The predictability of a target's motion influences gaze, head and hand movements when trying to intercept it

Cristina de la Malla[1,2,*], Simon K. Rushton[3], Kait Clark[3,4], Jeroen B.J. Smeets[2] and Eli Brenner[2]

[1] Vision and Control of Action (VISCA) Group, Department of Cognition, Development and Psychology of Education, Institut de Neurociències, Universitat de Barcelona (*contact author: c.delamalla@ub.edu)

[2] Department of Human Movement Sciences, Vrije Universiteit Amsterdam

[3] School of Psychology, Cardiff University

[4] Department of Health and Social Sciences, University of the West of England

Abstract

Does the predictability of a target's movement and of the interception location influence how the target is intercepted? In a first experiment, we manipulated the predictability of the interception location. A target moved along a haphazardly curved path, and subjects attempted to tap on it when it entered a hitting zone. The hitting zone was either a large ring surrounding the target's starting position (Ring condition) or a small disk that became visible before the target appeared (Disk condition). The interception location gradually became apparent in the Ring condition, whereas it was immediately apparent in the Disk condition. In the Ring condition subjects pursued the target with their gaze. Their head and hand gradually moved in the direction of the future tap position. In the Disk condition subjects immediately directed their gaze towards the hitting zone by moving both their eyes and heads. They also moved their hands to the future tap position sooner than in the Ring condition. In a second and third experiment we made the target's movement more predictable. Although this made the targets easier to pursue, subjects now shifted their gaze to the hitting zone soon after the target appeared in the Ring condition. In the Disk condition they still usually shifted their gaze to the hitting zone at the beginning of the trial. Together, the experiments show that predictability of the interception location is more important than predictability of target movement in determining how we move to intercept targets.

New and Noteworthy

We show that if people are required to intercept a target at a known location they direct their gaze to the interception point as soon as they can, rather than pursuing the target with their eyes for as long as possible. The predictability of the interception location rather than the predictability of the path to that location largely determines how the eyes, head and hand move.

Introduction

When interacting with objects people normally direct their gaze towards them (Land and Hayhoe, 2001; Johansson et al., 2001; Pelz et al., 2001; Mennie et al., 2007; Smeets, et al., 1996; for reviews see Hayhoe and Ballard, 2005; Land, 2006). When objects move in the environment, people almost automatically track them with their gaze (Lisberger et al., 1987; Dorr et al., 2010), often with a combination of eye and head movements (Orban de Xivry and Lefevre, 2007; Bahill and McDonald, 1983; Brenner and Smeets, 2007; 2009; Mrotek and Soechting, 2007; Soechting and Flanders, 2008). This allows them to keep the object of interest foveated, providing the maximal spatial resolution at the target (Schütz et al., 2009). Other advantages of looking at targets when one needs to interact with them are that it helps predict the target's future trajectory (Spering et al., 2011) leading to more precise interception (Brenner and Smeets, 2011; Fooken et al., 2016), and reduces the effects that irrelevant target features have on the object's apparent motion (Braun et al., 2008; de la Malla, et al., 2018; 2019) leading to more accurate performance (de la Malla et al., 2017).

An important factor that has received little attention in relation to how people interact with moving targets is how the predictability of the target's movement influences action. Most of what is known about intercepting moving objects is based on studying how targets such as balls with highly predictable movement trajectories are intercepted. However, predicting how a target will continue to move is not always so straightforward. Imagine for example that the wind blows away some notes that you were carrying to the other side of a lawn. The notes will be moving haphazardly across the lawn so you will probably try to track them with your gaze while gathering them. However, the notes probably cannot be tracked very smoothly, because inevitable inaccuracy in anticipating a note's future position will lead to tracking errors when this anticipated position is used to overcome the latency that is inherent in gaze control (van den Berg, 1988; Robinson, 1965).

If a target is moving predictably, the observer has the option of predicting where it will be some time in the future and moving their gaze to wait at that location. This would explain the anticipatory gaze shifts that are found when a target moves back and forth (Lisberger et al., 1981; Bahill and McDonald, 1983) or bounces off a hard surface (Land and McLeod, 2000; Diaz et al., 2013). Anticipating where a target will be a considerable time in the future makes it possible to successfully intercept targets even if they are not tracked accurately (Cesqui et al., 2015) or gaze is intentionally diverted from the target (López-Moliner et al., 2016). If a target is moving unpredictably, anticipating where it will be a considerable time in the future is not a reliable option, unless for some reason the future location is known. Here we systematically examine how being confronted with unpredictable target motion influences pursuit and interceptive behaviour, and the extent to which knowing where the target will be at some time in the future influences this.

In a first experiment we measured gaze, head and hand movements as subjects attempted to hit unpredictably moving targets. They were asked to hit the targets when the targets crossed into a hitting zone that was visible from the beginning of the trial. In one condition (the Ring condition) the hitting zone was a large ring so that the exact position at which the target will cross the ring gradually became clearer as time progressed (Graf et al., 2005). In the other condition (the Disk condition) the hitting zone was indicated by a small disk so the exact hitting position was evident from the start. In a second experiment the targets moved at a constant speed on straight paths to the same hitting zones, which made it easier to pursue the targets as well as always making it possible to predict where the targets had to be hit from the moment they started to move. In a last experiment the targets moved on a limited number of (straight) trajectories to make the target's motion even more predictable.

Methods

Subjects

Eight subjects (1 author, 1 male) took part in the first experiment (age range 26-39). Two of the subjects reported being left-handed. Five subjects (1 male, 1 left-handed) took part in both the second and third experiments (age range 27-33). Two of the subjects took part in all three experiments. Except for the author that took part in the first experiment, all subjects were naïve to the purposes of the experiments. All subjects had normal or corrected-to-normal vision. None had evident motor abnormalities. All subjects gave written informed consent. The study was part of a program that was approved by the ethical committee of the Faculty of Behavioural and Movement Sciences at the Vrije Universiteit Amsterdam. The experiments were carried out in accordance with the approved guidelines.

Apparatus

The three experiments were conducted in a normally illuminated room. Subjects stood in front of a large screen (Techplex 150, acrylic rear projection screen; width: 1.25 m; height: 1.00 m; tilted backwards by 30º to make tapping more comfortable) onto which the stimuli were projected (In-Focus DepthQ Stereoscopic Projector; resolution 800 by 600 pixels; screen refresh rate: 120 Hz; Figure 1A). The setup gave subjects a clear view of the stimuli as well as of their arm, hand and finger. Subjects were not restrained in any way and had to intercept the projected targets by tapping on them. An infrared camera (Optotrak 3020, Northern Digital) that was positioned at about shoulder height to the left of the screen measured (at 250 Hz) the position of an infrared marker attached to the nail of the index finger of the subjects' dominant hand.

Subjects were free to move in any way they wanted during the experiments. To measure their head movements, we had subjects use their teeth to hold a biteboard with a dental imprint. The positions of three infrared markers attached to the biteboard were monitored by the Optotrak. The movement of the head was inferred from the movement of the biteboard. The use of personal dental imprints means that the position of the head (and thus of the eyes) relative to the biteboard never changes, so their relative positions only need to be determined once.

Eye movements (rotations) with respect to the head were registered with a head-mounted eye-tracking system (Eyelink II, SR Research) at 500 Hz. Where subjects were looking on the screen was determined by combining the measurements of eye in head orientation from the eye tracking system with the position of the eyes and orientation of the head from the recorded biteboard marker positions.

Calibration

In order to relate our gaze measurements to positions of stimuli on the screen (details described in the next paragraph), we needed to know the spatial coordinates of the images on the screen. We used a pointer consisting of a rod with one tapered end and three infrared markers attached to a surface on the other end to calibrate the screen. This pointer was first calibrated by placing an additional marker at the tip of the tapered end to determine the position of the tip relative to the three markers. The rendering of images on the screen was then calibrated by placing the tip of the pointer at five consecutively indicated image positions on the screen. The coordinates of the image positions were determined from the positions of the three markers attached to the pointer.

The pointer and calibrated screen were used to determine the positions of the eyes relative to the biteboard. The pointer was attached to a tripod and was placed between the subject and the screen. Subjects were asked to look with one eye and move their head until the tip of the pointer was aligned with a small white dot presented on the calibrated screen. The markers of both the biteboard and the pointer were recorded by the Optotrak. Subjects could move their heads however they wanted. Once they considered the tip of the pointer to be aligned with the current dot on the screen, they had to press the button of a mouse that they were holding in their hand. If they had moved less than 1 mm during the last 300 ms before doing so, a new dot appeared at a different position and they had to repeat the procedure. Otherwise they had to press again after making sure that the alignment was still fine. Subjects had to align the tip of the pointer with 20 dots using only the left eye and then with 20 dots using only the right eye. Each time they considered the tip of the pointer and the dot to be aligned with one of their eyes, we converted the coordinates of the tip of the pointer and of the dot on the screen into a line with respect to the markers attached to the biteboard. These lines all pass through the eye, but with each measurement providing a different line with respect to the markers of the biteboard. The position with respect to the biteboard that minimized the sum of the distances to all lines was considered to be the position of the eye. From then on, we could determine the positions of the two eyes from measured positions of the markers on the biteboard.

Next, we calibrated the eye movement recordings. To do so, we presented a dot at the centre of the screen, and asked subjects to move their heads for 30 s while maintaining fixation on the dot. By combining the coordinates of the pupil with respect to the head from the Eyelink data with the position of the dot relative to the head (based on the calibrated screen and the biteboard marker coordinates), we determined the scaling of Eyelink coordinates that minimized the deviations in calculated gaze position throughout this period (for each eye). We verified this calibration by asking subjects to look at the screen and rendering dots at the positions at which we considered the subjects to be looking with their left and right eyes. If the two dots were at about the same place, and subjects reported that the dots were at the positions they were looking, the calibration was considered correct. If not, the calibration was repeated.

The final step in the calibration was to relate the position of the fingertip marker to where the subject perceived his or her finger to be relative to the projected images on the screen. For this, we measured the position of the marker on the fingertip when the subject placed the fingertip at four indicated positions on the screen. This step was performed to correct for the fact that the marker was attached to the nail rather than to the tip of the finger.

We synchronized the Optotrak recordings with the images projected on the screen by flashing a disk in the upper left corner of the screen whenever a new target appeared. A photodiode that was directed towards that part of the screen was used to briefly inactivate an additional Optotrak marker attached to the side of the screen (using custom built hardware with a delay of 1 ms). Detecting this inactivation provided information (to within the 4 ms sampling interval) about when the target appeared relative to the movement data, and allowed us to determine that the average latency with which we could adjust the images to events extracted from the online Optotrak data was 24 ms. All delays were accounted for, both in the analysis and in the feedback provided during the trials. Subjects did not notice that the target continued to move for about 24 ms before feedback about their hitting performance was provided, presumably partly because their own finger occluded the target and partly through backward masking (Breitmeyer and Ogmen, 2000).

Combining all these steps provided synchronized arm, head and gaze information in a common coordinate system. For convenience, we used a coordinate system that was aligned with the screen on which the target was moving, so that the target and gaze could be specified by two coordinates.

Stimulus and procedure

Experiment 1:

The experiment was performed in a single session with two randomly interleaved conditions. Subjects started each trial by placing their index finger at an indicated starting point (Figure 1A). The starting point was a 2 cm diameter red disk that was 35 cm below the screen centre. One of two possible hitting zones appeared at the same time as the starting point. The hitting zone was white and was 4 cm wide. It was either a ring (Ring condition, Figure 1B) or a disk (Disk condition, Figure 1C). After a random period between 0.5 and 0.7 s from when the subject placed his or her index finger on the starting position, the target appeared at the centre of the screen. The target moved along a seemingly unpredictable trajectory. The target was a 2 cm diameter black disk. We chose a target that was smaller than the hitting zones, because this often elicits pursuit of the target for at least part of its trajectory when intercepting predictably moving targets (Brenner and Smeets, 2011; de la Malla et al., 2017).

Subjects had to try to intercept the target by tapping on it when it was within the hitting zone. Taps were detected on-line. A tap was considered to have occurred if the deceleration of the movement orthogonal to the screen was at least 50 m/s^2 while the finger was less than 5 mm above the screen. To avoid inadvertently interpreting motion onset as a tap, we also checked that the finger was moving towards the screen, and that it had been lifted to at least 1 cm off the screen since being placed at the starting position. Whenever they wanted, subjects could rest between trials by not placing their finger at the starting position.

In the Ring condition (Figure 1B), the white ring always appeared at the same place, centred on the screen. The ring had a radius of 25 cm and was 4 cm wide. Consequently, it extended from 23 to 27 cm from the screen centre. Subjects had to hit the target when it was within the ring.

In the Disk condition (Figure 1C), the white disk appeared at one of twenty-four possible positions. The disk had a diameter of 4 cm (the same width as the ring) and its centre was 25 cm from the screen centre. The possible positions of the centres of these hitting zones were separated by 15 degrees. Subjects had to hit the target when it was within the disk. The same target trajectories were presented in the two conditions.

Figure 1 here

Figure 1. Schematic representation of the task and conditions. (A) Subjects started with their index finger at the red dot and had to intercept a moving target (black dot) by tapping on it when it reached the white hitting zone. (B) In the Ring condition, the hitting zone was always the same large white ring. (C) In the Disk condition, it was a small white disk at one of 24 possible positions. The white dashed lines in C indicate the other possible positions. They were not visible during the experiment. The six curves in B and C show the six possible paths that the target could take to one of the 24 hitting zones.

The target always appeared at the centre of the screen and could follow one of six possible trajectories in one of 24 directions. The different trajectories were constructed in polar coordinates using a constant increase in distance from the screen centre, with the polar angle φ changing according to Equation 1:

$$\varphi = D + \left(a + b\sin\left(2\pi\frac{t}{T}\right)\right)\left(\frac{t}{T}\right)^2 \qquad \text{(Equation 1)}$$

where the *D* is one of the 24 directions to the hitting zone (equally spaced), *t* is time to reach the centre of the hitting zone and *T* is the movement time of the target (1.2 s). There were six combinations of values of *a* and *b*: [-2π/3, π/2], [π/3, - π/2], [2π/3, - π/2], [-π/3, π/2], [π/2, π/2], [-π/2, - π/2]. The six possible target trajectories are shown in Figures 1B and 1C. All six trajectories crossed the centres of the hitting zones after 1.2 s. In trials of the Ring condition, subjects only gradually realised where the target would pass through the large hitting zone as the trial progressed, with the target approaching the ring along a curvy path. In trials of the Disk condition, subjects knew that the target was going to pass through the small hitting zone even before the target appeared.

Feedback was provided after each attempt to hit the target. A target was considered to have been hit if the tip of the finger (as calibrated) was within the outline of the target. If subjects hit the target, the target stopped moving and remained at the position at which it had been hit for 500 ms. If the tip of the finger was also within the hitting zone a sound indicated that the target was hit. If subjects missed the target, the target was deflected away from the finger at 1 m/s, remaining visible for 500 ms. All the trajectories and conditions were presented in random order in a single session. In total, there were 288 trials per subject: 2 conditions, 24 directions to the hitting zone, 6 trajectories for each direction. It took about 25 minutes to complete the experiment.

Experiment 2:

The second experiment was identical to the first, except that the targets followed a straight trajectory towards either the Ring or the Disk (*a* and *b* in Equation 1 were both zero). The purpose of this experiment was to determine which differences between how subjects intercepted the targets in the Disk and Ring conditions of Experiment 1 were due to the Disk revealing where the target could be hit even before the target appeared, and to determine which aspects of how subjects intercepted the targets in Experiment 1 were specific to targets that move unpredictably. In total there were 192 trials per subject: 2 conditions, 24 directions to the hitting zone, and 4 repetitions for each hitting zone. It took about 15 minutes to complete the experiment.

Experiment 3:

The third experiment was identical to the second, except that targets only moved in four of the 24 possible directions (0º, 90º, 180º or 270º). This made it even easier to judge where the target would cross the Ring. In total there were 40 trials per subject: 2 conditions, 4 directions to the hitting zone, and 5 repetitions for each hitting zone. It took about 8 minutes to complete the experiment.

Data analysis

All analyses were performed with custom written programs using RStudio (RStudio Team, 2018). In Experiment 1 we excluded 76 trials (3.3%) in which subjects clearly did not follow the instruction. These were 52 trials in which no tap was detected, 12 trials in which the distance between where subjects tapped (the tap position) and where the target was at the moment of the tap was larger than 20 cm, and 12 trials in which the distance between the tap position and the position at which the target's path crossed the centre of the hitting zone was larger than 20 cm. No trials were excluded due to missing data. In Experiments 2 and 3 we excluded 6 (0.5%) and 2 (0.8%) trials, respectively, all because subjects did not tap on the screen within 1.5 seconds.

The next step in our analysis was to align the Optotrak and Eyelink data with the presentation of the images on the screen using the timing signal from the photodiode. Since the data acquisition itself was not synchronised with the image projection, and was at different frequencies for the Optotrak and Eyelink, the first step in our analysis was to align the signals in time using linear interpolation to obtain a target position (on the screen), eye orientations (with respect to the head), eye positions (in space), head orientation (in three dimensions with respect to the world) and hand position (position of the finger with respect to the screen) at each moment from when the targets appeared until the moment of the tap. We refer to the average position of the two eyes as the head position, so the reported changes in head position include influences of both displacements and rotations of the head. We combined the temporally aligned positions of the eyes in space with the orientations of the eyes with respect to the head and the orientation of the head in space to calculate the line of sight for each eye.

We determined where subjects were looking on the screen (gaze) by averaging the estimates of where the lines of sight of the two eyes intersected the screen (except for 22 trials of Experiment 1 in which only one of the eyes was measured correctly, probably due to some light reflecting on glasses; for those trials we only used the estimates of one eye). We calculated the instantaneous speed and acceleration of gaze, head and hand movements by using finite difference approximations. We divided the change in position between 10 ms before and 10 ms after the moment in question by the 20 ms time difference between them. We calculated the gaze acceleration by dividing the difference between the gaze speeds 10 ms after and 10 ms before the moment in question by the 20 ms time difference between them. When calculating the speed of the head and of the hand we only considered the motion component parallel to the screen, because we wanted to determine the peak in the speed at which the hand moved towards the vicinity of the target. Including the motion component orthogonal to the screen would include the final tapping movement, which was often very fast so that the peak velocity would often be just before the tap. We also report the component parallel to the screen when reporting head and hand positions and distances moved.

To evaluate whether gaze, the head and the hand were following the target we examined how the distance from the interception point decreased during each trial. Given that the hand's starting position is below all possible target locations, the hand's initial distance differed considerably between hitting zones at the top and bottom of the screen (Figure 1B and 1C). To prevent changes in the hand's distance from the upper target locations from overshadowing those from the lower target locations when averaging across target locations, we averaged normalized distances. We obtained the latter by dividing the distance from the hand position to the tap position at each moment of time by the initial distance of the hand from the tap position. Unlike for the finger, there was no specified starting position for the head and gaze. To obtain somewhat comparable normalised distances for the head and gaze we assumed that subjects started each trial with their head approximately in front of the position at which the targets appeared and with their gaze directed at where the targets appeared. We divided the distances of the head and gaze from the tap position by the distance from the position at which the target appeared to where it was tapped. The latter distance was always approximately 25 cm, but not precisely so on each trial because the tap was not always exactly at the centre of the hitting zone. With these assumptions the initial normalized distance will be one unless subjects respond before the target appears. Gaze and the head are not required to end at any particular place, so they do not have to end at zero as the hand does, although we do expect gaze to end near the tap irrespective of whether subjects pursue the target or fixate where they tap. To compare how subjects moved in the different conditions we plotted the normalised distances of gaze, head and hand across time for each experiment and condition. To be able to evaluate the consistency of any visible differences the plots include the standard error across subjects at each moment.

The number of saccades per trial and whether the saccades were towards the target or towards the interception location provided additional measures of gaze behaviour. Determining the number of saccades towards the target can help evaluate to what extent differences in gaze behaviour result from being unable to predict how the target will move. We identified saccades using a similar method to that described in de la Malla et al. (2017). We considered the eyes to be making a saccade if the gaze speed remained above a threshold of three times the target's speed for more than 10 ms. Since the target did not move at a constant speed, this threshold differs slightly at different moments. Once we had detected a saccade we determined when it ended by first localizing the maximal deceleration of gaze and then finding the moment at which gaze no longer decelerated by more than 5 cm/s^2. We used the gaze position at the end of the saccade to distinguish between saccades that contribute to keeping gaze on the target and ones that direct gaze towards the hitting zone. If a saccade ended closer to the centre of the target than to the centre of the disk or to the midline of the ring (both at 25 cm from the screen centre), we considered it to be a saccade that served to keep gaze on the target. Otherwise, we considered it to be a saccade towards the hitting zone. We do not expect subjects to be able to pursue an unpredictably moving target very precisely, so we expect them to make more saccades when tracking the target in the Ring condition in which the precise position at which one would be able to hit the target was not known in advance. We tested whether this is the case using a one-sided paired t-test.

We also compared hand movements in the Disk and Ring conditions on a number of measures using one-sided t-tests on subject means. We compared (i) the proportion of targets hit, (ii) timing precision for hitting the target, (iii) peak speed of movement of the finger, (iv) time to peak speed (how rapidly subjects responded), and (v) the directness of the movement (the distance travelled: the sum of displacements across consecutive measurements until the time of the tap). In Experiment 1, knowing in advance where the finger's movement will need to end, as one did in the Disk condition, makes it possible to plan the movement as soon as the target appears, rather than having to track the target's meandering trajectory. We predicted that this might lead to (i) more targets being hit; (ii) timing being more precise; (iii) the mean peak speed being higher and (iv) occurring earlier; and (v) the movements being more direct in the Disk condition. As the subjects were the same in both conditions we used paired t-tests.

In Experiments 2 and 3 the position at which the finger's movement will end is still known earlier in the Disk condition, but the straight trajectories allow one to infer where the target is to be hit as soon as it starts moving (i.e. immediately after it appears) in the Ring condition. Thus, although the direction of any differences between the conditions would be expected to be the same as for Experiment 1, we expect all the differences between conditions to be smaller. We expect the behaviour of the finger in both conditions to be similar to that in the Disk condition of Experiment 1. The peak speed might still occur slightly later in the Ring condition because the interception point is only revealed by the target's motion, rather than being revealed even before the target appears (by the position of the Disk). Since the target trajectories were simpler in Experiment 2 than in Experiment 1, and were even more predictable in Experiment 3, we expected performance to become better in consecutive experiments (more targets hit and better timing) and the movements to possibly also become faster and occur earlier. We used one-sided paired tests when comparing Experiments 2 and 3, but tests were not paired when comparing those experiments with Experiment 1 because the subjects were not all the same.

Results

Experiment 1: unpredictable trajectories

The subjects' goal was to tap on the screen in such a manner that their fingertip was within both the target and the hitting zone at the time of the tap. Subjects successfully hit more targets in the Disk condition than in the Ring condition (Table 1). On average subjects tapped at the correct place (25 cm from the screen centre) and time (1.2 s after the target appeared) in both conditions, but the variability (standard deviation) in the time at which individual subjects tapped was smaller in the Disk condition than in the Ring condition (Table 2). Thus, their timing was more precise in the Disk condition.

Experiment	Disk	Ring	One-sided paired t-tests
1	72.2	57.4	t_7=3.36, p=0.006
2	83.8	85.4	t_4=2.02, p=0.94
3	86.0	94.0	t_4=1.73, p=0.92

Table 1. Percentage of targets hit. A target is considered to have been hit if the finger, as calibrated, was within the bounds of both the target and the hitting zone at the time of the tap. Performance only differed significantly between the Disk and Ring condition in Experiment 1. Performance in Experiments 2 and 3 differed significantly from that in Experiment 1 (Experiment 2, Disk: $t_{4,7}$=2.3, p=0.03; Ring: $t_{4,7}$=5.12, p=0.0003; Experiment 3, Disk: $t_{4,7}$=2.34,

p=0.03; Ring: $t_{4,7}$=7.02, p<0.001) but not from each other (Disk: $t_{4,4}$=0.33, p=0.38; Ring: $t_{4,4}$=1.46, p=0.09).

Experiment	Disk	Ring	One-sided paired t-tests
1	36	48	t_7=2.72, p=0.015
2	33	44	t_4=1.48, p=0.11
3	26	28	t_4=0.71, p=0.26

Table 2. Variability in the timing of the hits (standard deviation in ms). Performance only differed significantly between the Disk and Ring condition in Experiment 1. Performance in Experiment 3 differed significantly from that in Experiment 1 (Disk: $t_{7,4}$=1.92, p=0.04; Ring: $t_{7,4}$=3.05, p=0.008), but the other differences between experiments were not significant.

Figure 2 shows two example trials from a representative subject for Experiment 1. There are clear differences between how the subject moved to intercept the targets in the two conditions. When the position at which to hit the target was not known in advance (Ring condition, left panel), the gaze (blue) more or less followed the target's movement (grey) until the moment of the tap. It did so in quite a jerky manner, presumably because the eyes made many saccades to correct for errors in predicting how the target would proceed. Therefore, these saccades are not really to catch up with the target position, but anticipating where the target will be next and thus often anticipating incorrect positions because the target moves unpredictably. When the position at which to hit the target was known in advance (Disk condition, right panel), gaze was immediately directed towards this position: the blue curve starts and remains close to the disk rather than following the target. Both the head and the hand also moved sooner in the direction of the hitting zone in the Disk condition than in the Ring condition: a smaller part of the trajectory is clearly red or green. One can also see that the hand moves along a straighter path in the Disk than in the Ring condition.

Figure 2 here

Figure 2. Example of gaze, head and hand movements on single trials for a representative subject in the two conditions of Experiment 1. Data of two trials with the same target trajectory from the moment the target appeared until the time of the tap. The colours of the curves change with the remaining time to tap: from black to either grey, blue, red or green (for the target, gaze, head and hand, respectively).

The differences between the two example trials of Figure 2 are characteristic of the differences between the two conditions for this subject (Figure 3) as well as for other subjects. Due to the time period between the subject placing his or her finger at the starting position and the target appearing, gaze was usually no longer directed at the starting position by the time the target appeared. In the Ring condition gaze was usually directed at the centre of the screen, where the targets appeared, and then tracked the target. In the Disk condition gaze was often already directed towards the hitting zone by the time the target appeared, as is the case in the trial shown in Figure 2 (the hitting zone was visible well before the target appeared). On some other trials of this condition gaze was directed at the centre of the screen until the target appeared, but when the target appeared a saccade was made to the disk rather than gaze tracking the target.

Figure 3 here

Figure 3. Gaze, head and hand movements of all trials of the same representative subject in Experiment 1 shown in Figure 2. Colours change from black to blue (gaze), red (head) and green (hand) across time from when the target appears to the moment of the tap (as in Figure 2).

To illustrate the time-course of the gaze movements we plotted the average normalized distance of gaze from the tap position as a function of the time to hit the target (Figure 4A). There is a clear difference between the Ring and the Disk condition. In the Ring condition the distance between the gaze and the tap position decreases constantly across time at a similar pace as the target approaches the tap position (thin black dotted line). This is consistent with subjects trying to track the target with their eyes. As could be expected on the basis of Figures 2 and 3, on average subjects were already looking closer to the hitting zone when the target appeared in the Disk condition (dashed blue curve lower than solid blue curve from the start in Figure 4A). Consequently, the distance between gaze and the tap position changed much less across time. The average normalized distance between gaze and tap position only decreased to about 0.2 in both conditions (Figure 4A). This corresponds to a distance of about 5 cm at the moment of the tap. This could mean that gaze was not directed at the position that was tapped, but it could also arise from measurement errors (see Discussion). We never required subjects to fixate a specific position during the experiment, to avoid biasing where they looked, so we did not try to correct for systematic shifts (such as the overall shift to the upper right in the left panels of Figure 3), for instance by assuming that on average subjects were looking at the disks when they hit the targets, because we cannot be sure that this was the case. Importantly, the differences that we find between the two conditions cannot be due to eye-tracker shifts because the trials of the two conditions were interleaved.

A closer look at the tracking strategy (inset in Figure 4A) reveals that subjects made more than twice as many saccades in the Ring than in the Disk condition (t_7=8.9, p<0.001). In accordance with subjects trying to keep their eyes on the unpredictably moving target in the Ring condition, we see that the increase in the number of saccades is caused by an increase in the number of saccades directed to the target (t_7=11.4, p<0.001).

The movements of the head and the hand also differed between the two conditions (Figure 4B and 4C). The head was closer to the hitting zone in the Disk condition than in the Ring condition from the moment the target appeared (dashed red curve lower than solid red curve). At least part of this difference in head position is probably related to the above-mentioned difference in gaze: one can orient one's head towards the position at which the target is to be hit before the target appears in the Disk condition, but not in the Ring condition. The hand was not allowed to start moving before the target appeared, so it always started at a normalized distance of 1. It took some time for the hand to start moving when the target appeared. Once the hand did start moving it approached the tap position sooner in the Disk condition than in the Ring condition.

Figure 4 here

Figure 4. Analysis of the average gaze, head and hand movements of all eight subjects in Experiment 1 (A-C) and of all five subjects in Experiments 2 (D-F) and 3 (G-I). Normalized distance to the tap position as a function of the time until the target is hit for the gaze, head and hand. The lines (continuous for the Ring condition, dashed for the Disk condition) and shaded areas are the means and standard errors of the subjects' individual mean values. A normalized distance of zero corresponds to being at the tap position. A normalized distance of one corresponds to being where the target appeared for the gaze and the head, and corresponds to being at the finger's starting position for the hand. In the gaze panels, we also show the mean normalised distance of the target from the tap position (black dotted curve). The inset in A shows the number of saccades per trial in Experiment 1, split by whether saccades ended closer to the target (black bars) or closer to the tap position (white bars). Error bars are standard errors across the subjects' mean numbers of saccades.

In accordance with the impression one gets from the gaze panels of figures 2, 3 and 4A, the distance travelled by gaze while the target was present was longer in the Ring condition than in the Disk condition (53 ± 4 cm versus 32 ± 3 cm; mean ± standard error across subjects; t_7=6.3, p=0.0002). This is consistent with subjects trying to pursue the target in the Ring condition but not in the Disk condition.

Unlike gaze, the head does not travel significantly less in the Disk condition (t_7=1.11, p=0.15): it travels an average of 8.2 ± 0.9 cm. The peak speed of the head was not significantly higher (t_7=-6.2, p=0.99) in the Disk (18 ± 2 cm/s) than in the Ring condition (21 ± 2 cm/s). However, the head did reach the peak speed earlier in the Disk condition (t_7=4.86, p=0.0009): the peak speed occurred after 0.71 ± 0.05 s in the Disk condition and after 0.89 ± 0.03 s in the Ring condition. The hand trajectories were straighter (shorter) in the Disk condition (t_7=6.20, p=0.0002): the mean distance travelled by the hand was 43.4 ± 0.3 cm in the Disk condition and 51.6 ± 1.4 cm in the Ring condition. Despite the shorter distance, the peak speed of the hand was higher in the Disk condition: it was 122 ± 3 cm/s in the Disk condition and 112 ± 5 cm/s in the Ring condition (t_7=2.5, p=0.02). The peak speed of the hand also occurred earlier (t_7=3.44, p=0.005) in the Disk condition (0.52 ± 0.03 s) than in the Ring condition (0.65 ± 0.05 s). These findings support the idea that knowing in advance where they will hit the target allows subjects to move sooner, more directly and faster.

The location at which subjects will be able to hit the target only gradually became apparent in the Ring condition. When the ring appeared and the target started to move subjects could have followed the strategy of moving their hand directly to some position within the ring and adjust their movement along the ring as the target approached it. Figure 5 shows that they did not do this. They seldom moved along the ring (left panels). Furthermore, when the target was to be hit at the closest position to the hand's starting position subjects moved their hand towards the target, within the ring, before moving it back down to the ring as the target approached the ring (bottom left panel). In the Disk condition (right panels), subjects moved their hand to the hitting zone along a much straighter path, only moving beyond the hitting zone when the hitting zone was near the hand's starting position (bottom right panel) a single time.

Figure 5 here

Figure 5. Hand movements of all trials of all eight subjects for the furthest (top panels) and the nearest (bottom panels) hitting zones in Experiment 1. All trajectories start at the hand's starting point near the bottom of the panel. Colour changes from black to green across time as in Figures 2 and 3.

Experiment 2: predictable trajectories

The first experiment showed a marked difference in movement strategies between the two conditions. We attribute the difference to the predictability of the interception location. In the second experiment we kept the conditions the same, but the interception location was predictable from just after the targets appeared and started moving because the targets moved at a constant velocity along straight paths. Subjects managed to hit more targets when the targets moved more predictably, and there was no longer a significant difference between the Disk and Ring conditions (Table 1). The variability in the timing of the taps was also no longer significantly larger in the Ring than in the Disk condition (Table 2). The differences in performance between the two conditions were therefore not just due to the interception location being known before the target appeared in the Disk condition.

The tap accuracy and timing were similar in the Ring and Disk conditions (Table 1 and 2), but there were small differences between the two conditions. On average, gaze travelled less in the Disk (33.2±3 cm) than in the Ring (48.6±3 cm) condition. The difference was not consistent across subjects (t_4=1.7, p=0.08) and is easily explained by the interception location being known before the target appears in the Disk condition, while it only becomes apparent from the motion of the target in the Ring condition (it is evident as soon as the target moves because the target always moves along a straight path). Gaze was often already at the interception location by the time the target appeared in the Disk condition, whereas it could only move there after the target started moving in the Ring condition (Figure 4D). That the time at which the interception location is known is important is also evident from the difference between gaze in the Ring conditions of Experiments 1 and 2: gaze reaches the vicinity of the tap position earlier in Experiment 2 (compare Figure 4A and 4D). In Experiment 1 it took an average of 1.04 s for gaze to be within 10% of the final normalized distance to the tap position. In Experiment 2 it only took 0.79 s ($t_{4,7}$=3.84, p=0.003). This difference is undoubtedly the result of the predictable target motion revealing the interception location. However, the difference in performance between the Disk conditions of Experiments 1 and 2 (Table 1) suggests that there is also a direct effect of the predictability of target motion.

The difference in head position between the two conditions is smaller in Experiment 2 (Figure 4E) than in Experiment 1 (Figure 4B) from the moment that the target appears, although there is no difference between the experiments in terms of the available information at that moment. The difference is consistent with the difference in gaze at the moment the target appears also being smaller in Experiment 2 than in Experiment 1. Thus, the differences in head movement between the conditions are probably due to differences in gaze. The differences in gaze between the two experiments might be the result of the initial target trajectory always being informative in Experiment 2.

The hand movements were extremely similar in the Disk and Ring conditions of Experiment 2 (Figure 4F), with the hand traveling 42.1 cm in both cases. The small difference in movement onset is consistent with the hitting position becoming apparent slightly later for the Ring than for the Disk condition. The hand did not appear to move as quickly to the hitting zone in this experiment as it had in the Disk condition of Experiment 1. The peak speed was 110 ± 8 for the Disk condition and 107 ± 7 cm/s for the Ring condition (t_4=1.92, p=0.06), which are values close to the peak velocity of the hand for the Ring condition in Experiment 1 (113 cm/s). The peak speed occurred after 0.6 s, for both conditions, which is midway between the values that we found for the Disk and Ring conditions in Experiment 1. The results of this experiment support the idea that knowing that the target's initial movement will be informative of the interception location on all trials influences how subjects approach the task.

Experiment 3: predictable trajectories and tap positions

In Experiment 2 we found that the predictability of the hitting position influences interceptive actions. In Experiment 3 we investigated whether the degree of predictability was important. To do so we made it even easier to predict where the targets will be hit in the Ring condition. We repeated the second experiment but with only four of the 24 hitting zones (values of D in Equation 1 of 0, 90, 180 and 270º). The percentage of targets that were hit was highest in this experiment, though not significantly higher than in Experiment 2 (Table 1). The percentage of targets that were hit was not lower for the Ring condition (94%) than for the Disk condition (86%). The standard deviation in timing the hits was lowest in this experiment, though not significantly lower than in Experiment 2 (Table 2).

The time course of the movements in Experiment 3 was very similar to that in Experiment 2 (Figure 4G-I). Again, the main difference between the Ring and Disk conditions is that gaze was directed to the hitting zone before the target appeared in the Disk condition, whereas it obviously could not be in the Ring condition. Movements of the head hardly contributed to this difference, and the arm movements were not affected by knowing where the target would be hit in advance. Even the tiny delay in hand movement onset seems to have vanished, probably because it is easier to tell in which of the four directions the target is moving, than to distinguish between 24 directions. The peak speed of the hand (102 ± 6 cm/s) and the time at which it occurred (0.59 s after appearing, when the target was almost half way to the interception location) were similar to the values in Experiment 2 ($t_{4,4}$=1.51, p=0.90 and $t_{4,4}$=-0.06, p=0.52, for the peak speed and the time at which it occurred, respectively). The fact that, again, performance was slightly different from that of the Disk condition of Experiment 1, supports the notion that beside the target's path being relevant because it influences when one knows where the target is to be hit, it is presumably also easier to determine when the target will arrive at the position at which it is to be hit when the target is moving more predictably.

Discussion

What options does one have to successfully intercept a target that moves unpredictably? When one tries to catch a note that is blown away by the wind, the only option is to track it with one's gaze as one adjusts one's arm movement so that the hand reaches the note. When trying to intercept a predictably moving object one could follow the same strategy, but one could also predict where one will be able to intercept the target and immediately direct one's gaze and movement towards that location. We examined how the circumstances influence what people do and how the choice influences their performance.

The results of Experiment 1 suggest that even if the target moves in an unpredictable manner, so that it is essential to constantly monitor its motion, pursuing the target with one's gaze is not always the best strategy for guiding the hit. In order to pursue a target smoothly with no delay one must be able to anticipate how it will continue moving (Lisberger et al., 1981; Kowler and Steinman, 1979). If a target's trajectory is completely unpredictable (Ring condition of Experiment 1), gaze must track the target (Figures 2, 3 and 4A), even if this means that pursuit of the target will be interspersed with saccades (inset of Figure 4A). Such saccades will temporarily limit what one perceives (Zuber and Stark, 1966; Bridgeman et al., 1975; Burr et al., 1999; Castet and Masson, 2000; Maij et al., 2012; Ross et al., 2001) and give rise to errors in judging the target's position and motion (Matin and Pearce, 1965; Mateeff, 1978; Honda, 1989; Morrone et al., 1997; Schlag and Schlag-Rey, 2002; Maij et al., 2009; 2011; Matziridi et al., 2015, Goettker et al., 2018; 2019). If one knows where one will be able to hit the target in advance (imagine waiting for a fly to settle on a particular breadcrumb that it is clearly circling around; Disk condition), it appears to be better to quickly direct one's gaze towards that position and track its approach with peripheral vision (Figure 4A) because doing so appears to improve performance (Tables 1 and 2). That performance is better when fixating in such circumstances need not be due to the disadvantages associated with having to perform saccades to keep the target in central vision outweighing the disadvantages of relying on peripheral vision to track the target's motion, because being able to anticipate where one will be able to hit the target may be advantageous for other reasons. However, the fact that subjects did not consistently pursue the target in the Disk condition trials although they did pursue the target on the interleaved Ring condition trials suggests that fixating is advantageous under these circumstances.

As mentioned in the results, it seems surprising that subjects appeared not to direct their gaze exactly at the tap position at the moment of the tap (Figures 4A, 4D and 4G). In order to not bias their gaze behaviour we did not give them instructions about where to look at any time, except during the eye movement calibration during which subjects fixated a static dot (see Methods). The measured precision during calibration was about 0.7 degrees horizontally and 1.2 degrees vertically for each eye (root mean square deviation). However, recorded eye orientations are known to drift, mainly due to headband slippage, giving rise to systematic shifts. Therefore, we cannot determine with certainty which part of the distance between gaze and tap position at the moment of the tap is due to measurement errors and which to the fact that subjects may not have directed their gaze precisely at the tap position when tapping.

Our results are largely in agreement with previous studies on how people interact with unpredictable moving targets (Danion and Flanagan, 2018; Mrotek and Soechting, 2007; Xia and Barnes, 1999). Danion and Flanagan (2018) examined subjects' gaze strategy when tracking a target that moved along an unpredictable trajectory. In one condition their subjects had to track a target with their hand, without instructions about gaze. They found that gaze always also tracked the target. This is consistent with our observation that subjects track unpredictable target motion if they do not know how the target will move. Mrotek and Soechting (2007) examined subjects' gaze strategy in an interception task. In their task, subjects were free to choose when and where to hit the targets. They observed that subjects pursued the target, but also that saccades were suppressed just before the moment of interception. This is consistent with our proposal that making saccades near the time of interception comes at a cost. However, the cost cannot be very high because people do in some circumstances make saccades to where they are required to hit a target before reaching it with the hand (rather than pursuing it smoothly until it is hit) when the target moves predictably (de la Malla et al., 2017).

In both the Disk and Ring condition, the target has to be hit at a specific time and place. This restricts the adjustments that subjects can make when guiding the hand to the target (Brenner and Smeets, 2015). When the target's trajectory is unpredictable, knowing where to hit it in advance might not improve the timing of the tap (Experiment 1; Table 2) through its influence on the eye movements, but by making it easier to judge when to hit the target. The targets moved quite smoothly, so knowing that they will pass a certain position probably helped estimate when that would happen. However, judging when the target will cross the ring is less reliable because a small change in the trajectory, that is constantly curving, can change the position at which the target crosses the ring, and therefore also the time at which it does so at its current speed. The hand must also reach the changed position. The hand followed the target to some extent in the Ring condition of Experiment 1. Subjects did not quickly move their hand to the ring and then adjust its position along the ring (Figure 5), but the hand did not closely track the target either (Figure 2). This may just be due to physical limitations in how the hand can be moved, but subjects may intentionally avoid occluding the target with the hand, or even avoid occluding parts of the screen across which the target may move during its meanderings.

The predictability of the targets' trajectories also influenced head movements to some extent. Previous studies have reported that head movements contribute substantially to keeping moving targets in central vision when interacting with them (Bahill and LaRitz, 1984; Mann et al., 2013; Fogt and Zimmermann, 2014; Fogt and Persson, 2017). Most of those studies involved sports such as baseball or cricket, in which the ball's angular displacement near the time of the hit is so large that it is impossible to track the ball by moving the eyes only. In our study the distance between where the targets appeared and the hitting zone was only 25 cm (about 25 deg, depending on where the subject chose to stand), so large head movements were not necessary to keep track of the moving targets. However, head movements did contribute to the changes in gaze (Figure 4B, E and H). The contribution was modest, but the differences between the conditions were more or less consistent with the differences in gaze, although gaze changed more and more abruptly.

In summary, for the conditions used in the current study the preferred strategy was to quickly direct one's gaze at the position at which the target will be hit. Gaze only tracked the target when the interception point was initially unknown (Ring condition) and could not immediately be inferred from the target's motion (Experiment 1). In that case performance was relatively poor, presumably because it was impossible to keep one's eyes on the target and because the hand movement was constantly adjusted as a result of it being difficult to anticipate when and where the target could be hit. The experiments suggest that how people approach an interception task is mainly determined by how reliably they can predict the interception location rather than by how reliably they can predict the target's movement to that location, at least when an interception zone is specified.

References

Bahill, A.T. and McDonald, J.D. (1983) Smooth pursuit eye movements in response to predictable target motions. *Vision Res*, **23** (12), 1573-1583

Bahill, A.T. and LaRitz, T. (1984) Why can't batters keep their eyes on the ball. *American Scientist*, **72**(3), 249-253.

Braun, D. I., Mennie, N., Rasche, C., Schutz, A. C., Hawken, M. J., & Gegenfurtner, K. R. (2008). Smooth pursuit eye movements to isoluminant targets. *J Neurophysiol*, **100** (3), 1287-1300.

Breitmeyer, B.G. and Ogmen, H. (2000) Recent models and findings in visual backward masking: a comparison, review, and update. *Perception and Psychophysics*, **62** (8), 1572-1595.

Brenner, E. and Smeets, J.B.J. (2007) Flexibility in intercepting moving objects. *J Vision*, **7**(5):14, 1-17

Brenner, E. and Smeets, J.B.J. (2009) Sources of variability in interceptive movements. *Exp Brain Res*, **195** (1), 117-133

Brenner, E. and Smeets, J.B.J. (2011) Continuous visual control of interception. *Hum Movement Sci*, **30** (3), 475-494

Brenner, E. and Smeets, J.B.J. (2015a) How people achieve their amazing temporal precision in interception. *J Vision*, **15**(3):8, 1-21

Bridgeman, B., Hendry, D., and Stark, L. (1975) Failure to detect displacement of visual world during saccadic eye movements. *Vis. Res*. **15**, 719–722

Burr, D.C., Morgan, M.J. and Morrone, M.C. (1999) Saccadic suppression precedes visual motion analysis. *Curr. Biol.* **9**, 1207–1209

Castet, E. and Masson, G.S. (2000) Motion perception during saccadic eye movements. *Nat. Neurosci*. **3**, 177–183

Cesqui, B., Mezzetti, M., Lacquaniti, F., and d'Avella, A. (2015). Gaze behavior in one-handed catching and its relation with interceptive performance: what the eyes can't tell. *PLoS One*, **10**(3), e0119445.

Danion, F. R., and Flanagan, J. R. (2018). Different gaze strategies during eye versus hand tracking of a moving target. *Scientific Reports*, **8**(1), 10059.

de la Malla, C., Smeets, J.B.J. and Brenner, E. (2017) Potential systematic interception errors are avoided when tracking the target with one's eyes. *Scientific Reports*, **7** (1): 10793

de la Malla, C., Smeets, J.B.J. and Brenner, E. (2018) Errors in interception can be predicted from errors in perception. *Cortex* **98**, 49-59.

de la Malla, C., Brenner, E., de Haan, E. H., and Smeets, J. B. J. (2019). A visual illusion that influences perception and action through the dorsal pathway. Communications biology, 2(1), 38.

Diaz, G., Cooper, J., Rothkopf, C. and Hayhoe, M. (2013) Saccades to future ball location reveal memory-based prediction in a virtual-reality interception task. *J Vision*, **13** (20), 1-14

Dorr, M., Martinetz, T., Gegenfurtner, K. R., and Barth, E. (2010) Variability of eye movements when viewing dynamic natural scenes. *J Vision*, **10** (10), 1-17.

Fogt, N. F., and Zimmerman, A. B. (2014). A method to monitor eye and head tracking movements in college baseball players. *Optometry and Vision Science*, **91**(2), 200-211.

Fogt, N., and Persson, T. W. (2017). A pilot study of horizontal head and eye rotations in baseball batting. *Optometry and vision science: official publication of the American Academy of Optometry*, **94**(8), 789-796.

Fooken, J., Yeo, S. H., Pai, D. K., & Spering, M. (2016). Eye movement accuracy determines natural interception strategies. *J Vision*, **16** (14), 1-15.

Goettker, A., Braun, D. I., Schütz, A. C., and Gegenfurtner, K. R. (2018). Execution of saccadic eye movements affects speed perception. *Proc Natl Acad Sci USA*, **115** (9), 2240-2245.

Goettker, A., Brenner, E., Gegenfurtner, K. R., and de la Malla, C. (2019). Corrective saccades influence velocity judgements and interception. *Scientific Reports* 9(1): 5395

Graf, E.W., Warren, P.A. and Maloney, L.T. (2005) Explicit estimation of visual uncertainty in human motion processing. *Vision Res*, **45**(24), 3050-3059.

Hayhoe, M. and Ballard, D. (2005) Eye movements in natural behaviour. *Trends in Cognitive Sciences*, **9** (4), 188-194

Honda H. (1989) Perceptual localization of visual stimuli flashed during saccades. *Percept Psychophys*. **45**,162–174

Johansson, R.S., Westling, G., Bäckström, A. and Flanagan, J.R. (2001) Eye-hand coordination in object manipulation. *J Neurosci*, **21** (17), 6917-6932

Kowler, E. and Steinman, R. M. (1979) The effect of expectations on slow oculomotor control – I. Periodic target steps. *Vision Res*, **19**, 619-632.

Land, M.F. (2006) Eye movements and the control of actions in everyday life. *Progress in Retinal and Eye Research*, **25** (3), 296-324

Land, M.F. and Hayhoe, M. (2001) In what ways do eye movements contribute to everyday activities? *Vision Res*, **41**, 3559-3565

Land, M.F. and McLeod, P. (2000) From eye movements to actions: how batsmen hit the ball. *Nat Neurosci*, **3** (12), 1340-1345

Lisberger, S. G., Evinger, C., Johanson, G. W. and Fuchs, A. F. (1981) Relationship between acceleration and retinal image velocity during foveal smooth pursuit in man and monkey. *J Neurophysiol*, **46**, 229-249.

Lisberger, S. G., Morris, E. J., & Tychsen, L. (1987). Visual motion processing and sensory-motor integration for smooth pursuit eye movements. *Annual Rev Neurosci*, **10**(1), 97-129.

López-Moliner, J., and Brenner, E. (2016). Flexible timing of eye movements when catching a ball. *J Vis*, **16**(5), 13-13.

Maij F., Brenner E. and Smeets J.B.J. (2009) Temporal information can influence spatial localization. *J Neurophysiol*. **102**, 490–495

Maij, F., Brenner, E. and Smeets, J.B.J. (2011) Temporal uncertainty separates flashes from their background during saccades. *J Neurosci* **31**, 3708-3711.

Maij, F., Matziridi, M., Smeets, J.B.J., Brenner, E. (2012) Luminance contrast in the background makes flashes harder to detect during saccades. *Vision Res* **60**, 22-27.

Mann, D. L., Spratford, W., Abernethy, B. (2013). The head tracks and gaze predicts: how the world's best batters hit a ball. *PloS one*, **8**(3), e58289.

Matin, L. and Pearce, D.G. (1965) Visual perception of direction for stimuli flashed during voluntary saccadic eye movements. *Science*, **148**,1485–1488

Mateeff, S. (1978) Saccadic eye movements and localization of visual stimuli. *Percept Psychophys.* **24**, 215–224

Matziridi, M., Brenner, E. and Smeets, J. B. (2015). The role of temporal information in perisaccadic mislocalization. *PloS one*, **10**(9), e0134081.

Mennie, N., Hayhoe, M. and Sullivan, B. (2007) Look-ahead fixations: anticipatory eye movements in natural tasks. *Exp Brain Res*, **179** (3), 427-442

Morrone, M.C., Ross, J. and Burr, D.C. (1997) Apparent position of visual targets during real and simulated saccadic eye movements. *J Neurosci*. **17**, 7941–7953

Mrotek, L.A. and Soechting, J.F. (2007) Target interception: hand-eye coordination and strategies. *J Neurosci*, **27**, 7297-7309

Orban de Xivry, J. J., and Lefevre, P. (2007). Saccades and pursuit: two outcomes of a single sensorimotor process. *The Journal of Physiology*, **584**(1), 11-23.

Pelz, J., Hayhoe, M. and Loeber, R. (2001) The coordination of eye, head, and hand movements in a natural task. *Exp Brain Res*, **139** (3), 266-277

RStudio Team (2018). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA. URL http://www.rstudio.com/

Robinson, D. A. (1965) The mechanics of human smooth pursuit eye movements. *J Physiol* (Lond) **180**, 569-591.

Ross, J., Morrone, M. C., Goldberg, M. E., and Burr, D. C. (2001) Changes in visual perception at the time of saccades. *Trends Neurosci*, **24**(2), 113-121.

Schlag, J. and Schlag-Rey, M. (2002) Through the eye, slowly: delays and localization errors in the visual system. *Nat Rev Neurosci*. **3** (3), 191-215.

Schütz, A.C., Braun,. D.I. and Gegenfurtner, K.R. (2009) Object recognition during foveating eye movements. *Vision Res* **49**, 2241-2253

Smeets, J.B.J., Hayhoe, M. and Ballard, D.H. (1996) Goal-directed arm movements change eye-head coordination. *Exp Brain Res*, **109** (3), 434-440

Soechting, J.F. and Flanders, M. (2008) Extrapolation of visual motion for manual interception. *J Neurophysiol*, **99** (6), 2956-2967

Spering, M., Schütz, A. C., Braun, D. I., and Gegenfurtner, K. R. (2011) Keep your eyes on the ball: smooth pursuit eye movements enhance prediction of visual motion. *J Neurophysiol*, **105** (4), 1756-1767.

van den Berg, A. V. (1988) Human smooth pursuit during transient perturbations of predictable and unpredictable target movement. *Exp Brain Res*, **72**, 95-108.

Xia, R., & Barnes, G. (1999). Oculomanual coordination in tracking of pseudorandom target motion stimuli. *Journal of Motor Behavior*, **31**(1), 21-38.

Zuber, B. L. and Stark, L. (1966). Saccadic suppression: elevation of visual threshold associated with saccadic eye movements. *Experimental Neurology*, **16**(1), 65-79.

Author's note

Experimental data will be archived in a public data repository. The link to access it will be included in a final version of the paper, if accepted. At the moment it is available on request for checking.

ACKNOWLEDGEMENTS

This work was supported by grant NWO 464-13-169 from the Dutch Organization for Scientific Research to EB and by ESRC grant ES/M00001X/1 to SKR.

Figure legends

Figure 1. Schematic representation of the task and conditions. (A) Subjects started with their index finger at the red dot and had to intercept a moving target (black dot) by tapping on it when it reached the white hitting zone. (B) In the Ring condition, the hitting zone was always the same large white ring. (C) In the Disk condition, it was a small white disk at one of 24 possible positions. The white dashed lines in C indicate the other possible positions. They were not visible during the experiment. The six curves in B and C show the six possible paths that the target could take to one of the 24 hitting zones.

Figure 2. Example of gaze, head and hand movements on single trials for a representative subject in the two conditions of Experiment 1. Data of two trials with the same target trajectory from the moment the target appeared until the time of the tap. The colours of the curves change with the remaining time to tap: from black to either grey, blue, red or green (for the target, gaze, head and hand, respectively).

Figure 3. Gaze, head and hand movements of all trials of the same representative subject in Experiment 1 shown in Figure 2. Colours change from black to blue (gaze), red (head) and green (hand) across time from when the target appears to the moment of the tap (as in Figure 2).

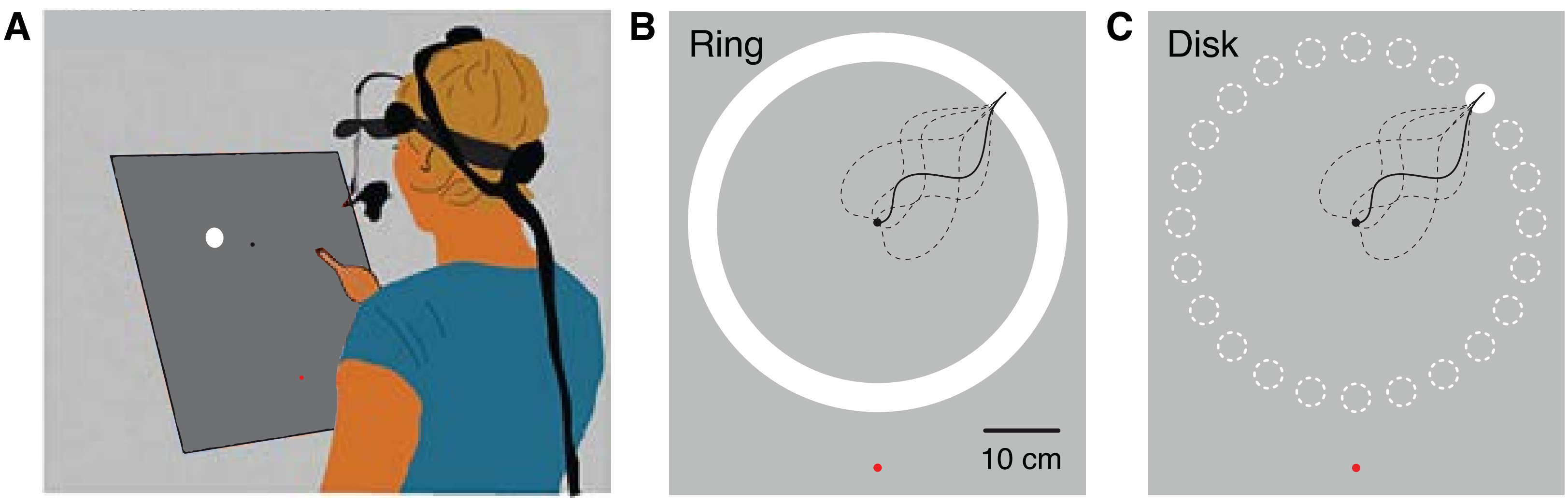
Figure 4. Analysis of the average gaze, head and hand movements of all eight subjects in Experiment 1 (A-C) and of all five subjects in Experiments 2 (D-F) and 3 (G-I). Normalized distance to the tap position as a function of the time until the target is hit for the gaze, head and hand. The lines (continuous for the Ring condition, dashed for the Disk condition) and shaded areas are the means and standard errors of the subjects' individual mean values. A normalized distance of zero corresponds to being at the tap position. A normalized distance of one corresponds to being where the target appeared for the gaze and the head, and corresponds to being at the finger's starting position for the hand. In the gaze panels, we also show the mean normalised distance of the target from the tap position (black dotted curve). The inset in A shows the number of saccades per trial in Experiment 1, split by whether saccades ended closer to the target (black bars) or closer to the tap position (white bars). Error bars are standard errors across the subjects' mean numbers of saccades.

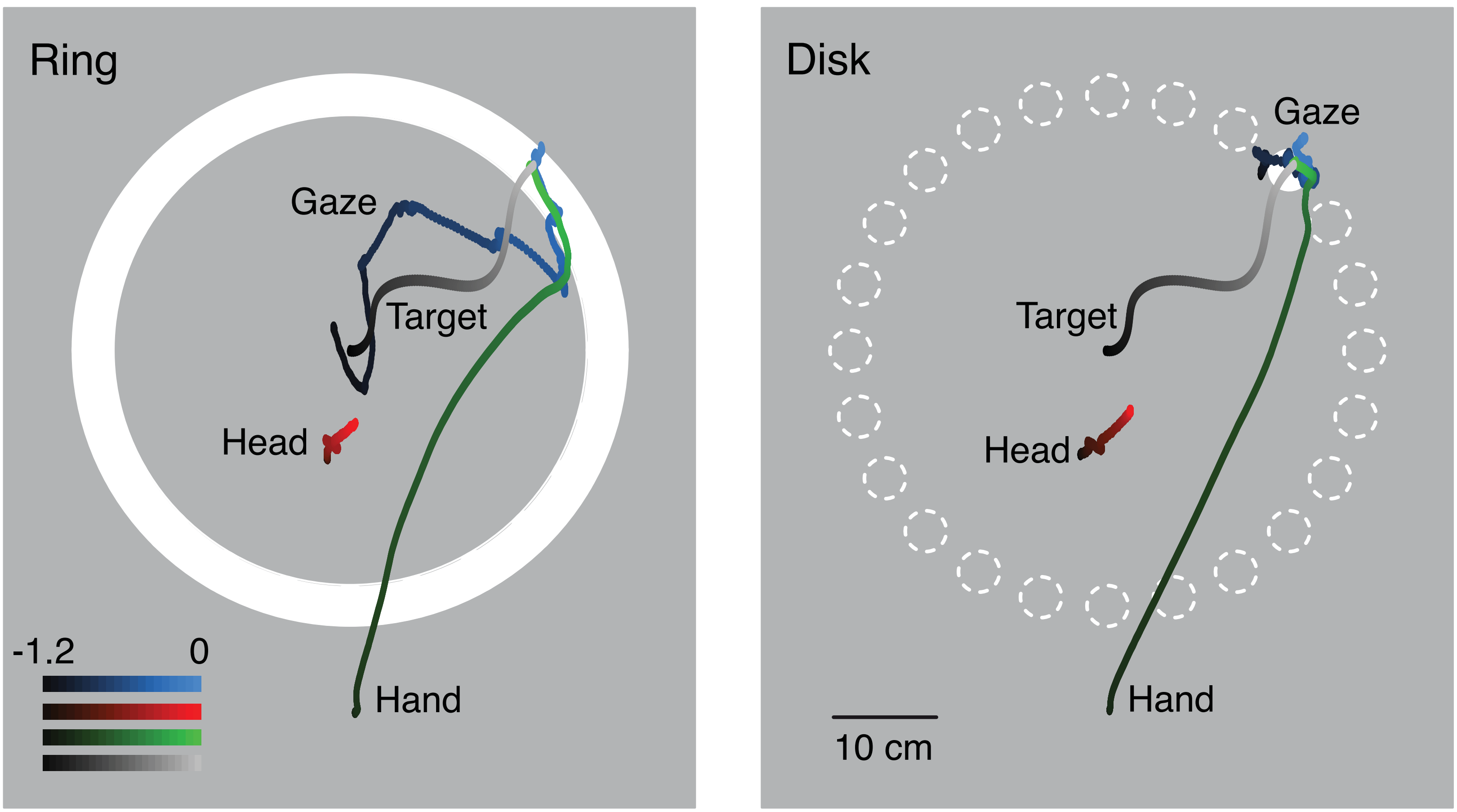
Figure 5. Hand movements of all trials of all eight subjects for the furthest (top panels) and the nearest (bottom panels) hitting zones in Experiment 1. All trajectories start at the hand's starting point near the bottom of the panel. Colour changes from black to green across time as in Figures 2 and 3.

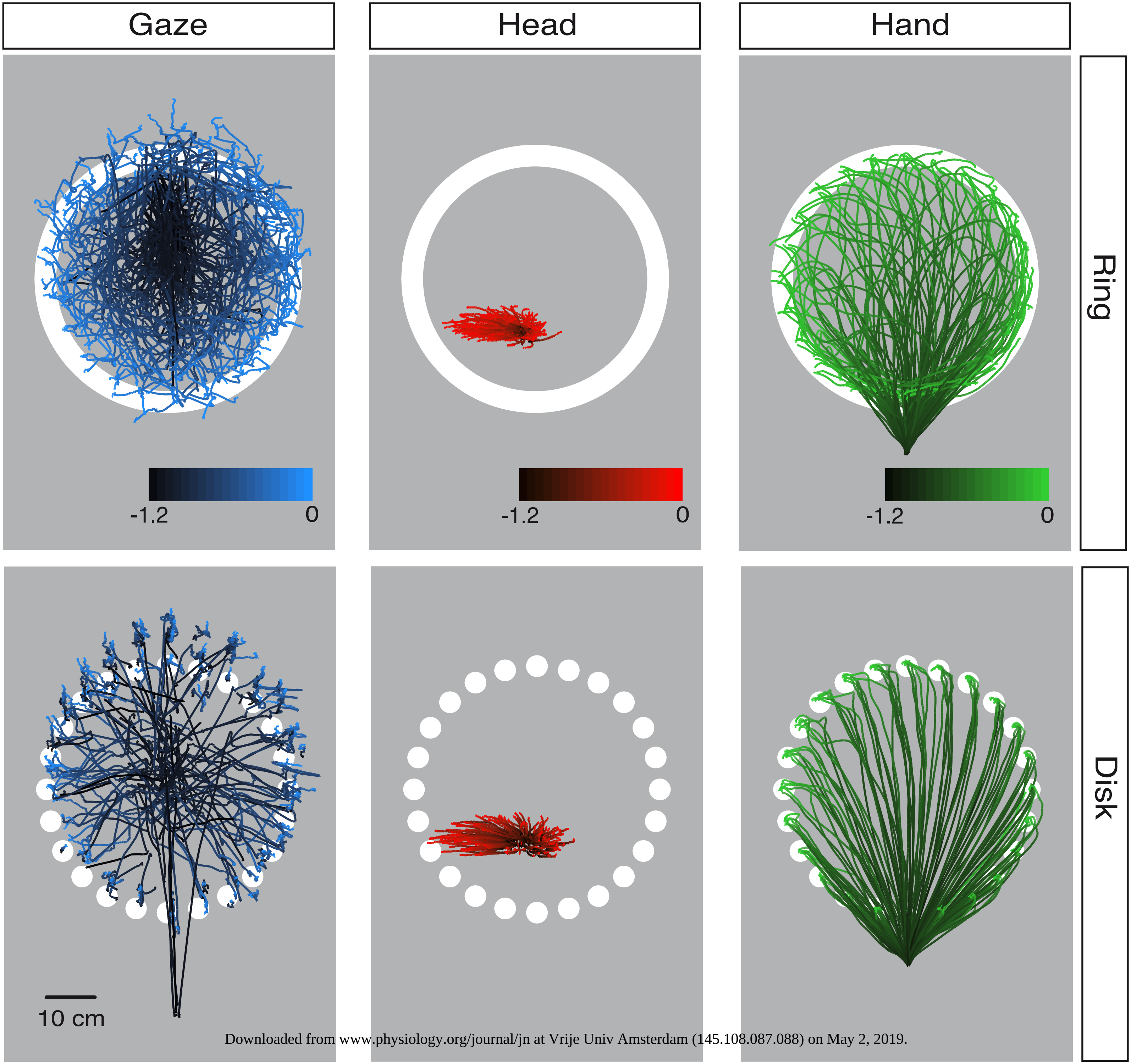
Table legends

Table 1. Percentage of targets hit. A target is considered to have been hit if the finger, as calibrated, was within the bounds of both the target and the hitting zone at the time of the tap. Performance only differed significantly between the Disk and Ring condition in Experiment 1. Performance in Experiments 2 and 3 differed significantly from that in Experiment 1 (Experiment 2, Disk: $t_{4,7}$=2.3, p=0.03; Ring: $t_{4,7}$=5.12, p=0.0003; Experiment 3, Disk: $t_{4,7}$=2.34, p=0.03; Ring: $t_{4,7}$=7.02, p<0.001) but not from each other (Disk: $t_{4,4}$=0.33, p=0.38; Ring: $t_{4,4}$=1.46, p=0.09).

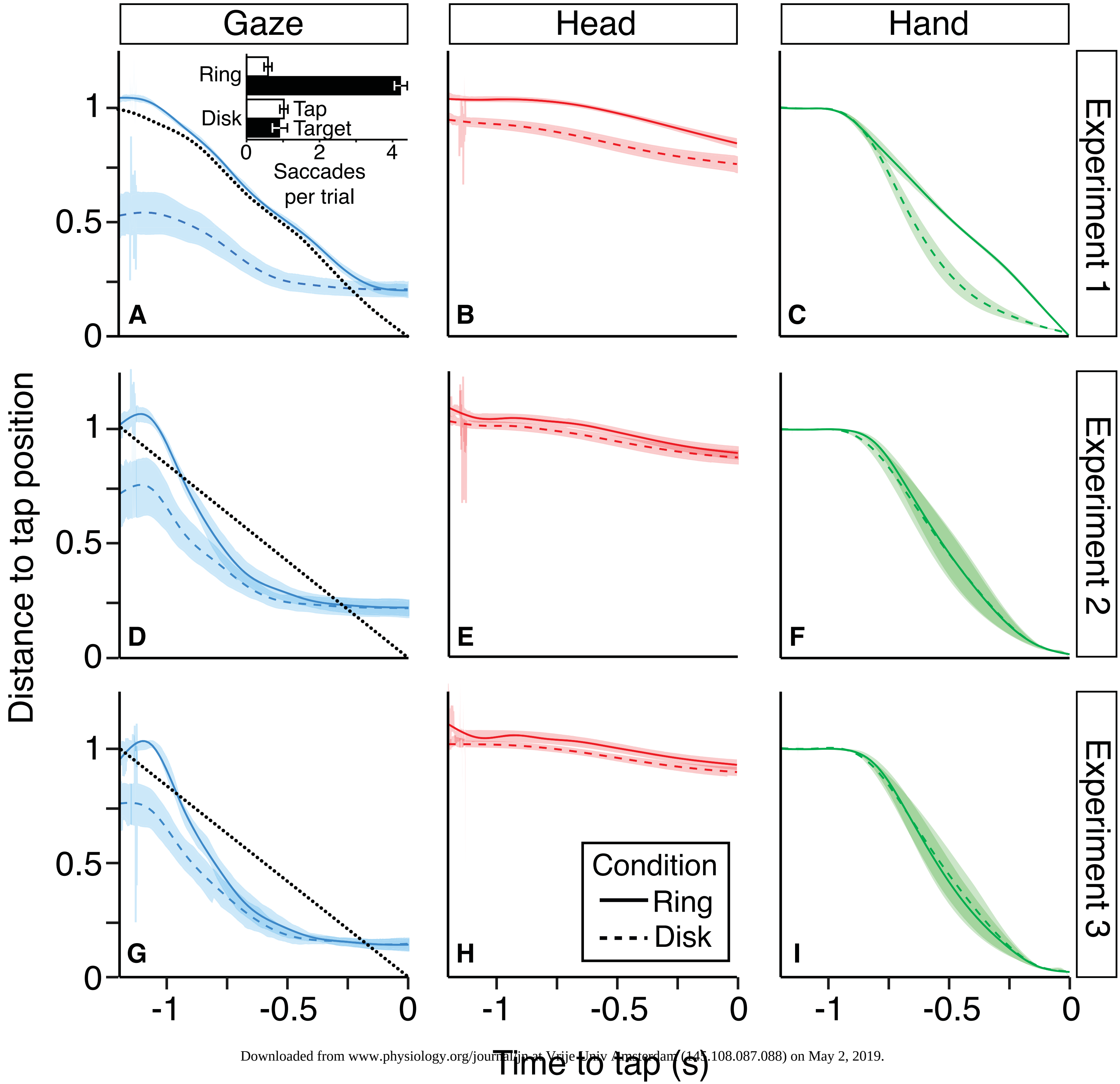
Table 2. Variability in the timing of the hits (standard deviation in ms). Performance only differed significantly between the Disk and Ring condition in Experiment 1. Performance in Experiment 3 differed significantly from that in Experiment 1 (Disk: $t_{7,4}$=1.92, p=0.04; Ring: $t_{7,4}$=3.05, p=0.008), but the other differences between experiments were not significant.

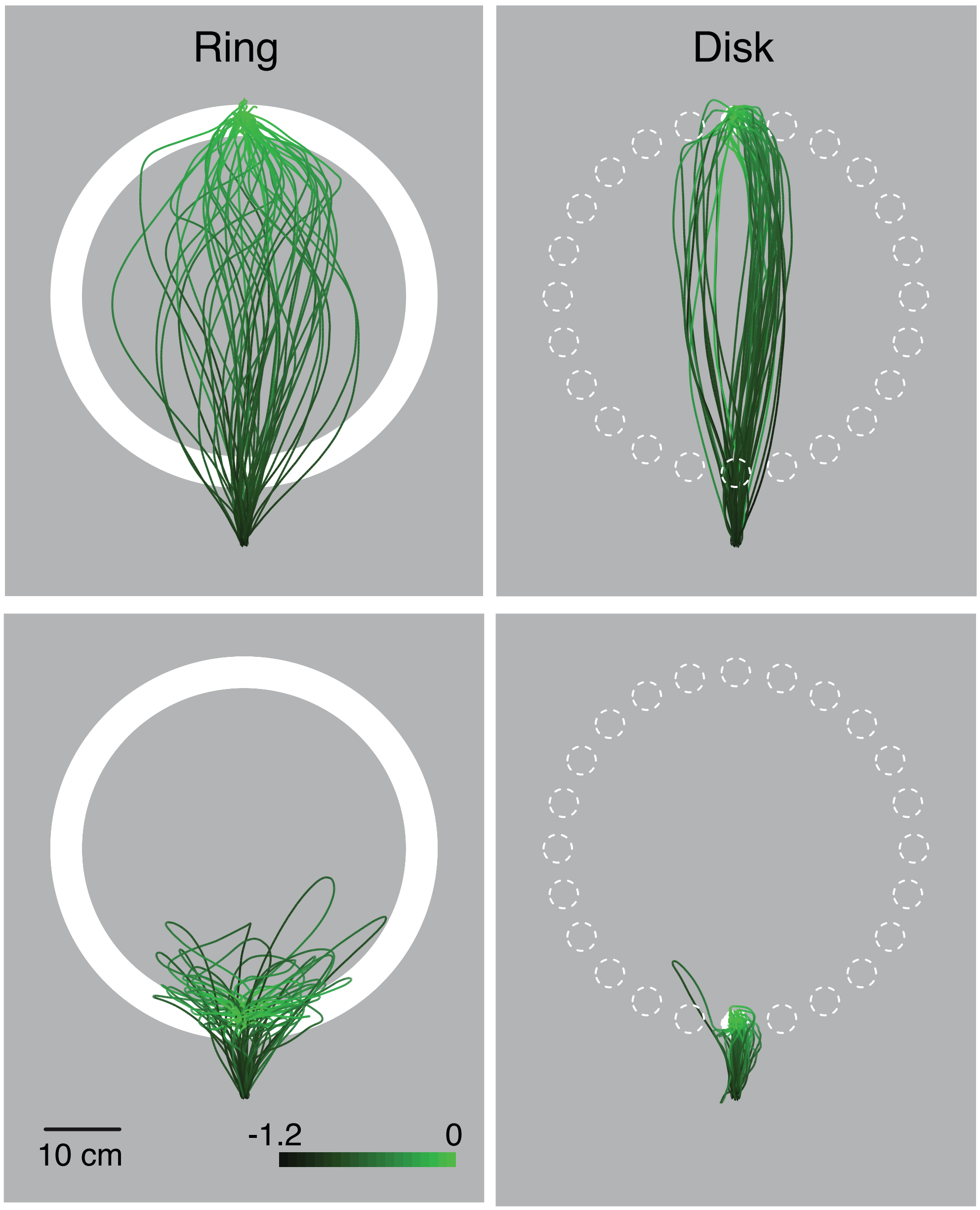

A
B
Ring
10 cm
C
Disk


Ring
Gaze
Target
Head
Hand
-1.2
0
Disk
Gaze
Target
Head
Hand
10 cm



Gaze
Head
Hand
Ring
Disk
-1.2
0
-1.2
0
-1.2
0
10 cm


Gaze
Head
Hand
Experiment 1
Experiment 2
Experiment 3
Distance to tap position
Time to tap (s)
Ring
Disk
Tap
Target
Saccades
per trial
0
2
4
A
B
C
D
E
F
G
H
I
1
0.5
0
-1
-0.5
Condition
Ring
Disk


Ring
Disk
10 cm
-1.2
0