



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Pictorial depth cues always influence reaching distance

Citation for published version:

McIntosh, RD, Iveson, MH, Sandoval Similä, S & Buonocore, A 2023, 'Pictorial depth cues always influence reaching distance', *Neuropsychologia*, vol. 190, 108701. <https://doi.org/10.1016/j.neuropsychologia.2023.108701>

Digital Object Identifier (DOI):

[10.1016/j.neuropsychologia.2023.108701](https://doi.org/10.1016/j.neuropsychologia.2023.108701)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Publisher's PDF, also known as Version of record

Published In:

Neuropsychologia

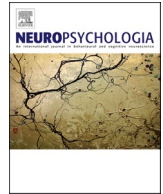
General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.





Pictorial depth cues always influence reaching distance

Robert D. McIntosh^{a,*}, Matthew H. Iveson^b, Sebastian Sandoval Similä^a, Antimo Buonocore^c

^a Human Cognitive Neuroscience, Psychology, University of Edinburgh, UK

^b Centre for Clinical Brain Sciences, University of Edinburgh, UK

^c Department of Educational, Psychological and Communication Sciences, Suor Orsola Benincasa University, Naples, 80135, Italy

ABSTRACT

We report five experiments to test the influence of pictorial depth on reaching. Our core method is to project a wide-field background of linear perspective and/or texture gradient onto a tabletop, and to measure the amplitude of reaches made to targets within it. In 63 healthy participants performing immediate open-loop reaches across Experiments 1–4, we observed a clear effect of pictorial depth. This effect was driven specifically by the convergence of the background pattern at the target position: for each additional degree of pictorial convergence, reaching distance increased by half a millimetre. In the individual experiments, we applied manipulations that might be expected to modify the influence of pictorial depth. We found no evidence that the effect was modified with monocular viewing, or when participants responded with the left hand, or if a memory delay was inserted before the response. Nor did participants become less susceptible to pictorial depth when visual feedback of terminal reaching errors was provided, although visual feedback during the reach did mitigate the influence of pictorial depth. Finally, the visual form agnostic patient DF showed an entirely normal effect of pictorial depth cues, which leads us to question the idea that this effect emanates from visual analyses of size and shape in the ventral stream, rather than from the dorsal stream, or from earlier stages of visual processing.

1. Introduction

Goodale and Milner's model of human cortical vision proposes a distinction between ventral-stream processing, which underpins visual object and pattern recognition, and dorsal stream processing, which guides target-directed actions (Goodale and Milner, 1992; Milner and Goodale, 1995). Their model is couched in terms of the behavioural outputs that each stream serves, encapsulated as the distinction between *vision-for-perception* and *vision-for-action*. Both vision-for-perception and vision-for action require spatial analyses, but encoded differently for their distinct behavioural roles. The perceptual system encodes environment-centred (allocentric) positions and relative sizes of objects (and object parts), independent of the current viewpoint, to create enduring representations. The action system encodes viewer-centred (egocentric) positions and absolute metrics, to enable accurate movements towards targets in the immediate field of view. This implies that the programming of immediate actions by the dorsal stream is preferentially based on absolute distance cues specifying egocentric position, independent of the surrounding visual context.

Relevant neuropsychological evidence comes from the study of

patients with damage to ventral stream structures, who exhibit visual form agnosia. The best-studied such patient is DF, who sustained bilateral damage affecting the Lateral Occipital Complex (LOC), a ventral stream area critical for object recognition (James et al., 2003; Milner et al., 1991). DF is adept at directing reaching and grasping actions to objects that she is otherwise unable to report the visual properties of (e.g. size, shape, orientation, distance), and this has been attributed to a relative sparing of her dorsal stream (Carey et al., 1998; Goodale et al., 1991, 1994; Milner et al., 1991).¹ DF's preserved visuomotor abilities have been found to rely heavily on extra-retinal cues to (absolute) distance (Mon-Williams et al., 2001a, 2001b; Wann et al., 2001), and binocular or dynamic monocular cues to (relative) depth (Dijkerman et al., 1996, 1999; Marotta et al., 1997). When these cues are perturbed or removed, her visuomotor performance deteriorates precipitously.

Healthy participants are much more robust to such perturbations, suggesting that they additionally make use of pictorial cues, presumably extracted within the ventral stream (Dijkerman et al., 1996; Mon-Williams et al., 2001a, 2001b; Tresilian et al., 1999; Tresilian and Mon-Williams, 2000). However, it has been debated whether pictorial depth cues are accessed only when extra-retinal and binocular cues are

* Corresponding author. Psychology, University of Edinburgh, 7 George Square, Edinburgh, EH8 9JZ, UK.

E-mail address: r.d.mcintosh@ed.ac.uk (R.D. McIntosh).

¹ DF's dorsal stream is not entirely spared. James et al. (2003) found a small lesion in left posterior parietal cortex, and Bridge and colleagues (2013) subsequently reported significant cortical thinning of the intra-parietal sulcus bilaterally. This dorsal stream involvement presumably accounts for the fact that DF has now been found to misreach severely for targets presented in peripheral vision, a pattern diagnostic of optic ataxia (Hesse et al., 2012, 2014; Rossit et al., 2018). However, her visuomotor performance is generally good for simple reaching and grasping tasks in which she can fixate the target, as in the present experiments.

unavailable, or are used routinely by action systems. For instance, a study by [Marotta and Goodale \(2001\)](#), repeatedly presented the same featureless sphere for grasping at eye-level, allowing participants to familiarise with its size. This familiarity could enable participants to infer absolute viewing distance from retinal image size, and Marotta and Goodale tested for this using occasional trials in which the usual object was swapped for a similar sphere of a different size. They concluded that familiar size influenced reaching responses only in a monocular condition, when binocular cues were excluded. However, a later study using more distinctive grasping targets (branded matchboxes) found a strong effect of familiar size even with binocular viewing ([McIntosh and Lashley, 2008](#)). So familiar size is one pictorial depth cue that may influence reaching under full cue conditions; but it could be a special case in this respect, because it offers absolute distance information likely to be especially useful for action systems.

A study by [Foley and Held \(1972\)](#) compared visually-directed (i.e. open-loop) reaching towards points of light viewed binocularly, with reaching in a multi-cue condition that also included relative size (of similar objects), linear perspective and accommodation cues. Participants' reaching distances were more accurate and precise in the multi-cue condition. This indicates that the additional depth cues were influential over and above the binocular extra-retinal cue of vergence, although it does not separate out their contributions. Nonetheless, subsequent literature has strongly emphasised binocular and extra-retinal cues as the pre-eminent sources of distance and depth information for direct manual actions towards objects ([Melmoth and Grant, 2006](#); [Servos, 2000](#); [Volcic et al., 2014](#)), although the initial programming of these actions may be somewhat approximate so that terminal accuracy is also reliant on online-control ([Bozzacchi et al., 2014](#); [Kopiske et al., 2019](#); [Melmoth and Grant, 2006](#)). Thus, although pictorial cues such as linear perspective and texture gradients have a large influence on our perceptual experience of a scene, it is often argued that the dorsal stream visuomotor system is refractory to such influences, at least under normal binocular conditions.

One popular approach to investigate this idea has been to study actions directed at objects within a salient pictorial context that causes an illusory misperception of the target. The prototypical 'illusions-in-action' study is that of [Aglioti et al. \(1995\)](#), who presented a 3D disc at the centre of a 2D ring of larger or smaller circles, and recorded participants' maximum finger-thumb grip aperture for picking it up. If the action system is concerned with metrical computations centred on the target itself ([Goodale, 2008a](#); [Goodale and Milner, 1992](#); [Milner and Goodale, 1995](#)), then grip aperture should be unaffected by the (Ebbinghaus-Titchener) size illusion that this display induces. This was the result originally claimed by Aglioti and colleagues, and it ignited a fierce debate, which has lasted more than two decades, over whether and how the effects of the Ebbinghaus-Titchener illusion on grip aperture differ from its effects on perception (e.g. [Franz, 2001, 2003](#); [Franz et al., 2000, 2003](#); [Franz and Gegenfurtner, 2008](#); [Haffenden et al., 2001](#); [Haffenden and Goodale, 1998, 2000](#); [Kopiske et al., 2016, 2017](#); [Whitwell and Goodale, 2017](#)).

The empirical and theoretical issues surrounding illusions in action studies are tangled and thorny (e.g. [Bruno, 2001](#); [Bruno et al., 2008](#); [Bruno and Franz, 2009](#); [Carey, 2001](#); [Goodale, 2008b](#); [Smeets and Brenner, 2006](#)). As a case in point, the pictorial basis of the Ebbinghaus-Titchener illusion is unclear: it might be a size contrast effect and/or arise from size-constancy scaling secondary to a pictorial illusion of depth in the display ([Doherty et al., 2010](#); [Gregory, 1963](#); [McCready, 1985](#)). The dependent measure of grip aperture is itself a complex behavioural variable, subject to more influences than just the size of the target, and it would be an indirect measure to choose if the primary illusion were one of depth. Further, when attempting to compare the effects of pictorial illusions on perception and action, many differences in task demands (over and above the perception-action distinction) could contribute to differences in measured outcome ([Franz and Gegenfurtner, 2008](#); [Mon-Williams and Bull, 2000](#); [Smeets](#)

[and Brenner, 2001](#)).

In the present study, we aim to sidestep many of the complexities of the prior illusions-in-action literature, by asking a more straightforward question: do pictorial depth cues influence reach distance? Our core method is to project a wide-field display of linear perspective and/or texture gradient—two idealised pictorial depth cues that depend on the relative sizes of elements in the scene—and to measure the terminal displacement of reaches made to targets within it. We also depart from the typical illusions-in-action paradigm by studying reaching alone, with no direct comparison to perceptual responses. However, whilst our experiments all use the same basic reaching task, Experiments 1–3 include manipulations that have been proposed to boost ventral stream influences on action (see [Goodale, 2008b](#)).

In Experiment 1 we will compare reaching in monocular and binocular viewing conditions. If monocular pictorial depth cues are available to the visuomotor system, then it would be reasonable to expect them to be more heavily weighted when binocular sources of information are removed. Indeed, if the dorsal stream visuomotor system normally relies exclusively on binocular and extra-retinal cues, then any influence of pictorial depth might arise only in monocular conditions (e.g. [Marotta and Goodale, 2001](#)). In Experiment 2, we will compare reaching with the left and right hand, because it has been proposed that the left hemisphere is specialised for visuomotor control, so that right hand actions are fluent and automatic, whilst the left hand requires more cognitive supervision, with increased ventral stream involvement. [Gonzalez et al. \(2006\)](#) based this proposal partly on their finding that, regardless of handedness, grasping responses of the left hand but not the right hand were susceptible to illusions of size induced by manipulation of pictorial depth cues (see also [Goodale, 2008b](#)). It has also been argued that the dorsal stream visuomotor system can act only towards targets in the immediate field of view, and that any delay between viewing and acting must be bridged by memory representations derived from ventral stream processing. Again this is supported by claims of increased susceptibility of grasping responses to pictorial illusions when the view of the target is removed prior to initiation of the response ([Gentilucci et al., 1996](#); [Goodale, 2008b](#); [Goodale et al., 2004](#); [Hu and Goodale, 2000](#); [Westwood and Goodale, 2003](#)). In Experiment 3, we will test the corresponding prediction that pictorial depth cues will have a greater influence in delayed than in immediate reaching.

Experiment 4 will address a different question, asking whether visual feedback on reaching errors can allow participants to learn to down-weight unreliable pictorial cues. The effect of learning will be assessed by comparing the influence of pictorial cues between a first and second block of trials. This experiment is intended to assess whether the influence of misleading pictorial depth cues is modified by experience of reaching errors that they induce. We will compare the standard open-loop reaching condition, in which no feedback is available, with a terminal feedback condition in which the final landing position can be observed at the end of the reach, and a closed-loop condition in which visual feedback is continuously available throughout. Finally, we have taken the opportunity to test the visual form agnostic patient DF (Experiment 5). If an influence of pictorial cues on action emanates from the ventral stream, then we would expect it to be reduced or eliminated in patient DF, who is known to rely heavily on binocular and extra-retinal cues ([Dijkerman et al., 1996](#); [Marotta et al., 1997](#); [Mon-Williams et al., 2001a, 2001b](#); [Wann et al., 2001](#)).

Despite a large illusions-in-action literature, our core question of whether linear perspective and texture gradients influence reach distance has not, to our knowledge, been asked before. This was surprising to us at the time that we ran these experiments (2007–2008), and it is even more so now. Our reasons for not publishing these experiments in the intervening years are noted in Methods. Our impetus for doing so now is to contribute novel data and perspectives to this special issue celebrating the career of Mel Goodale, whose creative and insightful experimental and theoretical work provides the inspiration and context for these experiments.

2. Methods

2.1. Participants

With the exception of patient DF, all participants were recruited from amongst students of the University of Edinburgh. All participants were right handed by self-report, and had a positive Laterality Quotient on the Edinburgh Handedness Inventory (Oldfield, 1971), except for left-handed participants in Experiment 2, who had a negative Laterality Quotient. All participants passed the screening plates of the TNO stereotest (Lam ris Ootech BV), confirming normal stereovision. Each participant took part in a single experiment only.

Experiment 1 had 24 right-handed participants (13 female, 11 male; age range was 22–28 years), with half assigned to a binocular and half to a monocular viewing condition.² Experiment 2 had 24 participants, half of them left-handed and half right-handed. Experiment 3 had 17 right-handed participants, in a single group, with all manipulations within-subjects. Experiment 4 had three feedback groups, with 10 participants per group. Experiment 5 was a single-case assessment of patient DF, a right-handed woman with visual form agnosia due to brain damage from carbon monoxide inhalation twenty-one years earlier. DF was 54 years old at the time of at the time of testing in 2008, having been the subject of many scientific studies since 1989 (Milner and Heywood, 1989). Her brain damage is bilaterally to the ventral stream of visual processing, particularly the ventrolateral occipital cortex, with more minor dorsal stream damage within the posterior parietal cortex. The anatomical details of her brain damage have been most fully documented by James et al. (2003) and Bridge and colleagues (2013) (see Footnote 1 in Introduction).

2.2. General set-up and procedure

All five experiments shared the same general set-up. The participant was seated centrally at a ‘projection table’, which houses a horizontal screen with an active display area of 1000 mm (wide) by 750 mm (front-to-back), 150 mm from the edge of the table. Stimuli were back-projected onto the screen at a resolution of 1024*768 via a projector and mirror arrangement within the undercarriage of the table. Fig. 1 shows what this back-projection looked like in practice; note that the laboratory lights were switched off during the experiments to increase the salience of the display. Participants rested their head on a chinrest so that eye-level was approximately 400 mm above the table surface and 50 mm proximal to the display. They wore PLATO liquid crystal display (LCD) goggles, which could switch between transparent and opaque, gating their view of the apparatus. The home position for the responding hand was a start button, elevated 40 mm above the table surface, and fixed centrally in front of the chinrest, 20 mm proximal to the display. An infra-red-emitting-diode (marker) was attached to the nail of the index finger of the responding hand, allowing reaching responses to be recorded at 200 Hz by an Optotrak Certus motion capture system (Northern Digital Inc.).

The five experiments shared a common core procedure, deviations from which are noted in the separate description for each experiment. Before each trial, the participant held down the start button with the index finger of their responding hand (usually the right); the other hand was kept below the table. Whilst the button was depressed, the LCD goggles were transparent, giving a view of the apparatus. Stimulus

projection was initiated remotely by the experimenter. On each trial, a yellow X-shaped target was presented, superimposed on a white-on-black pictorial depth background. The background was either consistent with normal depth (Baseline) or converged towards the far end to create exaggerated pictorial depth (Medium, or High), as shown in Fig. 2. The target was presented at different distances from the start button (see specific experiments), and at one of three sizes (small = 13 mm; medium = 23 mm; large = 33 mm). Target size was not a factor of interest but was manipulated to avoid an invariant size that might become a learned cue to distance. After a stimulus presentation period of 1000 ms, a high start tone sounded, cueing the participant to reach out ‘quickly and accurately’ to place their index finger on the target cross. When the finger left the start button, the LCD goggles turned opaque, so that the movement was made without visual feedback (open-loop). The participant was required to leave the finger where it landed until a low tone, 2000 ms after stimulus onset, cued them to return to the start button. Coincident with the tone, the target display was replaced by a black field. Each block of experimental trials was preceded by five unrecorded trials, sampled randomly from the stimulus set for that block, to establish the rhythm of the task.

2.3. Experiment-specific procedures

Experiment 1 tested the influence of viewing condition, using a between-subjects manipulation, with 12 participants in a monocular condition and 12 in a binocular condition. This was controlled by clearing both panes of the LCD goggles, or only one pane, set to the participant’s dominant eye (usually the right) as determined by a preliminary Porta test. Two types of pictorial background were used (linear perspective and texture gradient: see Fig. 2). Each participant performing one block of trials for each background type, with block order counterbalanced within each viewing condition. Each block had 81 experimental trials, with three repetitions for each combination of target distance (220 mm; 320 mm; 420 mm), pictorial depth (baseline; medium; high), and target size (small; medium; large). Each participant therefore performed a total of 162 trials.

Experiment 2 tested the influence of the hand of response, using a within-subjects blocked manipulation of responding hand (left or right), and included the between-subjects factor of handedness with 12 participants per group (left-handers, right-handers). For this and all subsequent experiments, the pictorial background was always of the ‘grid’ type (bottom row of Fig. 2). Each participant performed one block of trials for each hand, with block order counterbalanced within each handedness group. Each block had 81 experimental trials, with three repetitions for each combination of target distance (220 mm; 320 mm; 420 mm), pictorial depth (baseline; medium; high), and target size (small; medium; large). Each participant therefore performed a total of 162 trials.

Experiment 3 tested the effect of inserting a memory delay prior to action. Delay was manipulated within-subjects, with each of 17 participants performing the reaching task in four delay conditions. Participants were required to respond as soon as the high tone sounded after stimulus onset, but the temporal relationship between the tone and the occlusion of vision by the LCD glasses was manipulated. The no-delay condition was the standard immediate open-loop reaching task, with the tone delivered 1000 ms after stimulus onset, and vision occluded when the start button was subsequently released, so that the display was visible until the reach began. In the RT-delay condition, vision was occluded with the onset of the high tone, 1000 ms after stimulus onset, so that the display was already occluded at the moment the reach began. In the 2s-delay condition, vision was occluded 1000 ms after stimulus onset, and the high tone was delivered 2000 ms later. The 5s-delay condition was the same except that the tone was delivered 5000 ms after vision was occluded. To limit the total number of experimental trials, given four different delay conditions, only two target distances were used (280

² Due to loss of paper records, age and gender data for Experiments 2–4 are unavailable. This does not compromise the scientific purpose of the experiments; all participants were from a relatively homogenous undergraduate and postgraduate student population, and the specific demographics are not relevant to the hypotheses tested. The missing demographic details is the main reason that these experiments have not been written up previously: we had hoped that the paper records would be found, but they have not.



Fig. 1. The projection table in use, with the mid depth texture gradient pictorial display back-projected onto the 1000*750 mm screen. During the experiments the laboratory lights were switched off to increase the saliency of the display, and the participant wore LCD glasses, gating their view of the apparatus.

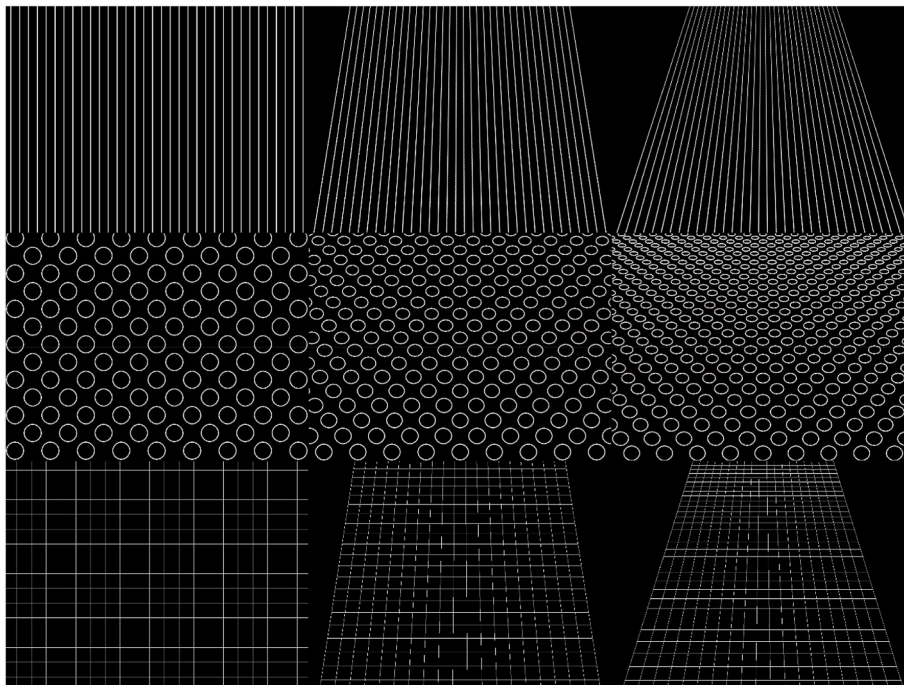


Fig. 2. The three sets of pictorial depth backgrounds in these experiments: linear perspective (top row); texture gradient (middle row); and grid pattern combining linear perspective and texture (bottom row). Three levels of pictorial depth were used: baseline (left column) with 0% convergence across the front-to-back axis (i.e. parallel elements); medium depth (middle column) with 25% convergence front-to-back; high depth (right column) with 50% convergence. The Pictorial Convergence at the level of the target emerged from the interaction of target distance with pictorial depth level (see [Table 1](#) and [Fig. 3b](#)).

mm; 360 mm), which were spaced evenly within the range of distances used in the other experiments. Each participant performed two blocks of 72 trials. Each block had one trial for each combination of delay (no-delay; RT-delay; 2s-delay; 5s-delay), target distance (280 mm; 360 mm), pictorial depth (baseline; medium; high), and

target size (small; medium; large). Each participant therefore performed a total of 144 trials.

Experiment 4 tested the influence of visual feedback, and whether participants could learn to down-weight pictorial depth cues when different forms of feedback were available. Feedback condition was a

between-subjects manipulation, with 10 participants in each of three feedback groups. The open-loop group performed the standard immediate reaching task, with visual feedback occluded at the start of the movement. The terminal-fb group performed the same task, except that the total stimulus display period was extended to 3 s, and the LCD glasses re-opened for the final second, so that participants could see their finger's terminal position in relation to the target. The closed-loop group had visual feedback continually available throughout the movement (the LCD glasses never closed). Each participant performed two blocks of 81 trials, with three repetitions for each combination of target distance (near; mid; far), pictorial depth (baseline; medium; high), and target size (small; medium; large). Each participant therefore performed a total of 162 trials.

Experiment 5 assessed the influence of pictorial depth cues on reach amplitude in patient DF, who has visual form agnosia. She was tested in the open-loop condition of Experiment 4, which means that she completed two blocks of the 81 trials of the standard immediate open loop reaching task, using the grid type of pictorial background, for a total of 162 trials.

2.4. Kinematic data processing

Kinematic data were recorded with the XY plane aligned with the display screen, and the Z plane orthogonal to it (aligned with gravitational vertical). The raw positional data were processed via custom analysis scripts in LabVIEW (National Instruments), being first filtered by a dual pass through a Butterworth filter with a cut-off of 20 Hz, and then differentiated to 3D speed. Movement onset was defined as the first frame in which the speed of the marker exceeded 50 mm/s, provided that this was maintained for at least 100 ms. Movement offset was defined as the first subsequent frame in which the speed of the marker fell back below 50 mm/s. The dependent measure of interest was the Amplitude of the reach, which was defined as the XY displacement of the marker from the start button at movement offset.

The following standard kinematic variables were also extracted by the analysis script: Reaction Time (RT) from start tone to movement onset; Movement Time (MT) from movement onset to offset; Peak Speed (PS), Time to Peak Speed (TPS) from movement onset; Peak Acceleration (PA); Time to Peak Acceleration (TPA) from movement onset. These variables were not analysed, because they were of little interest (RT, TPS, TPA) and/or were likely to be correlated with Amplitude but with greater variability (MT, PS, PA). All kinematic variables are included in the data archive for this study: <https://osf.io/ckv3w/>.

Trials were deemed invalid if missing samples precluded the unambiguous identification of movement onset or offset, if the RT was less than 100 ms (anticipation), or if the Amplitude was less than 100 mm (false start). Invalid trials accounted for 1.7% of trials overall: Experiment 1, 0.6%; Experiment 2, 3.7%; Experiment 3, 1.6%; Experiment 4, 1.0%; Experiment 5, 3.1%.

2.5. Statistical analysis

For each trial, we subtracted out that participant's mean Amplitude for the baseline depth background at the same target distance, thereby re-expressing Amplitude as Overshoot relative to baseline mean Amplitude for that distance. This numerically cancelled out the very large but uninteresting main effects of target distance, allowing us to focus on variations in reaching associated with pictorial depth and any other factors in the experiment. It also enabled us to recode the independent variables of target distance and pictorial depth, into a single higher-order variable (Pictorial Convergence), as explained below.

Target distance and pictorial depth were *manipulated* orthogonally in our design, but these two values interactively determine how converged the background is at the point that the target is located, and thus the *effective* pictorial depth context for each trial. The baseline pictorial depth display converges by 0% from front-to-back edges (i.e. the

Table 1

Pictorial Convergence of depth background at the position of the target for each combination of depth background and target distance (distances of 220, 320 and 420 mm were used in Experiments 1,2 and 4, and distances of 280 and 360 mm were used in Experiment 3). Pictorial Convergence is given by the total front-to-back convergence of the (baseline, medium or high) display multiplied by the proportional position of the target along the 750 mm front-to-back axis of the display. For instance, at 300 mm along the 750 mm display, the medium depth background has 10% Convergence [= 25% * (300/750)] and the high depth background has 20% Convergence [= 50% * (300/750)]. Note that an equal level of 13.33% Pictorial Convergence is obtained for the 220 mm target distance on the high depth background and the 420 mm target distance on the medium depth background.

| Target distance from start button (mm) | Target distance along display (mm) | Pictorial Convergence of depth background at position of target (%) | | |
|--|------------------------------------|---|--------------|------------|
| | | Baseline (0%) | Medium (25%) | High (50%) |
| 220 | 200 | 0 | 6.67 | 13.33 |
| 280 | 260 | 0 | 8.67 | 17.33 |
| 320 | 300 | 0 | 10.00 | 20.00 |
| 360 | 340 | 0 | 11.33 | 22.67 |
| 420 | 400 | 0 | 13.33 | 26.67 |

elements are parallel); the medium depth version converges by 25%; and the high depth version converges by 50%. If a reaching target is presented half-way along the front-to-back axis of the display, then the Pictorial Convergence at the level of the target will be 0% for the baseline condition, 12.5% for the medium depth condition, and 25% for the high depth condition. By the same token, if a target is presented one-quarter of the way along the high depth display, the Pictorial Convergence will be the same as if it were presented half-way along the medium depth display (12.5%). Table 1 gives the Pictorial Convergence created by each combination of pictorial depth and target distance across the five experiments (see also Fig. 3b).

Our original analysis plan was to include target distance and pictorial depth as independent variables, but an overall analysis of the data (see Results Section 3.2) shows that their influences on reach Overshoot can be fully and succinctly captured in terms of Pictorial Convergence (which is a linear combination of distance and depth). Statistical analyses for the individual experiments are made more parsimonious and more directly interpretable by using the higher-order variable of Pictorial Convergence in place of the manipulated variables of target distance and pictorial depth.

Statistical analyses were performed in R (R Core team, 2019). Separate ANOVAs were run for Experiments 1–4 (using the *ez* package: Lawrence, 2016), analysing the within-subjects effect of Pictorial Convergence, and any additional within- or between-subjects factors manipulated in that experiment, on overshoot. For Experiment 5 (patient DF), a factorial ANOVA by Pictorial Convergence and trial block was run (with Type 3 sums of squares), treating trials as independent observations. Pictorial Convergence was a continuous predictor, and all other factors were categorical. Data were plotted using the *ggplot2* package (Wickham, 2016; Wickham et al., 2023).

2.6. Sample size and sensitivity

The sample sizes for these experiments were based on precedent for the field at the time (e.g. Gonzalez et al., 2006; Marotta and Goodale, 2001; Westwood and Goodale, 2003), not on *a priori* power analyses. Given the modest sample sizes, these experiments would have high power (80%) to detect only large main effects of Pictorial Convergence (e.g. $d_z > 0.7$ or $r > 0.6$ for $n = 17$ in Experiment 3) (Faul et al., 2007). This may be reasonable for this main effect of interest, given the salient wide-field pictorial depth backgrounds we used; and strong effects of Pictorial Convergence were indeed observed in all experiments. However, it should be noted that there will be lower power to detect

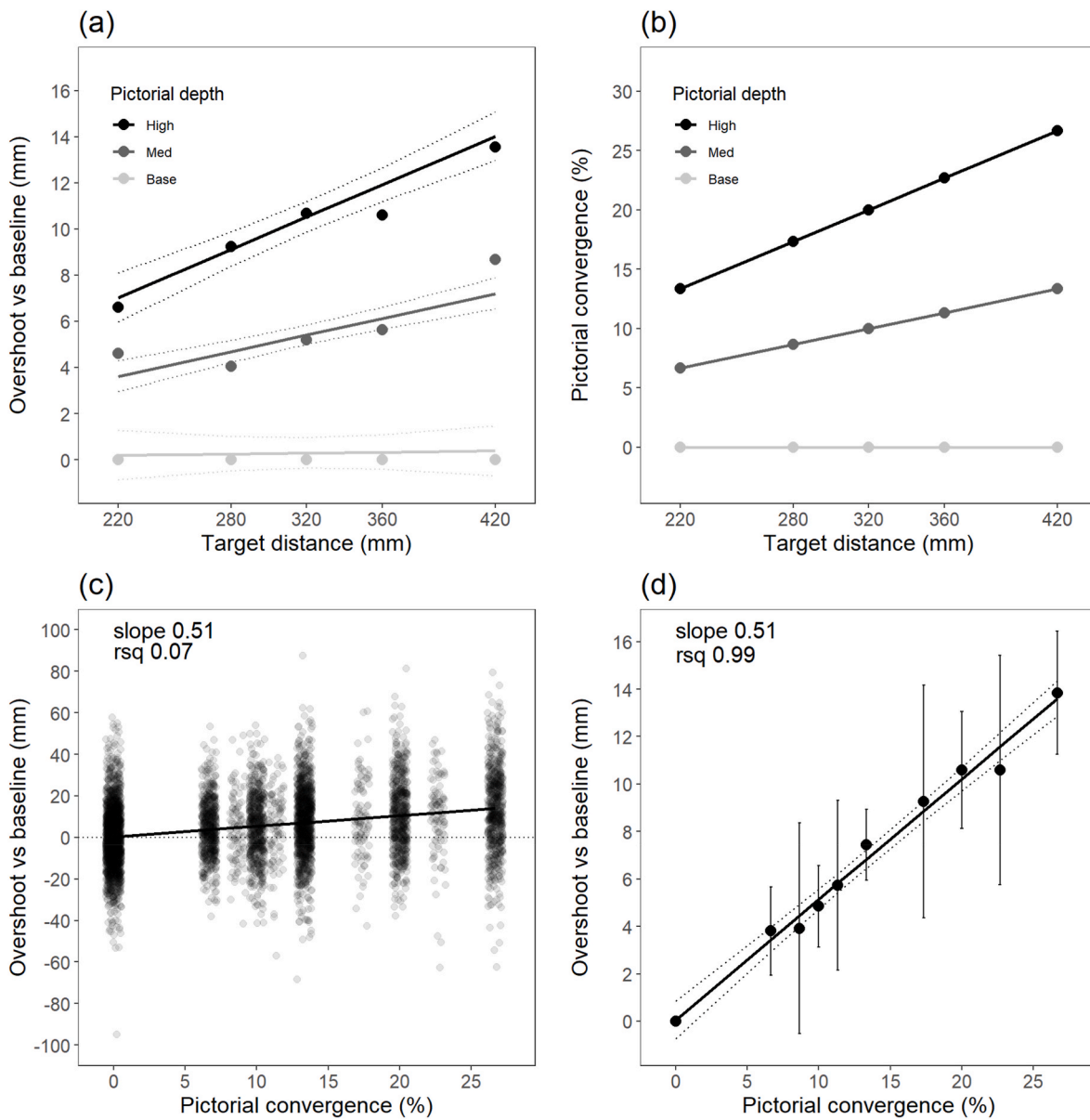


Fig. 3. Overall analysis of immediate right hand open-loop reaching with binocular viewing across Experiments 1–4 (distances of 220, 320 and 420 mm were used in Experiments 1, 2 and 4, and distances of 280 and 360 mm were used in Experiment 3). **(a)** Overshoot relative to baseline depth condition, by distance and depth. The lines show the best fitting linear model, with 95% CIs indicated, and the points are the observed group means. **(b)** Pictorial Convergence at the level of the target, for each combination of pictorial depth and target distance (i.e. this is a plot of the values in Table 1). The pattern matches that in Panel (a), implying that the interactive influences of depth and distance can be more succinctly captured in terms of Pictorial Convergence. **(c)** Linear model fitting Overshoot to Pictorial Convergence across all trials (6000 trials across 63 participants). **(d)** Linear model fitting Overshoot to Pictorial Convergence across group means, with 95% CIs indicated for the fit line and group means. Note that the sizes of the CIs differ because the number of trials collected across experiments at each level of convergence vary.

higher-level (and presumably smaller) interactions with other manipulations, particularly those involving between-subjects factors.

However, despite the modest sample sizes *per experiment*, the same immediate open-loop reaching condition with binocular viewing and using the right hand was included in all four group experiments (Experiments 1–4), with only minor variations in the background type (linear perspective and circle gradient in Experiment 1; grid gradient in Experiments 2–4), target distances (220, 320 and 420 mm in Experiments 1, 2 and 4; 280 and 360 mm in Experiment 3), and trial numbers (162 trials in Experiments 1 and 4; 81 trials in Experiment 2; 72 trials in Experiment 3). This makes possible an initial overall analysis of data across the group experiments, to provide a more reliable, quantitative estimate of the influence of pictorial depth on immediate open-loop reaching.

3. Results

3.1. Open-data statement

The inferential analyses focus on the dependent variable of reach Overshoot relative to baseline mean (see Methods Section 2.5), but complete data with all kinematic variables are archived with analysis scripts at <https://osf.io/ckv3w/>.

3.2. Overall analysis of immediate open-loop reaching

We begin with an overall analysis of the immediate open-loop

condition with binocular viewing and right hand reaching, to set the general context for the separate experiments reported subsequently.³ Across the four group experiments, there were 6000 valid trials for this condition, from 63 healthy participants ($n = 12$ from Binocular group of Experiment 1; $n = 24$ from Experiment 2; $n = 17$ from Experiment 3; $n = 10$ from open-loop group of Experiment 4). For each trial, reach Amplitude was re-coded as Overshoot relative to the mean of the baseline depth trials for the same participant at the same target distance. This focuses the analysis on how much further than usual participants reach when viewing targets on exaggerated pictorial depth displays. The mean Overshoot for the baseline pictorial depth condition is constrained by definition to zero at every target distance (see Fig. 3a). Nonetheless, these baseline conditions contain important information about the variability of reach Overshoot, and so are included in our statistical models of the data below.

We initially followed the original plan of analysing the influence of the *manipulated* variables of target distance and pictorial depth. We used a linear model with distance and depth as predictors, including their interaction, and Overshoot as the dependent measure. Note that this model does not explicitly account for the nesting of observations within participants, but the coding of the Overshoot variable zeroes the mean values per participant in the baseline condition, which removes idiosyncratic variability in Overshoot. Given this coding, it would be of little benefit to add a random intercept for participants, and so we used a simple fixed-effects linear model.

The best fitting model is shown by the lines plotted in Fig. 3a [$F(3, 5996) = 155, p = 2.2 \times 10^{-16}$, adjusted $r^2 = 0.07$, AIC 50663], with the observed mean Overshoot values overlaid as points. There is a strong interaction of distance by depth [$\beta = 0.02$, SE 0.003, $t(5996) = 5.17, p = 2.4 \times 10^{-7}$], and no significant effect for either distance [$\beta = 0.0009$, SE 0.004, $t(5996) = 0.21, p = .84$] or depth [$\beta = -0.33$, SE 1.09, $t(5996) = -0.31, p = .76$], above and beyond their interaction. The interaction is that the influence of exaggerated pictorial depth increases with target distance, with an increasing slope for Medium and High depth backgrounds.

However, there is a strikingly close similarity between the pattern of interaction in Fig. 3a, and the variations of Pictorial Convergence that emerge from the combinations of target distance and pictorial depth and target distance (compare Fig. 3b, which is simply a plot of the values derived in Table 1). This implies that the combined effects of distance and depth may be more succinctly captured in terms of Pictorial Convergence. A linear model to predict Overshoot from Pictorial Convergence showed that this single predictor [$\beta = 0.51$, SE 0.02, $t(5998) = -21.6, p = 2.2 \times 10^{-16}$] explained the same amount of variance as the model with depth and distance (R-squared 0.07) but had a slightly lower AIC (Δ AIC 3.9) reflecting its greater parsimony.

Pictorial Convergence explains only a relatively small amount of variance when considered in the context of all sources of trial-to-trial variation, as shown in Fig. 3c. High trial-to-trial variability is unsurprising given that participants were reaching without visual feedback. Nonetheless, Pictorial Convergence explains essentially all the variation in the group average Overshoot values, as shown in Fig. 3d. The slope of the fitted relationship means that, across the open-loop immediate right-hand reaching conditions of Experiments 1–4, reach amplitude increased by an estimated 0.51 mm for every additional percentage point of convergence of the pictorial background around the target.

This provides a clear answer to our main question of whether pictorial depth influences reach amplitude: it does so reliably, and the influence is linearly related to the degree of Pictorial Convergence

around the target. No additional explanatory value is gained by considering depth and distance as separate factors, an observation that we shall return to in Discussion. Given this pattern, it is simpler and more directly interpretable to substitute the higher-order variable of Pictorial Convergence for the manipulated variables of target distance and pictorial depth, in analysing whether this core effect was modulated by the manipulations introduced in the individual experiments.

3.3. Analysis of individual experiments

For the ANOVAs of Experiments 1–5, we report all statistically significant effects ($p < .05$), and (whether significant or not) additional effects of specific theoretical interest, usually the interaction of Pictorial Convergence with a factor proposed to modulate its influence.

Experiment 1 results are shown in Fig. 4a. The ANOVA had the between-subjects factor of viewing condition (monocular; binocular) and the within-subjects factors of Pictorial Convergence, and depth type (linear perspective; texture gradient). The main effect of Pictorial Convergence ($F_{(1, 22)} = 107.93, p = 6.0 \times 10^{-10}, \eta^2G = 0.33$) interacted significantly with depth type ($F_{(1, 22)} = 4.38, p = .048, \eta^2G = 0.02$). The interaction reflected stronger effects of Pictorial Convergence for the texture gradient (slope = 0.71) than the linear perspective background (slope = 0.52). (For subsequent experiments, we switched to the grid background type, which combines linear perspective with textural elements.) There was no indication that pictorial depth was more influential with monocular viewing ($F_{(1, 22)} = 0.27, p = .61, \eta^2G = 0.001$).

Experiment 2 results are shown in Fig. 4b. The ANOVA had the between-subjects factor of handedness (left-hander; right-hander) and the within-subjects factors of Pictorial Convergence, and responding hand (left hand; right hand). There was a strong main effect of Pictorial Convergence ($F_{(1, 22)} = 93.70, p = 2.2 \times 10^{-9}, \eta^2G = 0.14$, slope = 0.51), and a significant effect of hand dominance ($F_{(1, 22)} = 5.68, p = .03, \eta^2G = 0.02$), with slightly stronger overshoot effects of pictorial backgrounds in left-handers. However, there was no significant interaction of hand used by Pictorial Convergence ($F_{(1, 22)} = 1.89, p = .18, \eta^2G = 0.004$), giving no support to the idea that left hand use should lead to stronger effects of pictorial context (Gonzalez et al., 2006).

Experiment 3 results are shown in Fig. 5a. The ANOVA had the within-subjects factors of Pictorial Convergence and delay condition (no-delay; RT-delay; 2s-delay; 5s-delay). There were main effects of Pictorial Convergence ($F_{(1, 16)} = 37.97, p = 1.4 \times 10^{-5}, \eta^2G = 0.46$, slope = 0.52), and delay ($F_{(1, 16)} = 119.91, p = 7.7 \times 10^{-9}, \eta^2G = 0.81$), reflecting relatively increased undershoot with increasing delay. The interaction between these factors was not significant ($F_{(1, 16)} = 4.38, p = .053, \eta^2G = 0.02$), and gave no support to the idea that pictorial cues are more influential when actions are made on the basis of visual memory (Gentilucci et al., 1996; Goodale et al., 2004; Hu and Goodale, 2000; Westwood and Goodale, 2003).

Experiment 4 results are shown in Fig. 5b. The ANOVA had the between-subjects factor of feedback condition (open-loop; terminal-fb; closed-loop) and the within-subjects factors of Pictorial Convergence, and trial block (Block 1; Block 2). The significant main effects of Pictorial Convergence ($F_{(1, 27)} = 98.94, p = 1.6 \times 10^{-10}, \eta^2G = 0.47$) and feedback condition ($F_{(2, 27)} = 16.13, p = 2.5 \times 10^{-5}, \eta^2G = 0.28$) were modified by the interaction between these factors ($F_{(2, 27)} = 11.12, p = 3.0 \times 10^{-4}, \eta^2G = 0.17$). This interaction arose because the effect of pictorial depth was attenuated in the closed-loop condition, when online feedback was available (slope = 0.13), relative to the terminal feedback and open-loop conditions (slopes = 0.53 and 0.39). This implies that participants were able to use online feedback to minimise errors in the closed-loop trials, thereby limiting the influence of illusory depth. Nonetheless, there was no significant interaction of Pictorial Convergence with trial block ($F_{(1, 27)} = 0.01$,

³ We used the right-hand reaching data even for the left-handers ($n = 12$) of Experiment 2, because right hand reaching has been argued to be more natural and fluent, regardless of hand dominance (Gonzalez et al., 2006; Goodale, 2008b). The outcomes of this analysis would have been closely similar if we had used their left-hand reaching data, or if we had excluded these participants.

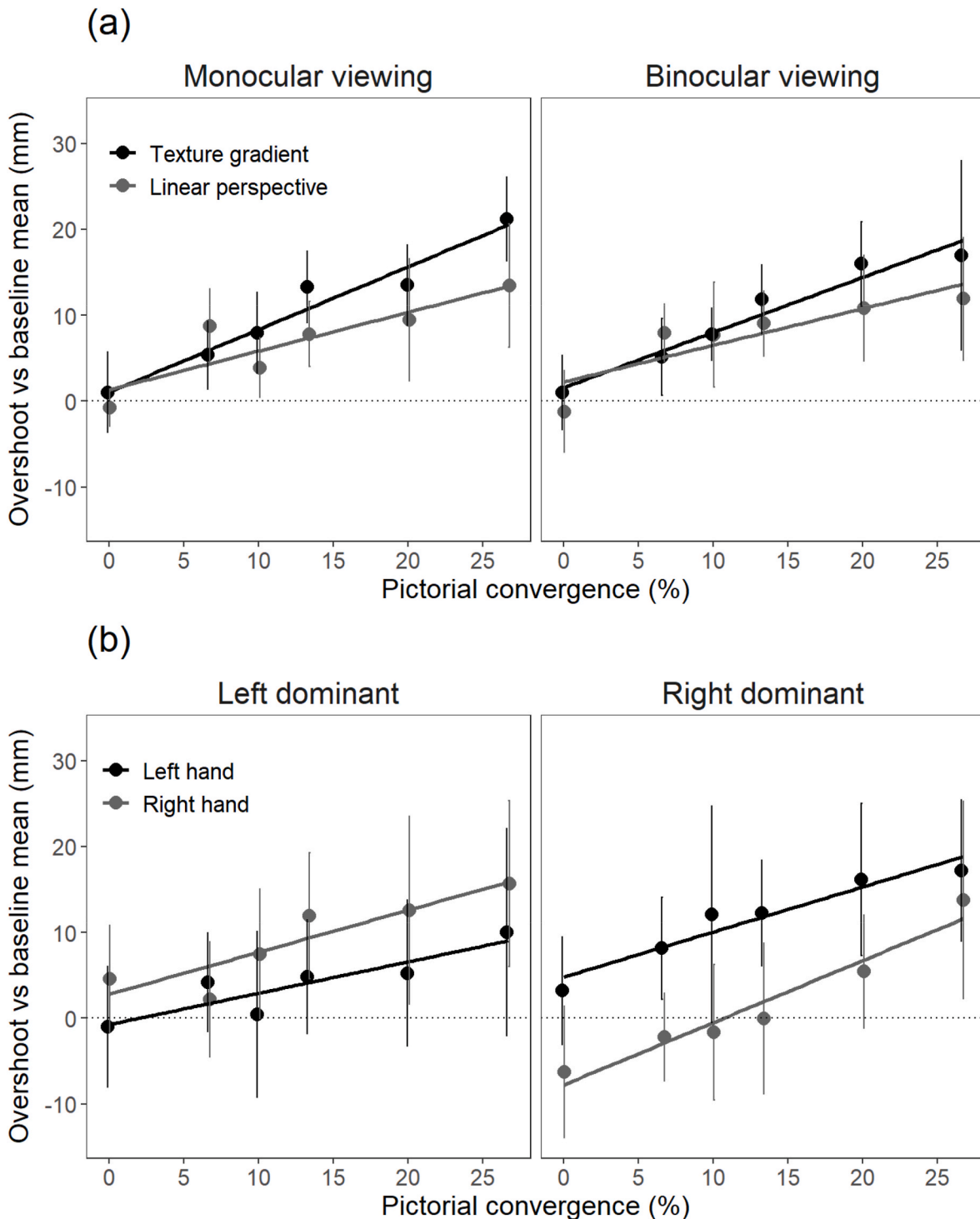


Fig. 4. (a) Results of Experiment 1: Overshoot vs baseline mean by Pictorial Convergence by pictorial background type (linear perspective or texture gradient) for each viewing group ($n = 12$ per group). (b) Results of Experiment 2: Overshoot vs baseline mean by Pictorial Convergence (of grid background) by responding hand for each handedness group ($n = 12$ per group). Error bars indicate 95% CIs for within-subjects designs (Morey, 2008).

$p = .93$, $\eta^2G = 0.00$), or three-way interaction with feedback condition and trial block ($F_{(2, 27)} = 0.12$, $p = .88$, $\eta^2G = 0.00$). Therefore, participants showed no sign of learning to ignore pictorial depth cues when online or terminal feedback was available.

Experiment 5 (Visual Form Agnosia). DF's results in two blocks of open loop immediate reaching are shown in Fig. 6a. Her reaches tended towards undershoot in the second block relative to the first, perhaps due to fatigue, but both blocks showed a clear influence of

pictorial depth. These patterns were confirmed by the factorial ANOVA, which found significant effects of Pictorial Convergence ($F_{(1, 153)} = 43.79$, $p = 5.8 \times 10^{-10}$, $\eta^2G = 0.22$, slope = 0.56) and block ($F_{(1, 153)} = 72.5$, $p = 1.45 \times 10^{-14}$, $\eta^2G = 0.32$). As illustrated in Fig. 6b, the slope of the effect of Pictorial Convergence on DF's reaching (0.56) was slightly larger the control mean value (0.51) derived from the overall analysis reported in Section 3.2. This

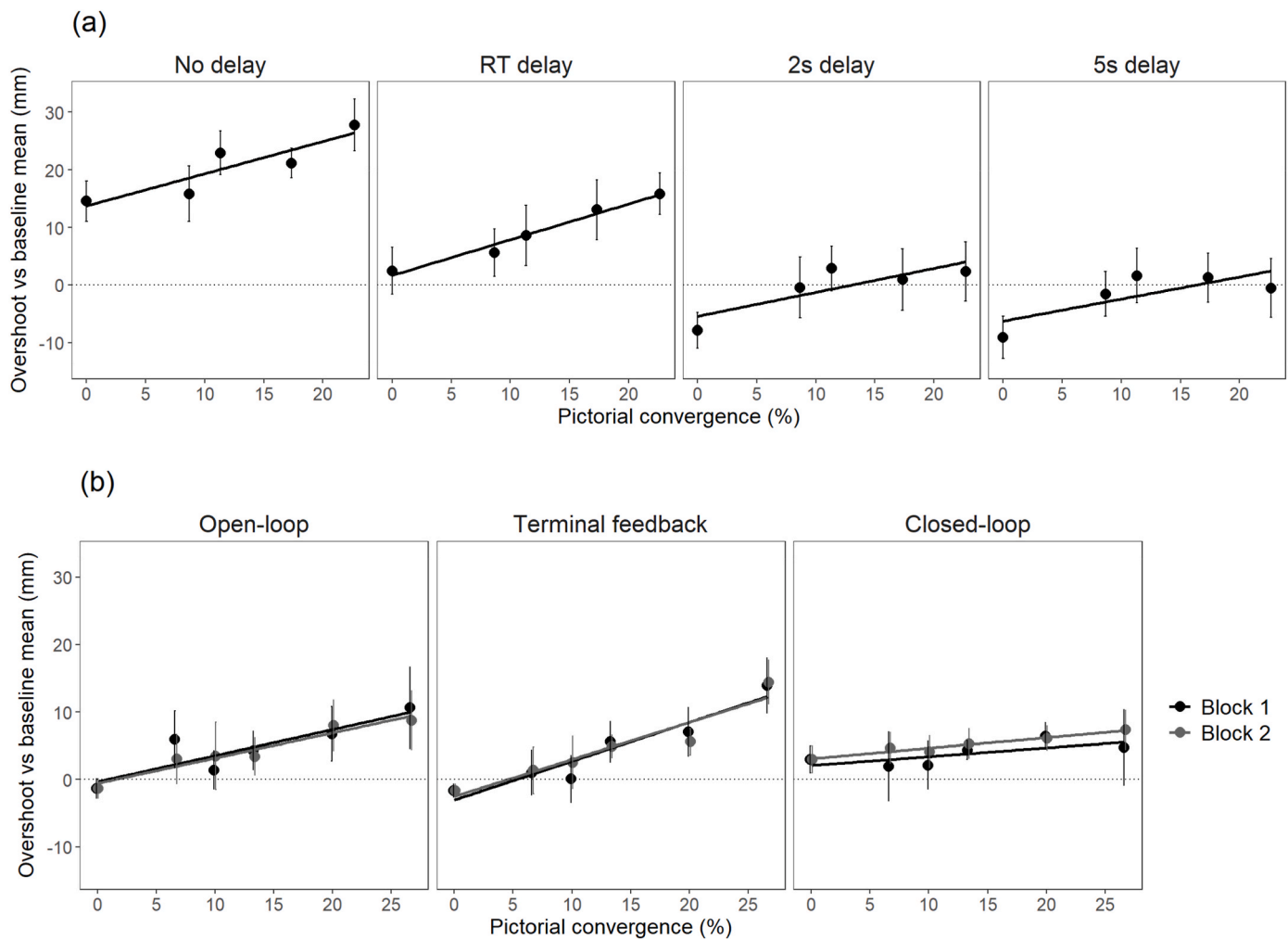


Fig. 5. (a) Results of Experiment 3: Overshoot vs baseline mean by Pictorial Convergence (of grid background) by Delay condition ($n = 17$). Error bars indicate 95% CIs. (b) Results of Experiment 4: Overshoot vs baseline mean by Pictorial Convergence (of grid background) by responding hand for each feedback group ($n = 10$ per group). Error bars indicate 95% CIs for within-subjects designs (Morey, 2008).

provides no support for the idea that damage to the ventral stream should reduce the influence of pictorial depth cues on action.

4. Discussion

The main question for this study was whether pictorial depth cues based on relative size, such as linear perspective and texture gradient, influence reaching distance. A clear positive answer was found, with exaggerated pictorial depth causing participants to reach further under all conditions of the four group experiments, and in a single case assessment of the visual form agnostic patient DF. The effect of pictorial depth was succinctly captured in terms of the variable Pictorial Convergence, which emerges from the interaction of the *manipulated* variables of pictorial depth level and target distance, and encodes the amount of convergence of the background immediately surrounding the target. For the standard immediate open-loop reaching condition (binocular, right hand), reach amplitude increased on average by 0.51 mm for every additional percent of Pictorial Convergence. This slope value, estimated from an overall analysis of the responses of 63 healthy participants across Experiments 1–4, was quite closely matched by the estimated slope for patient DF in Experiment 5.

Unlike many ‘illusions-in-action’ studies, our experiments did not seek to compare visuomotor and perceptual tasks. This approach was taken to sidestep the fraught issue of how to match perceptual and visuomotor tasks meaningfully, and to focus on the more fundamental

question of whether pictorial depth affects action. However, we did include manipulations that have been hypothesised to modulate the influence of ventral stream processing on visuomotor responses. We compared reaching under monocular and binocular viewing (Experiment 1), with the right and the left hand (Experiment 2), and reaching immediately and guided by visual memory (Experiment 3). If these manipulations modulate ventral stream influence, and the ventral stream emphasises pictorial cues, then this should predict more pronounced effects of pictorial depth with monocular viewing (e.g. Marotta and Goodale, 2001), with left hand use (e.g. Gonzalez et al., 2006), and in delayed reaching (Goodale et al., 2004; Westwood and Goodale, 2003). However, the effects of pictorial depth were statistically indistinguishable regardless of the manipulation applied.

One major caveat is that ‘statistically indistinguishable’ means that no difference could be detected in the modest sample sizes used (Experiment 1, $n = 12$ per viewing condition; Experiment 2; $n = 24$; Experiment 3, $n = 17$). These sample sizes are comparable to the key studies on which the predictions for Experiments 1–3 were based (e.g. Gonzalez et al., 2006; Marotta and Goodale, 2001; Westwood and Goodale, 2003), but they imply that we can rule out only very large changes in the weighting given to pictorial depth, and might have failed to detect more subtle modulations. It is also possible that some or all of these manipulations were ineffective in promoting ventral stream influence. Our monocular condition might have been insufficient to increase the reliance on pictorial cues because, although it eliminated the

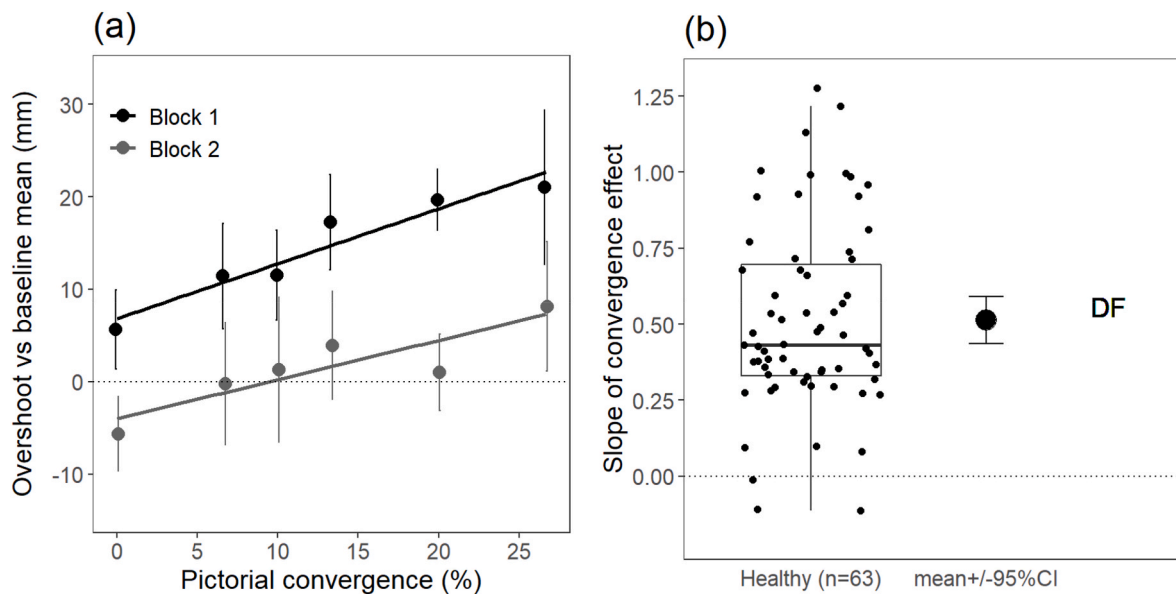


Fig. 6. (a) Results of Experiment 5, in which DF performed the open-loop condition of Experiment 4, showing Overshoot vs baseline mean by Pictorial Convergence (of grid background). Error bars indicate 95% CIs. (b) Comparison of the slope of the Pictorial Convergence effect between DF and the 63 healthy participants performing right hand immediate open-loop reaching across Experiments 1–4. The effect on DF's reaching behaviour is closely comparable to the average effect in healthy participants.

extra-retinal cue of vergence angle, participants could still potentially use vertical gaze angle to gauge absolute distance (Gardner and Mon-Williams, 2001; Mon-Williams et al., 2001a, 2001b). For the hand manipulation, the proposed link between ventral stream processing and left hand use is based on limited evidence (Gonzalez et al., 2006), and may anyway be more relevant to grasping, which requires much more digital dexterity than simple reaching. Finally, some authors have contested Westwood and Goodale's (2003) idea that interposing a delay causes a switch from dorsal to ventral modes of control, suggesting instead that delay merely causes decay of the memory trace, leading to a loss of precision of responding (Franz et al., 2009; Hesse and Franz, 2009). Reduced precision was visible in Experiment 3 as increasing relative undershoot with increasing memory delay (Fig. 5a).

As well as being consistent across experiments, the visuomotor influence of pictorial depth was persistent, being just as visible in a second block of trials as in an initial block (Fig. 5b). This is unsurprising for the standard open-loop feedback condition, because the participant would have no opportunity to observe their reaching errors and discover that pictorial depth was misleading. It is more surprising that there was no down-weighting of pictorial depth when terminal visual feedback allowed participants to see the outcome of each reach. Only the closed-loop condition, in which visual feedback was continuously available, showed a reduction in the effect of pictorial cues. This can be attributed to online correction improving terminal accuracy under these full-feedback conditions. But even in this closed-loop condition, the effect of pictorial depth appeared to be as strong in the second block as in the first, suggesting no down-weighting of pictorial cues with exposure to errors. From this perspective, it does not seem that the influence of pictorial depth is easily eliminated, or separated out from other cues informing reach distance.

We also had the opportunity to test the visual form agnostic patient DF, whose severe perceptual impairments should make her unable to perceive the visual forms within the depth displays (e.g. the shape and size of the grid elements). We did confirm informally that DF was unable to describe the stimuli, seeing them only as indistinct blurs of light. Nonetheless, her reaching responses showed a robust influence of Pictorial Convergence, closely comparable to that observed in healthy participants. In recent years, it has emerged that the specificity of DF's brain damage to ventral stream structures is less complete than has

sometimes been assumed, with areas of degeneration also within the dorsal stream (Bridge et al., 2013; James et al., 2003). DF has now been found to show visuomotor problems consistent with this dorsal stream involvement, including misreaching in peripheral vision and a lack of fast online corrections (Hesse et al., 2012, 2014; Rossit et al., 2018). Even so, given her more extensive ventral stream lesions, and profound perceptual problems, DF's apparently normal susceptibility to pictorial depth might indicate that these contextual effects arise within the dorsal stream itself, or in earlier visual areas.

Perhaps the most remarkable discovery of this experiment is that the influence of pictorial depth, across a range of reachable distances, seems to be captured simply in terms of the degree of Pictorial Convergence around the target. For instance, a convergence of 13.3% caused a similar overshoot, relative to a baseline (normal) depth display, whether that level of convergence was created by a high depth display at a near distance, or a medium depth display at a far distance (see Table 1, and Fig. 3a). This is surprising, because one would intuitively expect a given degree of convergence to always imply the same *relative* increase in depth (e.g. 50% convergence implies a doubling of depth), which should translate into a greater absolute increase when the viewing distance is greater. The viewing distance in our set-up is from the participant's eye-level, which is not the same as the target distance from the start button, but viewing distance did always increase with target distance. Nonetheless, the influence of Pictorial Convergence did not increase, but seemed to be a simple linear function, with every extra degree adding around half a millimetre to the distance reached.

This surprisingly simple relationship will require further research to corroborate and to test the generality of across different viewing and response arrangements. Several factors may need to be considered. Pictorial Convergence is only one cue to depth in the display, and the subjective depth illusion it creates is not of greater distance in the plane of the display; the viewer rather has the impression of a slanted surface, receding in depth. But although their perception may be of a slanted surface, the participant knows from experience that it is a horizontal plane that they cannot reach through. This physical constraint may considerably limit the influence that pictorial depth can have. If so, then the observation of a distance-independent influence of Pictorial Convergence may be relatively specific to our set-up, and might not generalise to situations with fewer physical constraints (such as virtual

set-ups). Moreover, it should be acknowledged that, although consistent, the biasing influence of Pictorial Convergence is metrically small, especially considering the striking subjective illusion of depth when looking at these displays *in situ*. This suggests that, although pictorial cues have statistically robust effects on action, they may be weighted only weakly into action plans, perhaps much less than into perceptual experience (cf. Knill, 2005).

Finally, although we manipulated the effects of pictorial depth in terms of the idealised depth cues of linear perspective and texture gradient, it is possible that the results are influenced by a more general factor of visual density. For any given target position in a converging display, there will more visual content between the viewer and the target than there would be in the non-converging version. Thus, in addition to the lawful changes of size and shape of the textural elements, there is just more visual ‘stuff’ between the viewer and the target in the converging version. It is possible that visual density could account for some of the effect of pictorial displays, as a form of Oppel-Kundt illusion, whereby the presence of more visual elements within a space leads to it to be perceived as more extensive. This might also help explain the fact that the circle texture gradient background was more effective than the sparser linear perspective background (Fig. 4a). We did consider having an additional, reversed-depth condition in which the display converged towards the viewer. This would boost the density of the display between viewer and target, but in the direction opposite to the (reversed) depth cue, potentially teasing apart any general effect of density from the specific influence of the depth cue. We decided against this manipulation, because reversed pictorial depth violates ecological expectations, so its influence could be hard to interpret. In any case, whether our effects derive exclusively from idealised depth cues, or are modulated by visual density is not critical for confirming a general influence of pictorial depth on reaching.

Overall, the main conclusion is clear: pictorial depth cues influence reach distance. Given the salient pictorial depth in our displays, this conclusion is perhaps unsurprising. Somewhat more surprising is that the observed influence was a linear function of Pictorial Convergence, and did not seem to depend upon the target distance. Whether this observation generalises beyond our specific set-up remains to be seen. Not only was the effect of pictorial depth robust, in the sense of being replicable across experiments, it was also impervious to a range of different experimental manipulations, including monocular viewing, and organic damage to ventral stream areas critical for size and shape processing. Whether the influence of pictorial depth is really immune to all such factors, or just shows more subtle modulations than we could detect here, will require further, higher-powered experiments to resolve.

CRediT authorship contribution statement

Robert D. McIntosh: Conceptualization, Methodology, Software, Formal analysis, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision. **Matthew H. Iveson:** Investigation, Writing – review & editing, Project administration. **Sebastian Sandoval Similá:** Investigation, Writing – review & editing, Project administration. **Antimo Buonocore:** Software, Writing – review & editing.

Data availability

Data and code shared at <https://osf.io/ckv3w/>

Acknowledgements

The creation of the projection table was funded by a Small Project Grant (HSS/2678/102006) from the University of Edinburgh Development Trust. We are grateful to Dougie Howie and Hugh Cameron for the detailed design and construction of this equipment.

References

- Aglioti, S., DeSouza, J.F.X., Goodale, M.A., 1995. Size-contrast illusions deceive the eye but not the hand. *Curr. Biol.* 5 (6), 679–685. [https://doi.org/10.1016/S0960-9822\(95\)00133-3](https://doi.org/10.1016/S0960-9822(95)00133-3).
- Bozzacchi, C., Volcic, R., Domini, F., 2014. Effect of visual and haptic feedback on grasping movements. *J. Neurophysiol.* 112 (12), 3189–3196. <https://doi.org/10.1152/jn.00439.2014>.
- Bridge, H., Thomas, O.M., Minini, L., Cavina-Pratesi, C., Milner, A.D., Parker, A.J., 2013. Structural and functional changes across the visual cortex of a patient with visual form agnosia. *J. Neurosci.* 33 (31), 12779–12791.
- Bruno, N., 2001. When does action resist visual illusions? *Trends Cognit. Sci.* 5 (9), 379–382. [https://doi.org/10.1016/S1364-6613\(00\)01725-3](https://doi.org/10.1016/S1364-6613(00)01725-3).
- Bruno, N., Bernardis, P., Gentilucci, M., 2008. Visually guided pointing, the Müller-Lyer illusion, and the functional interpretation of the dorsal-ventral split: conclusions from 33 independent studies. *Neurosci. Biobehav. Rev.* 32 (3), 423–437. <https://doi.org/10.1016/j.neubiorev.2007.08.006>.
- Bruno, N., Franz, V.H., 2009. When is grasping affected by the Müller-Lyer illusion?: a quantitative review. *Neuropsychologia* 47 (6), 1421–1433. <https://doi.org/10.1016/j.neuropsychologia.2008.10.031>.
- Carey, D.P., 2001. Do action systems resist visual illusions? *Trends Cognit. Sci.* 5 (3), 109–113. [https://doi.org/10.1016/S1364-6613\(00\)01592-8](https://doi.org/10.1016/S1364-6613(00)01592-8).
- Carey, D.P., Dijkerman, H.C., Milner, A.D., 1998. Perception and action in depth. *Conscious. Cognit.* 7 (3), 438–453. <https://doi.org/10.1006/ccog.1998.0366>.
- Dijkerman, H.C., Milner, A.D., Carey, D.P., 1996. The perception and prehension of objects oriented in the depth plane. *Exp. Brain Res.* 112 (3), 442–451. <https://doi.org/10.1007/BF00227950>.
- Dijkerman, H.C., Milner, A.D., Carey, D.P., 1999. Motion parallax enables depth processing for action in a visual form agnostic when binocular vision is unavailable. *Neuropsychologia* 37 (13), 1505–1510. [https://doi.org/10.1016/S0028-3932\(99\)00063-9](https://doi.org/10.1016/S0028-3932(99)00063-9).
- Doherty, M.J., Campbell, N.M., Tsuji, H., Phillips, W.A., 2010. The Ebbinghaus illusion deceives adults but not young children. *Dev. Sci.* 13 (5), 714–721. <https://doi.org/10.1111/j.1467-7687.2009.00931.x>.
- Faul, F., Erdfelder, E., Lang, A.-G., Buchner, A., 2007. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39 (2), 175–191. <https://doi.org/10.3758/bf03193146>.
- Foley, J.M., Held, R., 1972. Visually directed pointing as a function of target distance, direction, and available cues. *Percept. Psychophys.* 12 (3), 263–268. <https://doi.org/10.3758/BF03207201>.
- Franz, V.H., 2001. Action does not resist visual illusions. *Trends Cognit. Sci.* 5 (11), 457–459. [https://doi.org/10.1016/S1364-6613\(00\)01772-1](https://doi.org/10.1016/S1364-6613(00)01772-1).
- Franz, V.H., 2003. Manual size estimation: a neuropsychological measure of perception? *Exp. Brain Res.* 151 (4), 471–477. <https://doi.org/10.1007/s00221-003-1477-6>.
- Franz, V.H., Bühlhoff, H.H., Fehle, M., 2003. Grasp effects of the Ebbinghaus illusion: obstacle avoidance is not the explanation. *Exp. Brain Res.* 149 (4), 470–477. <https://doi.org/10.1007/s00221-002-1364-6>.
- Franz, V.H., Gegenfurtner, K.R., 2008. Grasping visual illusions: consistent data and no dissociation. *Cogn. Neuropsychol.* 25 (7–8), 920–950. <https://doi.org/10.1080/02643290701862449>.
- Franz, V.H., Gegenfurtner, K.R., Bühlhoff, H.H., Fehle, M., 2000. Grasping visual illusions: No evidence for a dissociation between perception and action. *Psychol. Sci.* 11 (1), 20–25. <https://doi.org/10.1111/1467-9280.00209>.
- Franz, V.H., Hesse, C., Kollath, S., 2009. Visual illusions, delayed grasping, and memory: No shift from dorsal to ventral control. *Neuropsychologia* 47 (6), 1518–1531. <https://doi.org/10.1016/j.neuropsychologia.2008.08.029>.
- Gardner, P.L., Mon-Williams, M., 2001. Vertical gaze angle: absolute height-in-scene information for the programming of prehension. *Exp. Brain Res.* 136 (3), 379–385. <https://doi.org/10.1007/s002210000590>.
- Gentilucci, M., Chieffi, S., Daprati, E., Saetti, M.C., Toni, I., 1996. Visual illusion and action. *Neuropsychologia* 34 (5), 369–376. [https://doi.org/10.1016/0028-3932\(95\)00128-X](https://doi.org/10.1016/0028-3932(95)00128-X).
- Gonzalez, C.L.R., Ganel, T., Goodale, M.A., 2006. Hemispheric specialization for the visual control of action is independent of handedness. *J. Neurophysiol.* 95 (6), 3496–3501. <https://doi.org/10.1152/jn.01187.2005>.
- Goodale, M.A., 2008a. Action without perception in human vision. *Cogn. Neuropsychol.* 25 (7–8), 891–919. <https://doi.org/10.1080/02643290801961984>.
- Goodale, M.A., 2008b. Action without perception in human vision. *Cogn. Neuropsychol.* 25 (7–8), 891–919. <https://doi.org/10.1080/02643290801961984>.
- Goodale, M.A., Meenan, J.P., Bühlhoff, H.H., Nicolle, D.A., Murphy, K.J., Racicot, C.I., 1994. Separate neural pathways for the visual analysis of object shape in perception and prehension. *Curr. Biol.* 4 (7), 604–610. [https://doi.org/10.1016/S0960-9822\(00\)00132-9](https://doi.org/10.1016/S0960-9822(00)00132-9).
- Goodale, M.A., Milner, A.D., 1992. Separate visual pathways for perception and action. *Trends Neurosci.* 15 (1), 20–25. [https://doi.org/10.1016/0166-2236\(92\)90344-8](https://doi.org/10.1016/0166-2236(92)90344-8).
- Goodale, M.A., Milner, A.D., Jakobson, L.S., Carey, D.P., 1991. A neurological dissociation between perceiving objects and grasping them. *Nature* 349 (6305), 6305. <https://doi.org/10.1038/349154a0>.
- Goodale, M.A., Westwood, D.A., Milner, A.D., 2004. Two distinct modes of control for object-directed action. In: *Progress in Brain Research*, vol. 144. Elsevier, pp. 131–144. [https://doi.org/10.1016/S0079-6123\(03\)14409-3](https://doi.org/10.1016/S0079-6123(03)14409-3).
- Gregory, R.L., 1963. Distortion of visual space as inappropriate constancy scaling. *Nature* 199 (4894), 678–680. <https://doi.org/10.1038/199678a0>.
- Haffenden, A.M., Goodale, M.A., 1998. The effect of pictorial illusion on prehension and perception. *J. Cognit. Neurosci.* 10 (1), 122–136.

- Haffenden, A.M., Goodale, M.A., 2000. Independent effects of pictorial displays on perception and action. *Vis. Res.* 40 (10–12), 1597–1607. [https://doi.org/10.1016/S0042-6989\(00\)00056-0](https://doi.org/10.1016/S0042-6989(00)00056-0).
- Haffenden, A.M., Schiff, K.C., Goodale, M.A., 2001. The dissociation between perception and action in the Ebbinghaus illusion: nonillusory effects of pictorial cues on grasp. *Curr. Biol.* 11, 177–181.
- Hesse, C., Ball, K., Schenk, T., 2012. Visuomotor performance based on peripheral vision is impaired in the visual form agnostic patient DF. *Neuropsychologia* 50 (1), 90–97. <https://doi.org/10.1016/j.neuropsychologia.2011.11.002>.
- Hesse, C., Ball, K., Schenk, T., 2014. Pointing in visual periphery: is DF's dorsal stream intact? *PLoS One* 9 (3), e91420. <https://doi.org/10.1371/journal.pone.0091420>.
- Hesse, C., Franz, V.H., 2009. Memory mechanisms in grasping. *Neuropsychologia* 47 (6), 1532–1545. <https://doi.org/10.1016/j.neuropsychologia.2008.08.012>.
- Hu, Y., Goodale, M.A., 2000. Grasping after a delay shifts size-scaling from absolute to relative metrics. *J. Cognit. Neurosci.* 12 (5), 856–868. <https://doi.org/10.1162/089892900562462>.
- James, T.W., Culham, J., Humphrey, G.K., Milner, A.D., Goodale, M.A., 2003. Ventral occipital lesions impair object recognition but not object-directed grasping: an fMRI study. *Brain* 126 (11), 2463–2475. <https://doi.org/10.1093/brain/awg248>.
- Knill, D.C., 2005. Reaching for visual cues to depth: the brain combines depth cues differently for motor control and perception. *J. Vis.* 5 (2), 2. <https://doi.org/10.1167/5.2.2>.
- Kopiske, K.K., Bozzacchi, C., Volcic, R., Domini, F., 2019. Multiple distance cues do not prevent systematic biases in reach to grasp movements. *Psychol. Res.* 83 (1), 147–158. <https://doi.org/10.1007/s00426-018-1101-9>.
- Kopiske, K.K., Bruno, N., Hesse, C., Schenk, T., Franz, V.H., 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. <https://doi.org/10.1016/j.cortex.2016.03.020>.
- Kopiske, K.K., Bruno, N., Hesse, C., Schenk, T., Franz, V.H., 2017. Do visual illusions affect grasping? Considerable progress in a scientific debate. A reply to Whitwell & Goodale, 2016. *Cortex* 88, 210–215. <https://doi.org/10.1016/j.cortex.2016.10.012>.
- Lawrence, M.A., 2016. *ez: Easy Analysis And Visualization Of Factorial Experiments* (4.4-0 [Computer software]). <https://cran.r-project.org/web/packages/ez/index.html>.
- Marotta, J.J., Behrmann, M., Goodale, M.A., 1997. The removal of binocular cues disrupts the calibration of grasping in patients with visual form agnosia. *Exp. Brain Res.* 116 (1), 113–121. <https://doi.org/10.1007/PL00005731>.
- Marotta, J.J., Goodale, M.A., 2001. The role of familiar size in the control of grasping. *J. Cognit. Neurosci.* 13 (1), 8–17. <https://doi.org/10.1162/089892901564135>.
- McCready, D., 1985. On size, distance, and visual angle perception. *Percept. Psychophys.* 37 (4), 323–334. <https://doi.org/10.3758/BF03211355>.
- McIntosh, R.D., Lashley, G., 2008. Matching boxes: familiar size influences action programming. *Neuropsychologia* 46 (9), 2441–2444. <https://doi.org/10.1016/j.neuropsychologia.2008.03.003>.
- Melmoth, D.R., Grant, S., 2006. Advantages of binocular vision for the control of reaching and grasping. *Exp. Brain Res.* 171 (3), 371–388. <https://doi.org/10.1007/s00221-005-0273-x>.
- Milner, A.D., Goodale, M.A., 1995. *The Visual Brain in Action*. Oxford University Press pp. xvii, 248.
- Milner, A.D., Heywood, C.A., 1989. A disorder of lightness discrimination in a case of visual form agnosia. *Cortex* 25 (3), 489–494. [https://doi.org/10.1016/S0010-9452\(89\)80062-0](https://doi.org/10.1016/S0010-9452(89)80062-0).
- Milner, A.D., Perrett, D.I., Johnston, R.S., Benson, P.J., Jordan, T.R., Heeley, D.W., Bettucci, D., Mortara, F., Mutani, R., Terazzi, E., 1991. Perception and action in 'visual form agnosia'. *Brain* 114 (1), 405–428.
- Mon-Williams, M., Bull, R., 2000. The Judd illusion: evidence for two visual streams or two experimental conditions? *Exp. Brain Res.* 130 (2), 273–276. <https://doi.org/10.1007/s002219900258>.
- Mon-Williams, M., McIntosh, R.D., Milner, D.A., 2001a. Vertical gaze angle as a distance cue for programming reaching: insights from visual form agnosia II (of III). *Exp. Brain Res.* 139 (2), 137–142. <https://doi.org/10.1007/s002210000658>.
- Mon-Williams, M., Tresilian, J.R., McIntosh, R.D., Milner, D.A., 2001b. Monocular and binocular distance cues: insights from visual form agnosia I (of III). *Exp. Brain Res.* 139 (2), 127–136. <https://doi.org/10.1007/s002210000657>.
- Morey, R.D., 2008. Confidence intervals from normalized data: a correction to Cousineau (2005). *Tutorials Quant. Methods Psychol.* 4 (2), 61–64. <https://doi.org/10.20982/tqmp.04.2.p061>.
- Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9 (1), 97–113. [https://doi.org/10.1016/0028-3932\(71\)90067-4](https://doi.org/10.1016/0028-3932(71)90067-4).
- Rossit, S., Harvey, M., Butler, S.H., Szymanek, L., Morand, S., Monaco, S., McIntosh, R.D., 2018. Impaired peripheral reaching and on-line corrections in patient DF: optic ataxia with visual form agnosia. *Cortex* 98, 84–101. <https://doi.org/10.1016/j.cortex.2017.04.004>.
- Servos, P., 2000. Distance estimation in the visual and visuomotor systems. *Exp. Brain Res.* 130 (1), 35–47. <https://doi.org/10.1007/s002210050004>.
- Smeets, J.B., Brenner, E., 2001. Action beyond our grasp. *Trends Cognit. Sci.* 5 (7), 287. [https://doi.org/10.1016/S1364-6613\(00\)01684-3](https://doi.org/10.1016/S1364-6613(00)01684-3).
- Smeets, J.B., Brenner, E., 2006. 10 years of illusions. *J. Exp. Psychol. Hum. Percept. Perform.* 32, 1501–1504. <https://doi.org/10.1037/0096-1523.32.6.1501>.
- Tresilian, J.R., Mon-Williams, M., 2000. Getting the measure of vergence weight in nearness perception. *Exp. Brain Res.* 132 (3), 362–368. <https://doi.org/10.1007/s002210000333>.
- Tresilian, J.R., Mon-Williams, M., Kelly, B.M., 1999. Increasing confidence in vergence as a cue to distance. *Proc. R. Soc. Lond. Ser. B Biol. Sci.* 266 (1414), 39–44. <https://doi.org/10.1098/rspb.1999.0601>.
- Volcic, R., Vishwanath, D., Domini, F., 2014. In: Rogowitz, B.E., Pappas, T.N., de Ridder, H. (Eds.), *Reaching into Pictorial Spaces*, 901413. <https://doi.org/10.1117/12.2045458>.
- Wann, J.P., Mon-Williams, M., McIntosh, R.D., Smyth, M., Milner, D.A., 2001. The role of size and binocular information in guiding reaching: insights from virtual reality and visual form agnosia III (of III). *Exp. Brain Res.* 139 (2), 143–150. <https://doi.org/10.1007/s002210000654>.
- Westwood, D.A., Goodale, M.A., 2003. Perceptual illusion and the real-time control of action. *Spatial Vis.* 16, 243–254. <https://doi.org/10.1163/156856803322467518>.
- Whitwell, R.L., Goodale, M.A., 2017. Real and illusory issues in the illusion debate (Why two things are sometimes better than one): commentary on Kopiske et al. (2016). *Cortex: A J. Devoted Study Nervous Syst. Behav.* 88, 205–209. <https://doi.org/10.1016/j.cortex.2016.06.019>.
- Wickham, H., 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer Verlag.
- Wickham, H., Chang, W., Henry, L., Pedersen, T.L., Takahashi, K., Wilke, C., Woo, K., Yutani, H., Dunnington, D., Posit, & PBC, 2023. *ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics* (3.4.2) [Computer software]. <https://cloud.r-project.org/web/packages/ggplot2/index.html>.