention driven by stimulus salience. *J. Cogn. Neurosci.* 21, 1584–1601. doi:10.1162/jocn.2009.21103
- Goodale, M. A. (2014). How (and why) the visual control of action differs from visual perception. *Proc. R. Soc. B* 281, 20140337. doi:10.1098/rspb.2014.0337
- Goodale, M. A., Króliczak, G., and Westwood, D. A. (2005). Dual routes to action: Contributions of the dorsal and ventral streams to adaptive behavior. *Prog. Brain Res.* 149, 269–283. doi:10.1016/S0079-6123(05)49019-6
- Goodale, M. A., and Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in neurosciences* 15 (1), 20–25.
- Goodale, M. A. (2011). Transforming vision into action. *Vis. Res.* 51, 1567–1587. doi:10.1016/j.visres.2010.07.027
- Harris, D. J., Buckingham, G., Wilson, M. R., and Vine, S. J. (2019). Virtually the same? How impaired sensory information in virtual reality may disrupt vision for action. *Exp. Brain Res.* 237, 2761–2766. doi:10.1007/s00221-019-05642-8
- Heath, M., Manzone, J., Khan, M., and Davarpanah Jazi, S. (2017). Vision for action and perception elicit dissociable adherence to Weber's law across a range of 'graspable' target objects. *Exp. Brain Res.* 235, 3003–3012. doi:10.1007/s00221-017-5025-1
- Heath, M., and Manzone, J. (2017). Manual estimations of functionally graspable target objects adhere to Weber's law. *Exp. Brain Res.* 235, 1701–1707. doi:10.1007/s00221-017-4913-8
- Hibbard, P. B., Haines, A. E., and Hornsey, R. L. (2017). Magnitude, precision, and realism of depth perception in stereoscopic vision. *Cogn. Res.* 2, 25. doi:10.1186/s41235-017-0062-7

- Holmes, S. A., and Heath, M. (2013). Goal-directed grasping: The dimensional properties of an object influence the nature of the visual information mediating aperture shaping. *Brain Cogn.* 82, 18–24. doi:10.1016/j.bandc.2013.02.005
- Holzwarth, V., Gislser, J., Zurich, E., Christian Hirt, S., and Andreas Kunz, S. (2021). “Comparing the accuracy and precision of SteamVR tracking 2.0 and Oculus quest 2 in a room scale setup,” in ICVARS 2021 (ACM), Melbourne, VIC, Australia, 1–5. doi:10.1145/3463914.3463921
- Hornsey, R. L., and Hibbard, P. B. (2021). Contributions of pictorial and binocular cues to the perception of distance in virtual reality. *Virtual Real.* 25, 1087–1103. doi:10.1007/s10055-021-00500-x
- Hu, B., and Knill, D. C. (2011). Binocular and monocular depth cues in online feedback control of 3D pointing movement. *J. Vis.* 11, 23. doi:10.1167/11.7.23
- Huber, T., Paschold, M., Hansen, C., Wunderling, T., Lang, H., Kneist, W., et al. (2017). New dimensions in surgical training: Immersive virtual reality laparoscopic simulation exhilarates surgical staff. *Surg. Endosc.* 31, 4472–4477. doi:10.1007/s00464-017-5500-6
- Jamiy, F. E., and Marsh, R. (2019a). “Distance estimation in virtual reality and augmented reality: A survey,” Brookings, SD, USA, IEEE International Conference on Electro Information Technology 20–22 May 2019, 063–068. doi:10.1109/EIT.2019.8834182
- Jamiy, F. E., and Marsh, R. (2019b). Survey on depth perception in head mounted displays: Distance estimation in virtual reality, augmented reality, and mixed reality. *IET Image Process.* 13, 707–712. doi:10.1049/iet-ipr.2018.5920
- Jost, K., and Mayr, U. (2016). Switching between filter settings reduces the efficient utilization of visual working memory. *Cogn. Affect. Behav. Neurosci.* 16, 207–218. doi:10.3758/s13415-015-0380-5
- Juliano, J. M., and Liew, S. L. (2020). Transfer of motor skill between virtual reality viewed using a head-mounted display and conventional screen environments. *J. Neuroeng. Rehabil.* 17, 48. doi:10.1186/s12984-020-00678-2
- Juliano, J. M., Schweighofer, N., and Liew, S.-L. (2021). Increased cognitive load in immersive virtual reality during visuomotor adaptation is associated with decreased long-term retention and context transfer. *Res. Sq.* doi:10.21203/rs.3.rs-1139453/v1
- Kelly, J. W., Cherep, L. A., and Siegel, Z. D. (2017). Perceived space in the HTC vive. *ACM Trans. Appl. Percept.* 15, 1–16. doi:10.1145/3106155
- Kelly, J. W. (2022). Distance perception in virtual reality: A meta-analysis of the effect of head-mounted display characteristics. *PsyArXiv*, 1–11.
- Kelly, J. W., Doty, T. A., Ambourn, M., and Cherep, L. A. (2022). Distance perception in the Oculus quest and Oculus quest 2. *Front. Virtual Real.* 3, 1–11. doi:10.3389/frvir.2022.850471
- Levac, D. E., Huber, M. E., and Sternad, D. (2019). Learning and transfer of complex motor skills in virtual reality: A perspective review. *J. Neuroeng. Rehabil.* 16, 121. doi:10.1186/s12984-019-0587-8
- Levin, M. F. (2020). What is the potential of virtual reality for post-stroke sensorimotor rehabilitation? *Expert Rev. Neurother.* 20, 195–197. doi:10.1080/14737175.2020.1727741
- Mao, R. Q., Lan, L., Kay, J., Lohre, R., Ayeni, O. R., Goel, D. P., et al. (2021). Immersive virtual reality for surgical training: A systematic review. *J. Surg. Res.* 268, 40–58. doi:10.1016/j.jss.2021.06.045
- Milner, A. D., Perrett, D. I., Johnston, R. S., Benson, P. J., Jordan, T. R., Heeley, D. W., et al. (1991). Perception and action in “visual form agnosia. *Brain* 114, 405–428. doi:10.1093/brain/114.1.405
- Minini, L., Parker, A. J., and Bridge, H. (2010). Neural modulation by binocular disparity greatest in human dorsal visual stream. *J. Neurophysiol.* 104, 169–178. doi:10.1152/jn.00790.2009
- Mon-Williams, M., and Dijkerman, H. C. (1999). The use of vergence information in the programming of prehension. *Exp. Brain Res.* 128, 578–582. doi:10.1007/s002210050885
- Ozana, A., Berman, S., and Ganel, T. (2018). Grasping trajectories in a virtual environment adhere to Weber’s law. *Exp. Brain Res.* 236, 1775–1787. doi:10.1007/s00221-018-5265-8
- Ozana, A., Berman, S., and Ganel, T. (2020a). Grasping Weber’s law in a virtual environment: The effect of haptic feedback. *Front. Psychol.* 11, 573352. doi:10.3389/frpsyg.2020.573352
- Ozana, A., and Ganel, T. (2018). Dissociable effects of irrelevant context on 2D and 3D grasping. *Atten. Percept. Psychophys.* 80, 564–575. doi:10.3758/s13414-017-1443-1
- Ozana, A., and Ganel, T. (2019a). Obeying the law: Speed–precision tradeoffs and the adherence to Weber’s law in 2D grasping. *Exp. Brain Res.* 237, 2011–2021. doi:10.1007/s00221-019-05572-5
- Ozana, A., and Ganel, T. (2019b). Weber’s law in 2D and 3D grasping. *Psychol. Res.* 83, 977–988. doi:10.1007/s00426-017-0913-3
- Ozana, A., Namdar, G., and Ganel, T. (2020b). Active visuomotor interactions with virtual objects on touchscreens adhere to Weber’s law. *Psychol. Res.* 84, 2144–2156. doi:10.1007/s00426-019-01210-5
- Parker, A. J. (2007). Binocular depth perception and the cerebral cortex. *Nat. Rev. Neurosci.* 8, 379–391. doi:10.1038/nrn2131
- Ping, J., Liu, Y., and Weng, D. (2019). “Comparison in depth perception between virtual reality and augmented reality systems,” in 26th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2019 - Proceedings, Osaka, Japan, 23–27 March 2019, 1124–1125. doi:10.1109/VR.2019.8798174
- Renner, R. S., Velichkovsky, B. M., and Helmert, J. R. (2013). The perception of egocentric distances in virtual environments - a review. *ACM Comput. Surv.* 46, 1–40. doi:10.1145/2543581.2543590
- Thoma, V., and Henson, R. N. (2011). Object representations in ventral and dorsal visual streams: fMRI repetition effects depend on attention and part-whole configuration. *Neuroimage* 57, 513–525. doi:10.1016/j.neuroimage.2011.04.035