ORIGINAL ARTICLE



Towards accurate cursorless pointing: the effects of ocular dominance and handedness

Katrin Plaumann¹ · Matthias Weing¹ · Christian Winkler¹ · Michael Müller¹ · Enrico Rukzio¹

Received: 16 December 2016 / Accepted: 14 November 2017 / Published online: 7 December 2017 © Springer-Verlag London Ltd., part of Springer Nature 2017

Abstract

Pointing gestures are our natural way of referencing distant objects and thus widely used in HCI for controlling devices. Due to current pointing models' inherent inaccuracies, most of the systems using pointing gestures so far rely on visual feedback showing users where they point at. However, in many environments, e.g., smart homes, it is rarely possible to display cursors since most devices do not contain a display. Therefore, we raise the question of how to facilitate accurate pointing-based interaction in a cursorless context. In this paper we present two user studies showing that previous cursorless techniques are rather inaccurate as they lack important considerations about users' characteristics that would help in minimizing inaccuracy. We show that pointing accuracy could be significantly improved by acknowledging users' handedness and ocular dominance. In a first user study (n=33), we reveal the large effect of ocular dominance and handedness on human pointing behavior. Current ray-casting techniques neglect both ocular dominance and handedness as effects onto pointing behavior, precluding them from accurate cursorless selection. With a second user study (n=25), we show that accounting for ocular dominance and handedness yields to significantly more accurate selections compared to two previously published ray-casting techniques. This speaks for the importance of considering users' characteristics further to develop better selection techniques to foster more robust accurate selections.

Keywords Cursorless distant pointing \cdot Ocular dominance \cdot Handedness \cdot Ray casting \cdot Smart environments \cdot Interaction \cdot Smart objects

1 Introduction

Pointing gestures have been a topic in human-computerinteraction research for a long time. They are easy to perform and our natural means of referencing objects when communicating with each other and thus particularly suitable for interactions with distant objects and devices. So far, most research covering distant pointing dealt with cursor-based pointing [12, 16, 21, 23]. Cursors are particularly useful for mediated pointing, where a device translates the indication of users into a cursor position.

Submitted to Personal and Ubiquitous Computing

This is still the current method of choice to interact with personal computers (indirect cursor steering with mouse or touchpad) as well as large displays (direct relation between pointing target and cursor) [16, 21]. The main advantage of cursor-based pointing is that inaccuracies in recognizing the user's pointing intent are largely uncritical as they are automatically compensated for by the user. This compensation takes place as users consciously adjusts the cursor so that it matches the position of the objects they intend to select. This is true for indirect pointing (e.g., mouse), but as well for direct pointing gestures (e.g., pointing at large displays).

In smart environments, e.g., next-generation factories and offices and especially smart homes, ever more devices are "smartified," i.e., radiators, air conditioning, blinds, curtains, and lights become digitally and remotely controllable, other systems such as hi-fi systems receive network connectivity and APIs that likewise increase their possibilities for remote interaction. However, many of these controllable devices do not offer a display for



Insitute of Media Informatics, Ulm University, Helmholtzstraße 16, 89081 Ulm, Germany

cursor-based feedback, let alone one that is large enough to be visible from a distance. Physical interaction such as pressing buttons or by means of touchscreens has and will be a good alternative, but requires the device to be within physical reach. Remote controls already pile up in users' homes and require to be in physical reach as well. This leaves cursorless pointing as one of the few promising alternatives for future distant interaction with devices in smart environments. Examples of a desired support for interaction range from simple scenarios (like switching individual lights on and off or moving blinds up and down) to more complex ones like changing the source of a hi-fi system or setting the air conditioning to a specific temperature (see Fig. 1).

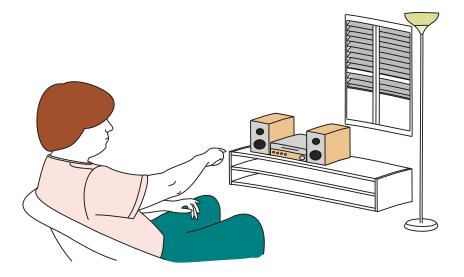
More generally, the types of interaction with smart devices can be broken down into macro and micro selections. By macro selection, in this context, we mean selecting one of the available objects in a room for further interaction. Conversely, by micro selections we mean selecting functions of a device such as moving the blinds. Of course, with simple devices macro selection can be sufficient to directly execute an associated function, for example, turning a light on or off by pointing at it. Instead of using pointing, macro selection could also be achieved using voice input or through a handheld device (remote control or smartphone app); however, this can be cumbersome in case of many or identical objects in the environment due to the entailed list navigation and disambiguation of identical objects (e.g., select one of several identical lamps in a room). In contrast, pointing allows to leverage the environment for direct interaction that should be much quicker and easier to perform. When a device is selected for interaction, micro selections may allow for more fine-grained interactions such as setting a temperature, selecting a type of coffee to prepare for brewing, choosing the position of blinds, and controlling

Fig. 1 A person interacting with appliances in a smart home. The lamp, the blinds, and the hi-fi system are all devices with no display suitable for cursor-based pointing. To facilitate cursorless pointing, a very accurate model of human pointing behavior is needed

home entertainment systems. These are only a few examples for interactions that might occur in a smart home that require precise distant interaction without a cursor. With currently available techniques, none of those are easy to perform. Today, one instead has to use an additional device for interaction with smart devices, such as a remote control, an app on a (previously locked) mobile phone—or one even has to physically approach the device in order to control it.

In this paper, we report on research aiming at improving current cursorless, direct selection methods for smart environments. Our research questions were (1) how ocular dominance and handedness influence the accuracy of existing cursorless, direct selection methods and (2) if a cursorless, direct selection method accounting for the effects ocular dominance and handedness have on pointing behavior would be more accurate than previously presented techniques. Thus, two user studies were conducted. In the first study, 33 participants with normal or corrected to normal vision (left and right handed as well as with left and right dominant eye) were instructed to point at 16 different targets. Their motions were recorded with motion capturing, providing us with a ground truth of their exact pointing positions. The motion capturing data was that used to compare several ravcasting techniques depending on participants' handedness and ocular dominance. Additionally, the motion capturing data we gained from this first study was used to calculate ideal starting points for our proposed ray-casting technique, HRC optimized (Head-based Ray Casting optimized for ocular dominance and handedness).

Unfortunately, as our first user study will illustrate, even with a perfect tracking of hands and gaze, current ray-casting techniques only allow macro selections under strict conditions, while micro selections are currently hard to achieve. Another limitation of current ray-casting techniques is that they imply a certain pointing model that





may be counter-intuitive at least for some users and require prior training and mental demand to always follow exactly this model. In our study, even the best tested previously published technique, eye-finger ray casting, showed an average offset of $18\,cm$ to the target. As the intended direction cannot be known in advance, this enlarges the range of uncertainty to both sides resulting in a possible offset up to $36\,cm$. Though the reduction of this deviation is topic of ongoing research, previous results are still not accurate enough for micro selections in smart homes. A possible reason for this might be their disregard of ocular dominance and handedness as an influential factor on pointing behavior.

To answer our second research question, we used the motion capture data to then calculate ideal starting points for our proposed ray-casting technique, HRC optimized. Here, ocular dominance as well as handedness is accounted for by using different starting points for individuals depending on their handedness and ocular dominance. This leads to an increased pointing accuracy when no cursor is visible. In our second study, conducted to validate the results of the first study, participants were instructed to not only point at different targets but also to simultaneously walk to create arbitrary pointing angles. The results confirm the results of the first study. We thus show that by accounting for ocular dominance and handedness, the offset between target and ray casting result can significantly be improved compared to two previously published ray-casting techniques that depict the state of the art. In particular, we make the following contributions:

- An assessment of the accuracy of several ray-casting techniques for cursorless pointing. It shows how pointing behavior is systematically influenced by ocular dominance, for instance leading to 95.5% of our right-eyed right-handed participants missing the target from the south-east direction, further showing interesting differences between right- and left-eyed users and left-and right-handed users, as for instance left-handed users tending to position their finger more on the left side and right-handed users positioning their finger more on the right side of a target. This effect, however, is weakened when the eye opposite to the dominant hand is the dominant one.
- HRC optimized (Head-based Ray Casting optimized for ocular dominance and handedness), the new ray-casting technique we propose based on the results of our first study, accounting for ocular dominance and handedness of users and leading to significantly more accurate cursorless distant pointing, as a cross-validation shows.
- The validation of HRC optimized in a separate study with greater focus on external validity, including arbitrary pointing positions. We show that even in

such an uncontrolled setting, our technique yields to significantly more accurate results without the need to instruct users. Based on these results, we argue that considering users' characteristics like handedness and ocular dominance allows us to develop more accurate ray-casting techniques and to further improve interactions in smart spaces.

Our findings help interaction designers and system developers to better understand human pointing behavior, thus helping to optimize existing systems and build more accurate future systems. Albeit this paper focuses on *cursorless* pointing, our findings can also be used to improve *cursor-based* pointing. Through a more accurate interpretation of the user's pointing intent, in theory, less correctional movements should be required to finally steer the cursor onto the intended target.

We will start the rest of this paper by presenting related work, followed by the first user study, the newly proposed technique, and the second user study.

2 Related work

2.1 Ocular dominance

Porac and Coren describe ocular dominance as one manifestation of the phenomenon of lateral dominance, where the image of one eye was more heavily relied on than the image of the other eye [24]. They describe three types of ocular dominance: sighting dominance, sensory dominance, and acuity dominance. Sighting dominance could be the most comparable to handedness, describing a behavioral preference for either one image or the other. This can either be measured consciously (selecting one eye for a monocular task) or unconsciously (unconscious selection of one eye for a binocular task). Sensory dominance describes the alternation of the images of both eyes. For acuity dominance, the eye with the greater acuity is seen as the dominant one. Following the argumentation of Porac and Coren [24], we consider sighting dominance as ocular dominance in this paper. Further, we use an unconscious sighting test to asses ocular dominance, since the result of a conscious test could be affected by handedness or training [24]. Porac and Coren further analyzed the effect of sighting dominance on egocentric localization [25]. Although indicating that ocular dominance affects the localization of objects, their study did not involve pointing at targets, what might alter the results. Kahn and Crawford investigated eye-finger alignment for pointing at targets presented on a wide range of horizontal angles [15]. They found that ocular dominance has an effect, stating that the finger is more likely aligned with the eye-target line of the



dominant eye. However, they did not further analyze the position the finger has when aligned with the target. Kahn and Crawford showed that ocular dominance is not merely a static concept, but rather depends on the position one has towards the target one is looking at [14]. They found that this comes into effect at angles beyond 15.5° off center. Yet, in our study, the area participants had to point at targets was limited to 14° off center.

2.2 Motor control theory and hand-eye coordination

Motor control theory describes the process of moving muscles and limbs according to perceived information [9, 19]. Within the field, several theories describing human motor control are discussed [19]. Within motor control theory, hand-eye coordination is of most interest for analyzing pointing behavior, since the sensory information from the eye is processed into the movement of limbs. Regarding trajectory planning, Kawato proposes that both dynamic and kinematic approaches are applied [13]. Todorov and Jordan address the variety and unrepeatability of motor movements and propose a model were feedback is used to only correct deviations interfering with the actual goal [29]. Both Kawato and Todorov and Jordan propose more general, basic models of motor control, whereby we focus on the application in form of distant pointing and target acquisition.

Based on an empirical experiment, Biguer et al. propose a model which enhances target position encoding based on eye and hand movements [4]. Bowman et al. analyzed how the timing of gaze shifts in relation to hand movements when sequentially acquiring targets with a physical handle [5]. They observed a proactive gaze strategy, meaning that the gaze shift to the next target is triggered when the sensor-motor system anticipates that the target will be reached instead of waiting for the visual confirmation. In contrast to the work of Bowman et al. and Biguer et al., we focus on distant pointing instead of physically touching the targets.

Helsen et al. looked at the coupling of gaze and hand [10], as well as the coupling of eye and whole arm, including shoulder and finger [11] while pointing. They focused mainly on the aiming process and the motions involved, not on the pointing position. Wnuczko and Kennedy analyzed pointing behavior for pointing with open and closed eyes [30]. They stated that with open eyes, participants visually aligned the pointing tool with the target. However, they did not take sight dominance into account, and also did not further analyze how the pointing tool is aligned with the target. Rhythmical pointing tasks were subject to the work of Lazzari et al. [20]. They state that hand and eye movements are dynamically synchronized. In contrast to their work, we focus on discrete distant pointing.

2.3 Distant pointing

A large body of work exists covering pointing with pointing devices [7, 17, 28]. This work, however, focuses on pointing without a special pointing device.

To increase distant pointing accuracy, several techniques have been developed. Among them are adaptive pointing [16] and magnetic cursor [21]. Mayer et al. proposed a correction function for several ray-casting techniques useful to increase accuracy for cursorless pointing [22]. Argelaguet and Andujar compared the performance of hand- and eye-based ray-casting in cluttered virtual environments [2]. Their results show that eye-based ray-casting outperforms hand-based ray-casting regarding selection errors. Also, the main focus of improvements of pointing interaction lies on the technical side, e.g., developing algorithms to increase accuracy. Focusing on distant pointing at ultra walls, Nancel et al. developed several methods to increase pointing accuracy for cursor-based pointing [23], whereby we focus on cursorless pointing in smart environments.

Among the little research done in the area of cursorless distant pointing is the work from Cockburn et al. [7]. They evaluated the acquisition of distant targets through pointing with decreasing visual feedback. They state that with reduced visual feedback, the position pointed at more and more drifts away from the actual target. These results substantiate our hypotheses that currently used pointing metaphors are unlikely to represent human pointing behavior. Caon et al. addressed pointing at devices in smart environments. Their solution allowed macro selection, e.g., turning the devices users point at on or off [6]. In contrast to Caon et al., our focus is less on the technical implementation of a gesture recognition system for macro interaction and more on finding a suitable raycasting method for micro selections, allowing for a more fine-grained interaction.

2.4 Ray casting

A common approach for finding the target users point at is ray casting. With ray casting, a ray is cast from the user towards the targets and its intersection with potential targets is calculated [23]. Therefore, a starting point for the ray and either a reference point the ray is supposed to go through or the orientation the ray should follow are defined. Then, a vector is calculated, starting at the starting point and either running through the reference point or following a certain orientation. The vector is then scaled until it runs through a target object. The intersection point between the vector and the target object is calculated and regarded as the point of selection, thus the result of the ray cast. This principle is used by finger ray casting and cyclops eye (or eye finger ray casting). Corradini and Cohen proposed



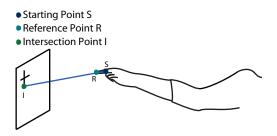


Fig. 2 This image shows how finger ray casting based on [8] (*FRC*) works. Starting at the first knuckle of ones index finger (starting point S colored in dark blue), the ray is calculated as running through the tip of ones finger (reference point R colored in light blue). The ray is then extended so that the intersection point I (green) with the target object, in this case a plain the target cross is displayed on, can be calculated. The intersection point I is regarded as the point of selection and thus the result of the ray cast

finger ray casting [8] as depicted in Fig. 2, were a ray is cast from the base and the tip of the index finger. As can be seen in Fig. 2, the starting point for this ray-casting technique is the person's first finger knuckle. The reference point the ray runs through is the tip of one's finger. Depending on the fingertip's position, the intersection point and thus the point regarded for target selection changes. We will refer to this technique as FRC. An eye-rooted ray-casting technique is cyclops eye [18], where the ray starts from the point between the eyes and goes through the tip of the index finger (EFRC in this paper). EFRC is depicted in Fig. 3 where the starting point S is set to the point between ones eyes. The reference point is again the tip of one's finger. Depending on the fingertip's position, the intersection point and thus the point regarded for target selection changes. Jota et al. compared several ray-casting techniques, stating that the used technique should be selected depending on the task [12]. They also stated the importance of ocular dominance for eye-rooted ray casting. However, they did not analyze effects of ocular dominance on ray casting.

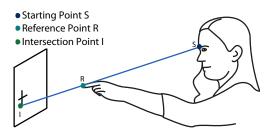


Fig. 3 This image shows how eye-finger ray-casting based on [18] (*EFRC*) works. Starting at the point between one's eyes (starting point S colored in dark blue), the ray is calculated as running through the tip of ones finger (reference point R colored in light blue). The ray is then extended so that the intersection point I (green) with the target object, in this case a plain the target cross is displayed on, can be calculated. the intersection point I is regarded as the point of selection and thus the result of the ray cast

We will compare *FRC* and *EFRC* against our own optimizations and analyze the effect of ocular dominance and handedness on the mentioned techniques.

2.5 User elicitation in HCI

Since Wobbrock et al. presented their approach for finding user-defined gestures [31], user elicitation has been applied to a wide range of gestural interfaces, for example midair gestures for smartphones [27] or in-car interactions [1]. Though we have not conducted a user elicitation study per se, we still consider our work related to user elicitation, since the presented results are based on observed behavior of humans.

3 Study one - understanding pointing and modeling *HRC optimized*

3.1 Approach and goals of the study

The goal of our study is to systematically analyze human pointing behavior with special regard to ocular dominance and handedness. We thus recorded the participants' point of view with sight recording glasses. This data was used to compare (1) the alignment of the fingertip to the target while pointing. Additionally, we observed the body positions while pointing using a motion capturing system to capture a ground truth. This data was used to derive (2) the ideal starting point for ray casting through casting rays from the target back towards the users (further referred to as reverse ray casting). Based on the recorded motion capture data, we then compared (3) the accuracy of known raycasting techniques in the context of cursorless pointing, applied to users' natural pointing behavior. Since (1) and (2) rather observe natural behavior and (3) compares interaction techniques, this study is both exploratory and comparative. The overall aim was to capture natural pointing behavior without giving specific instructions on how to point, and then applying diverse ray-casting techniques on the body position data captured of these pointing positions to analyze how well diverse ray-casting techniques model natural pointing behavior. To achieve this, it is valid to capture natural pointing behavior once and then apply several techniques, all based on calculations on the same body positions while pointing, to the once recorded data. Similar approaches have been taken in the past [22] in similar study settings. Therefore, we do not expect a negative impact. In the following, each type of analysis is explained in detail.

(1) Alignment of the fingertip Previous research in the psychomotoric field indicates that the pointing tool (e.g., the tip of the index finger) is visually aligned to the target [30],



Table 1 The mean age and height as well as the number of participants for the four tested groups right dominant hand and eye, right dominant hand and left dominant eye, left dominant hand and right dominant eye, and both left dominant hand and eye

| Dom. hand | Dom. eye | Age (SE) | Hight (SE) in m | n (females) |
|-----------|----------|------------|-----------------|-------------|
| Right | Right | 25.6 (1.2) | 180.0 (3.9) | 10 (2) |
| Right | Left | 25.6 (2.3) | 179.3 (2.6) | 10 (3) |
| Left | Right | 28.5 (1.5) | 181.5(1.6) | 7 (0) |
| Left | Left | 27.3 (1.9) | 181.0(2.5) | 6 (1) |

but not how the finger is aligned. Thus, one goal was to analyze how the finger is aligned and if there is a dependency with the dominant eye and dominant hand. Our first hypothesis thus was that the alignment of the fingertip with the target differs for persons with different dominant eyes as well as for person with left and right dominant hand (H1). The second hypothesis was that for persons with left dominant eye and right dominant hand as well as for persons with right dominant eye and left dominant hand, the alignment of the fingertip is more centered around the middle of the target, since the effect of ocular dominance balances the effect of handedness on the alignment of the fingertip (H2).

(2) Reverse ray casting Reverse ray casting means casting rays from the target towards the users, going through the tip of the index finger. The tip of the index finger was used because of its importance for pointing [15, 30]. The intersections of the rays with a plane parallel to the projection plane through users' body were calculated. This helped to find a suitable starting point for ray casting leading to our own proposed ray-casting technique, *HRC optimized*.

(3) Comparison of ray casting Based upon the results of the previous steps and upon previous work [8, 18, 22], we applied *FRC*, *EFRC*, and our own technique *HRC optimized* to the collected data and systematically compared their accuracy regarding ocular dominance and handedness. Our hypothesis here was that *HRC optimized* outperforms *FRC* and *EFRC* in terms of accuracy (H3).

3.2 Study design and participants

Participants were positioned 3m in front of the plane targets were projected on (Fig. 4). As pointing happens from

different angles we arranged 16 targets in a 4-by-4 grid across a $2m \times 2m$ space with 48 cm spacing, resulting in a range spanning horizontal and vertical angles of 27° . The lowest row was situated 48 cm above the floor. Targets were resembled by cross-hairs measuring $7 \times 7 cm$ to preclude ambiguity during pointing.

3.3 Apparatus

Motion-capturing and point-of-view video recording were used for data collection to assess the exact positions of the participants' joints and view when pointing (Figs. 4 and 5). Participants wore a jacket, gloves, and a cap equipped with 33 retro-reflective markers for the tracking of the upper body and fingers using an OptiTrack system. The positions of the markers on the arm and fingers are depicted in Fig. 6. The makers were placed and the system was calibrated according to guidelines delivered with the system, resulting in sub-millimeter accuracy. Besides the skeleton data, we also recorded the participant's sight from the cyclops eye, that is the point between the eyes. This was achieved by using glasses equipped with a camera in its bridge (SMI ETG 2.0). The camera records a 1280×960 px video at a rate of 24 fps with an horizontal angle of 60° and a vertical angle of 46°. By equipping the glasses with markers, too, the orientation and location of the glasses' camera in relation to the skeleton and the targets could be calculated. We refrained from collecting eye tracking data

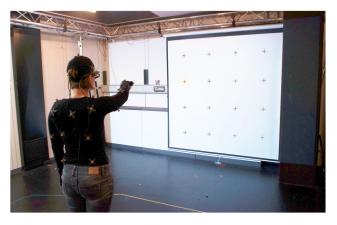


Fig. 4 A participant pointing at the next target (the leftmost target in the second row) highlighted in orange





Fig. 5 The sight recording glasses used in the study. Markers are outlined in red

due to its inherent inaccuracy of one to several degrees which results in large aberrations on a 3 m distance.

3.4 Procedure

Participants started from a relaxed standing position, arms facing downwards. They were instructed to point at the targets as they would naturally do using their whole arm and without a time constraint as we are ultimately looking for a technique that can be used intuitively without prior training or imposed limitations. To reduce the search task when a new target was activated, all targets have always been visible. Participants started pointing at the next target by saying "OK". The experimenter then signaled the system to highlight (in orange) the next randomly selected target for pointing in order for the participant to perform the pointing task. Upon reaching their pointing position, participants said "Yes", which the experimenter forwarded to the system, which presented the target in dark red for two seconds. During this time, participants were advised to hold their position, which was required to calculate an average position that compensates for natural hand tremor. Participants did not receive any feedback indicating where they were pointing at. To avoid or at least reduce arm fatigue, participants were given as much time as they needed to recover before the next pointing gesture, which again started by participants saying "OK". Every participant completed 5 rounds, each consisting of all 16 targets presented in random order. We conducted a Lang-Stereotest to avoid lacking stereo vision influence our results and tested for the dominant eye with the Miles test



Fig. 6 Placement of 9 markers on arm, index, and little finger. Markers are outlined in red

as described by Aswathappa et al. [3]. At last, participants were asked about their demographic data.

3.5 Results

We let 33 participants point 5 times at 16 targets, and thus collected 2,640 pointing gestures. Due to tracking issues, we could use 2614 of them.

3.5.1 Alignment of the fingertip (H1 and H2)

The alignment of the fingertip with the targets was analyzed using the recorded videos taken with the glasses' camera. In a preliminary screening of the videos, nine areas were identified where participants could possibly place their finger: directly on the target and to all eight cardinal directions around it (cf. Fig. 7). Subsequently, all videos were watched and the position of the fingertip for each pointing gesture were assigned to one of the nine areas.

Our hypothesis was that the arrangement of the fingertip differed for persons with different ocular dominance and handedness. Figure 7 shows the areas were the fingertips were positioned and the percentage of gestures were the fingertip was positioned at that area for each dominant eye and dominant hand. As can be seen, the difference is rather high, with right-handed participants with left dominant sight arranging their fingertip more often at the lower left

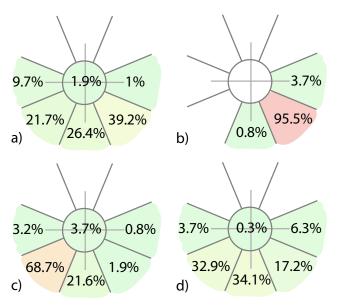


Fig. 7 The fingertips positions towards the target for right-handed participants with dominant left (a) and right (b) eye, as well as left-handed participants with dominant left (c) and right (d) eye. Percentages in the charts show the distribution of actual areas participants placed their finger at in terms of percent of all pointing gestures of each group. We only regarded pointing gestures where no tracking issues occurred



corner of the targets than participants with right dominant sight. Also, for 95.5% of the gestures performed by right eyed right-handed participants, the finger was aligned in the lower right corner, while this position was only used for 39.2% of the pointing gestures performed by left-eyed right-handed participants. For left-handed participants, the pattern is shifted. Here, participants with left dominant eye tended to place their finger more in the lower left corner (68.7%), in contrast to 32.9% of participants with right dominant eye. A chi-squared test confirmed the different finger alignments for the four groups being significant ($\chi^2(15) = 1,704.3, p < 0.001$).

3.5.2 Reverse ray casting

We cast rays from the center of the targets over participants' index fingertips, towards a plane through their bodies. Subsequently, the intersections of these rays with the body plane were calculated. Aiming to find a suitable starting point for ray casting, we calculated the average of all intersection points and its distance to the position of the glasses' camera position. In Fig. 8, the camera's position and the calculated intersection points are depicted as dots (the whiskers show twice the standard error in x and y direction). We calculated the mean intersections for all four groups (right-handed with right dominant eye RR, right-handed with left dominant eye RL, left-handed with right dominant eye LR, and left-handed with left dominant eye LL). For right-handed and right-eyed participants, the intersection point was 1.4 cm right and 2.2 cm below the cyclops eye, and for right-handed and left-eyed participants, it was located 1.6 cm left and 1.4 cm below the cyclops eye. For left-handed participants, the intersection point for those with right dominant eye occurred 2.4 cm left of the cyclops eye and 2.1 cm below it, while for those with a left dominant eye the intersection point was 4.3 cm to left and 2.2 cm below the cyclops eye. The intersection point seems to be influenced by ocular dominance. If the dominant eye is the

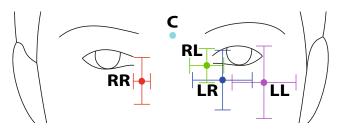


Fig. 8 The different starting points for ray casting based on reverse ray casting: cyclops eye (C, cyan dot), optimized starting point for right handedness and right dominant sight (RR, red dot), right handedness and left dominant sight (RL, green dot), left handedness and right dominant sight (LR, blue dot), and optimized starting point for left handedness and left dominant sight (LL, magenta dot). The whiskers are twice the standard error in x and y direction

one opposite to the dominant hand, the intersection point is shifted in the direction of the dominant eye, while when the dominant eye is on the same side as the dominant hand, the intersection point is on that side of the face.

Based on our assumption that ocular dominance and handedness influence ray-casting accuracy, we assumed further that by using the newly derived intersection points as starting points, pointing accuracy would overall increase. Depending on their ocular dominance and handedness, individuals would have different starting points for ray casting. The reference point for casting the ray would still be ones fingertip, and the intersection point (e.g., point regarded for selecting a target) would be calculated in the same way as with other ray-casting techniques (see Figs. 2 and 3).

Since ocular dominance and handedness introduce a natural shift between target and the result of the ray casting, a technique accounting for that shift and automatically correcting it would foremost be easier and faster to learn, since users would not have to learn to apply a correction factor themselves.

3.5.3 Comparison of ray casting (H3)

Based on previously published ray-casting metaphors, we compared the following ray-casting methods: finger ray casting (further referred to as FRC, see Fig. 2) where a ray is cast through the used finger's knuckle and tip, and eye-finger ray casting (further referred to as EFRC, see Fig. 3) where the ray is cast from the cyclops eye (the point between the eyebrows) through the fingertip. We restrained from implementing elbow finger ray casting, since a recently published similar experiment showed that elbow finger ray-casting results in huge offsets and thus seems not suitable for cursorless pointing [22]. To prove that the starting points for ray casting retrieved in the previous step are also suitable, we included ray casting with our previously retrieved starting points (further referred to as HRC optimized, standing for Headbased Ray Casting optimized for ocular dominance and handedness) for a ray going through the tip of the index finger. To avoid testing a model with the same data it was generated on, a leave one out cross validation was applied. Thus, new starting points were calculated by applying reverse ray casting to the data of all participants of one group except the one the newly calculated starting point was used for. Afterwards, this new starting point for each participant was used for HRC optimized. This procedure was applied to all participants. For each ray-casting method the offset between the intersection of the ray with the target plane and the actual target center was calculated. The rays were cast through the average position of the retrieved skeleton joints while pointing (for eye finger ray casting, the position of the glasses' camera's position was used as the position of the cyclops eye).



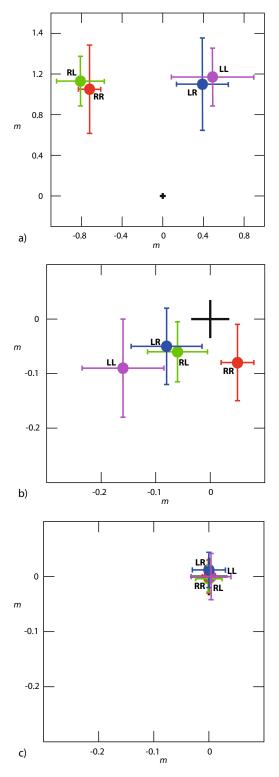


Fig. 9 Average intersection points with the target plane for different ray casting techniques FRC (a), EFRC (b), and HRC optimized (c). The dots marked with RR (red) are for right-handed right eyed participants, RL (green) represents the right-handed left-eyed participants, LR (blue) stands for left-handed right eyed participants, and LL (magenta) for left-handed left-eyed participants. The error bars indicate one standard deviation in x and y direction. The black cross resembles the target. Please note the different scale of Figure (a)

Figure 9 shows the average position of all hitpoints relative to the target. As can be seen, *HRC optimized* (using leave one out cross validation) was the most accurate ray-casting technique. The average offsets for *FRC* were 141.5(2.3 *SE*) for RR, 149.5(1.5 *SE*) for RL, 128.4(4.3 *SE*) for LR, and 141.3 (3.5 *SE*) for LL. For *EFRC*, the average offsets were 14.7 (0.4 *SE*) for RR, 15.9 (0.3 *SE*) for RL, 18.5 (0.5 *SE*) for LR, and 25.0 (0.7 *SE*) for LL. With average offsets of 13.3 (0.1 *SE*) for RR, 14.7 (0.1 *SE*) for RL, 16.8 (0.2 *SE*) for LR, and 20.5 (0.4 *SE*) for LL *HRC optimized* also theoretically achieves smaller offsets (14.6% for all groups).

Since the data did not meet the criteria for parametric tests (homogeneity of variances and normal distribution), only non-parametric statistical tests could be performed. We thus conducted non-parametric Aligned Rank Transforms for mixed designs [32] with handedness and ocular dominance as between subjects and ray-casting technique as within subjects factor for offsets in x and y direction. Using Aligned Rank Transforms allowed us to conduct fullfactorial ANOVAs as well as corresponding post hoc tests [32]. For offsets in x direction, significant main effects occurred for ray-casting method ($F_{2,58} = 35.33$, p < 0.001) and hand ($F_{1.29} = 45.45$, p < 0.001), and a significant interaction between hand and ray-casting method ($F_{2.58}$ = 502.77, p < 0.001). Post hoc comparisons of ray-casting methods with Tukey's HSD method adjusted significance levels showed significant differences between all three raycasting methods (p < 0.05).

For offsets in y direction, ray-casting method was the only significant main factor ($F_{2,58} = 183.84$, p < 0.001). As with offsets in x direction, post hoc comparisons with Tukey's HSD method adjusted significance levels revealed differences between all three ray-casting methods being significant (p < 0.01).

3.6 Discussion

Our results clearly show that ocular dominance has a large effect on distant pointing. Following on work by Porac and Coren [25], people tend to perceive targets offset to the side of their dominant eye and position their finger accordingly. Because the finger is closer to the eye than the target, the positional error affecting the finger will be even bigger resulting in a large angular error between eye and finger, eventually leading to surprisingly large pointing errors on the target plane (especially for longer distances). This inherent inaccuracy in human pointing makes it very difficult for tracking systems to accurately infer where the user is pointing at. Even with gaze tracking in place, without knowledge of the user's dominant eye, this inherent issue cannot be mitigated as it happens in the user's brain. Fortunately, testing for ocular dominance is simple and can be done using manual tests, e.g., focusing



Table 2 The mean age and height as well as the number of participants for the four tested techniques *FRC*, *EFRC*, and *HRC optimized*

| Technique | Age (SE) | Hight (SE) in m | n (females) |
|-----------|-----------|-----------------|-------------|
| FRC | 24.1(0.8) | 180.9(2.5) | 8 (3) |
| EFRC | 23.2(0.6) | 168.3(2.9) | 9 (8) |
| HRC opt. | 25.5(1.2) | 178.8(4.0) | 8 (2) |

on an object through a small hole in the hand, then closing one eye and checking whether it is still visible and centered or through semi-automatic tests, for instance by asking users to tell the middle of a rectangle while wearing shutter glasses that keep only one eye open at a time.

Besides ocular dominance, handedness plays also an important role. Similar to ocular dominance, people also tend to position their finger more to their dominant hand's side of the target, resulting in different error patterns for left and right handed persons. Luckily, handedness is by far easier to detect than ocular dominance and thus straightforward to account for. As we have shown, our ray-casting technique HRC optimized integrating ocular dominance and handedness delivers a theoretical increase by 14.6% in pointing accuracy, as Fig. 9 depicts, and also significantly less aberrations in x and y direction. The achieved offsets still do not allow full micro selections; however, they show that there is still potential to further improve existing ray-casting techniques. By better accounting for users' characteristics, more accurate ray casting methods can be found. And with increased accuracy, micro selections in smart environments become more and more possible. Furthermore, through knowing about the influences of ocular dominance and handedness, existing systems could also achieve better accuracy. Knowledge about users' ocular dominance and handedness could be used to better interpret their pointing intentions.

However, the presented study is limited with regard to pointing position and angle towards the target. In a real smart environment, it is rather unlikely to always stand 3 m in front of the smart devices and point from a relaxed starting position. Also, the angle towards the target might be more extreme. To tackle these issues, we conducted a second study with less controlled starting positions and angles, and compared the ray-casting technique derived from this study (*HRC optimized*) to the finger and finger-eye ray-casting (*FRC* and *EFRC*, respectively) techniques.

4 Study two—validation of HRC optimized

4.1 Approach and goal

With this study, we wanted to validate our previously retrieved ray-casting technique and compare it to *FRC* and *EFRC* in a more realistic setting involving varying distances

and arbitrary angles towards the targets. In particular, we were interested in the general performance of our proposed technique compared to *FRC* and *EFRC* in the mentioned setting. Thus, our hypothesis was that even with arbitrary angels and varying disctances, *HRC optimized* outperforms *FRC* and *EFRC* in terms of accuracy (H4).

4.2 Study design and participants

We chose a one-factorial between-subjects design, the only factor being the used pointing technique. Pointing techniques tested were our own proposed technique, HRC optimized, EFRC, and FRC. We decided to compare our technique to EFRC and FRC because they represent the two categories of existing ray-casting techniques (hand and eye rooted). We recruited 25 participants, but only righthanded and right-eyed ones as the previous study had already investigated differences between left- and right-eyed participants. Further, the majority of the population (around 63% [3, 26]) are right handed and right eyed. None of the participants had participated in study one. For participants' mean age, height, and percentage of female participants see Table 2. Participants were recruited subsequently for each group at our local institution, during a semester break, and it just so happened that in the group for EFRC were more women than men. All participants were rewarded with \in 10. Participants were advised to walk along a path in the form of an "8," and to stop and point at a target when signaled (for a more detailed description, see subsection *Procedure*). The path ran parallel and between a 2.5- and 3.5-m distance to the projected targets. Due to horizontal angles being induced by different pointing positions, we only used one vertical column of four targets. The distances between the targets were the same as in study one. To account for higher targets being harder to point at, we adjusted the targets' vertical position to the shoulder height of participants, thus accounting for height differences between participants.

4.3 Apparatus

The same apparatus as in the first study was used.

4.4 Procedure

The procedure for each group differed only in their instruction. While the *HRC optimized* group was only told



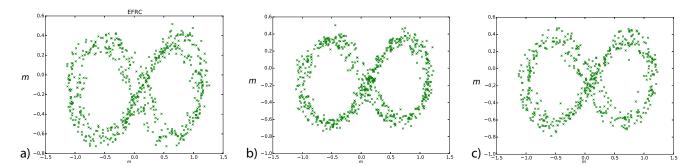


Fig. 10 The pointing positions for the groups *FRC* (**a**)), *EFRC* (**b**)), and *HRC optimized* (**c**)). In all groups, participants walked along an "8" shaped path. The center of the path was 3 m in front of the displayed column of targets

to point at the targets, the other two groups were told to imagine a ray either extending their finger (FRC) or starting from the glasses' camera (cf. Fig. 5) position through their fingertip (EFRC). The point where that ray intersects the wall the targets were projected on was the point they where pointing at. Participants of all groups were advised to walk at their usual pace while all four targets were visible. Upon an auditory signal, a random target was marked orange. Participants were advised to stop at their current standing position and point at the target upon that signal. As in study one, they were asked to indicate when they reached their pointing position, so that the target could be colored red. As in study one, we instructed the participants to hold the pointing position for as long as the target was marked red (2 s). After that, they continued to walk. The next auditory signal occurred 2s later. As in study one, no feedback regarding the pointing position was given. The number of rounds (5) and targets to point at (80, 16 per round) remained as in study one. After each round, participants could take a break for as long as they liked. After the five rounds we proceeded as in study one.

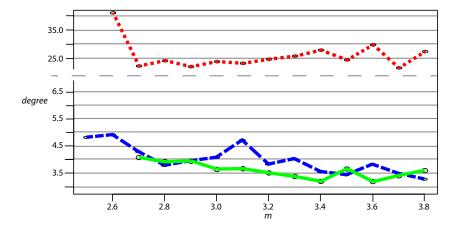
4.5 Results

We let 25 participants point 5 times at 16 targets, and thus collected 2,000 pointing gestures. Due to missing

tracking data, 1,892 of these gestures could be used. The higher amount of not usable data is attributed to participants walking around resulting in more tracking difficulties. The design of this study involved an uncontrolled pointing position to gain gestures from as much random angles as possible. Figure 10 shows the different pointing positions. As can be seen, we achieved an even distribution around the path participants followed for all three groups.

We calculated the offsets by applying the ray-casting technique assigned to each group, and calculated the pointing position as in study one (Fig. 11). For HRC optimized, the starting point retrieved in study one was applied. Subsequently, the offsets towards the targets were calculated. Due to the setting involving several distances towards the targets, all offsets were first calculated as angles. Reported metric offsets are angular offsets calculated in meters for a 3 m distance towards the targets. Figure 12 shows the resulting average hit points for each group. As can be seen, FRC still achieved the highest offset and thus the lowest performance. As in our first study, EFRC's performance was considerably better than FRC's (3.9° and 25.7°, respectively), yet still had a greater offset than HRC optimized (3.7°). HRC optimized is 6.8% more accurate than EFRC. Kruskall-Wallis tests (the data was not normally distributed) confirmed this differences being significant for total offset $(\chi^2(2) = 1183.1, p < 0.001)$, the offset in x direction

Fig. 11 The average angular errors for the groups *FRC* (red dotted line), *EFRC* (blue dashed line), and *HRC optimized* (green solid line). The offset is given in degrees, the distance to the targets in meters





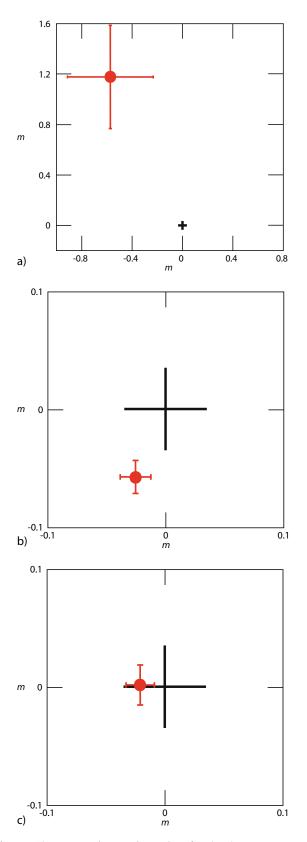
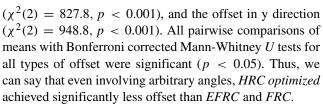


Fig. 12 The average intersection points for the three groups *FRC* (a), *EFRC* (b), and *HRC optimized* (c). Error bars are one standard deviation. The cross resembles the target. Please not the different scales between Figure a) and Figures b) and c)



Next, we looked at the angular error (offset as angle between the ideal ray and the actually cast ray) dependent on the distance towards the target. As distance to the target, we defined the distance between the center between the two shoulder joints and the target's position on the xz-plane. The distance was rounded to the first decimal and the average offsets for each technique and distance were calculated. The results are depicted in Fig. 11. The image clearly shows that *FRC* produces huge offsets (over 20°) from every distance. As the distance independent offsets, *EFRC* and *HRC optimized* are close, however there is a sweet spot were *HRC optimized* clearly outperformes *EFRC*. For distances between 3 and 3.4 m, *HRC optimized* achieves up to 1° lesser offsets than *EFRC*.

4.6 Discussion

Our second study showed that HRC optimized not only significantly decreases offsets for frontal interaction with smart devices, but also for pointing from arbitrary positions and extreme angles, up to one degree when the distance to the targets is between 3 and $3.4 \, m$. We think this is a direct result of people's tendency to visually align their finger with the target, and this process being influenced by the hand used to point and the dominant eye. We have also observed that all three tested ray-casting techniques achieved smaller offsets in the first study. We attribute this to the more relaxed starting position and to the small variation in angle towards the targets in study one. Considering the extreme angles included in this study, the variance of ocular dominance at these angles [14], and that pointing gestures were performed shortly after stopping walking, it is not surprising that the offsets in this study increased.

Besides hard measures like offset, we have to keep in mind that both *FRC* and *EFRC* require users to first learn and then continuously apply a certain ray-casting technique, in our context without the assistance of visual feedback informing them about their performance. *HRC optimized*, on the contrary, does not require users to learn any pointing technique. We thus argue that of the ray-casting techniques covered in this paper, *HRC optimized* is the best suited technique for pointing in smart environments. However, micro interactions are not yet fully possible even when employing *HRC optimized* B. To further explore how micro selections could be fully supported for cursorless pointing, our future research aims at analyzing how other human characteristics influence pointing behavior.



5 Limitations

To reduce the complexity of the studies, we focused on pointing in a standing position only. Pointing from a sitting position may lead to different results, although the results of Mayer et al. kept trends between techniques between standing and sitting conditions [22]. Similarly, the environment of the lab study and the worn equipment may have affected the results, although the latter did not impede movement nor was it mentioned negatively by any of the participants. Another limitation to our studies is the distance towards the target. HRC optimized was developed based on pointing at targets at only a 3-m distance. In the second study, distances varied from 2.6 m up to 3.8 m and overall, HRC optimized still performs well, yielding to up to one degree lesser offsets than EFRC. However, we cannot fully say how the tested ray-casting techniques perform for significantly shorter or larger distances.

The sample sizes (n=33 and 25) are okay as considered by most HCI researchers, but larger samples sizes could possibly lead to more robust starting points for HRC optimized, expecting even greater accuracy. Though we have shown that ocular dominance and handedness do affect pointing behavior and leveraging this behavior helps to create a more accurate technique, we still do not know whether other body characteristics have a similar or even stronger influence and should be accounted for. Addressing this issue is topic of our future research.

6 Conclusion

This paper focused on distant cursorless pointing in smart environments, where pointing is an input modality for both macro and micro selections. Current ray-casting techniques only allow for macro selection when smart objects are positioned around 40 cm apart when users are 3 m away from the objects. Micro selections, which would allow a more fine grained interaction by enabling selections of specific functions of appliances, are still not possible. We thus aimed to examine the effects of ocular dominance and handedness on ray-casting techniques and to show that by accounting for both characteristics, more accurate selections could be possible. In the first study presented in this paper, we used motion capturing and simultaneous video recordings of user's sight to learn how the finger is aligned with the target and what influence ocular dominance and handedness have on the accuracy of ray-casting in a cursorless setting. Subsequently, we used the data to find the best suited ray-casting technique by means of reverse ray casting. The obtained new ray-casting technique, HRC optimized, was validated in a second user study in which we compared it against two previously published ray-casting techniques, namely FRC and EFRC. HRC optimized did not only show an overall better experience, but was and can be used with uninstructed users. This not only allows it to be used right out of the box, but further enables the application to user groups that likely have more difficulties to follow a certain pointing model, elderly people for instance, all the more if it was not their innate one. Notwithstanding, prior training may improve the accuracy of HRC optimized as well, which we will investigate in the future.

Overall, we have shown that ocular dominance and handedness have a strong influence on how the finger is aligned with the target, thus effecting accuracy of raycasting techniques. With this knowledge, users's intentions are better interpretable and correction mechanisms can be better tailored for users as well as for specific use cases. Albeit taking ocular dominance and handedness into account leads to more accuracy, HRC optimized is not yet accurate enough to allow full micro selections. To further explore how micro selection could be possible in cursorless settings is part of our future research. Also, further evaluation is needed to show if HRC optimized is not only suitable for the use by young healthy adults, but also in ambient assistive living settings where users are mostly senior citizens. Nevertheless, our results contribute to the understanding of human pointing, and will serve in constructing a model of human pointing behavior needed for distant cursorless interaction in the future.

Acknowledgements This work was conducted within the Emmy Noether research group Mobile Interaction with Pervasive User Interfaces funded by DFG and the SenseEmotion project funded by BMBF.

References

- Angelini L, Carrino F, Carrino S, Caon M, Khaled OA, Baumgartner J, Sonderegger A, Lalanne D, Mugellini E (2014) Gesturing on the steering wheel: a user-elicited taxonomy. In: Proceedings of automotiveUI. ACM, New York, pp 31:1–31:8. https://doi.org/10.1145/2667317.2667414
- Argelaguet F, Andujar C (2009) Efficient 3d pointing selection in cluttered virtual environments. IEEE Comput Graph Appl Mag 29(6):34–43
- Aswathappa J, Kutty K, Annamalai N (2011) Relationship between handedness and ocular dominance in healthy young adults - a study. Int J Pharm 2(2):76–78
- Biguer B, Prablanc C, Jeannerod M (1984) The contribution of coordinated eye and head movements in hand pointing accuracy. Exp Brain Res 55(3):462–469
- Bowman MC, Johannson RS, Flanagan JR (2009) Eye-hand coordination in a sequential target contact task. Exp Brain Res 195(2):273–283
- Caon M, Yue Y, Tscherrig J, Mugellini E, Abou Khaled O (2011) Context-Aware 3D Gesture interaction based on multiple kinects, IARIA. In: Proceedings of AMBIENT, pp 7–12



- Cockburn A, Quinn P, Gutwin C, Ramos G, Looser J (2011) Air pointing: design and evaluation of spatial target acquisition with and without visual feedback. Int J Hum-Comput St 69(6):401–414
- Corradini A, Cohen PR (2002) Multimodal speech-gesture interface for handfree painting on a virtual paper using partial recurrent neural networks as gesture recognizer. In: Proceedings of IJCNN, vol 3. IEEE, pp 2293–2298
- 9. Friston K (2011) What is optimal about motor control? Neuron 72(3):488–498
- Helsen WF, Elliott D, Starkes JL, Ricker KL (1998) Temporal and spatial coupling of point of gaze and hand movements in aiming. J Motor Behav 30(3):249–259
- Helsen WF, Elliott D, Starkes JL, Ricker KL (2000) Coupling of eye, finger, elbow, and shoulder movements during manual aiming. J Motor Behav 32(3):241–248
- Jota R, Nacenta MA, Jorge JA, Carpendale S, Greenberg S (2010) A comparison of ray pointing techniques for very large displays. In: Proceedings of GI 2010, pp 269–276
- Kawato M (1999) Internal models for motor control and trajectory planning. Curr opin neurobiol 9(6):718–727
- Khan AZ, Crawford JD (2001) Ocular dominance reverses as a function of horizontal gaze angle. Vis Res 41(14):1743–1748
- Khan AZ, Crawford JD (2003) Coordinating one hand with two eyes: optimizing for field of view in a pointing task. Vision res 43(4):409–417
- König WA, Gerken J, Dierdorf S, Reiterer H (2009) Adaptive pointing design and evaluation of a precision enhancing technique for absolute pointing devices. In: Proceedings of INTERACT, 5726. Springer, Berlin, pp 658–671
- Kopper R, Bowman DA, Silva MG, McMahan RP (2010) A human motor behavior model for distal pointing tasks. Int J Hum-Comput St 68(10):603–615
- Kranstedt A, Lücking A, Pfeiffer T, Rieser H, Staudacher M (2006) Measuring and reconstructing pointing in visual contexts. In: Proceedings of the brandial
- Latash ML, Levin MF, Scholz JP, Schöner G (2010) Motor control theories and their applications. Medicina (Kaunas, Lithuania) 46(6):382

- Lazzari S, Mottet D, Vercher JL (2009) Eye-hand coordination in rhythmical pointing. J Motor Behav 41(4):294–304
- Mäkelä V, Heimonen T, Turunen M (2014) Magnetic cursor: improving target selection in freehand pointing interfaces. In: Proceedings of PerDis. ACM, New York, pp 112–117
- Mayer S, Wolf K, Schneegrass S, Henze N (2015) Modeling distant pointing for compensating systematic displacements. In: Proceedings of CHI 2015
- Nancel M, Pietriga E, Chapuis O, Beaudouin-Lafon M (2015)
 Mid-air pointing on ultra-walls. ACM Trans Comput-Hum Interact 22(5):21:1–21:62. https://doi.org/10.1145/2766448
- Porac C, Coren S (1976) The dominant eye. Psychol Bull 83(5):880–897
- Porac C, Coren S (1986) Sighting dominance and egocentric localization. Vis Res 26(10):1709–1713
- Roth HL, Lora AN, Heilman KM (2002) Effects of monocular viewing and eye dominance on spatial attention. Brain 125(9):2023–2035
- Ruiz J, Li Y, Lank E (2011) User-de ned motion gestures for mobile interaction. In: Proceedings of CHI, ACM, New York, CHI '11, pp 197–206. https://doi.org/10.1145/1978942.1978971
- Teather RJ, Stuerzlinger W (2011) Pointing at 3d targets in a stereo head-tracked virtual environment. In: Proceedings of 3DUI. IEEE, pp 87–94 20
- Todorov E, Jordan MI (2002) Optimal feedback control as a theory of motor coordination. Nat neurosci 5(11):1226–1235
- Wnuczko M, Kennedy JM (2011) Pivots for pointing: visuallymonitored pointing has higher arm elevations than pointing blindfolded. J Exp Psychol Hum 37(5):1485–1491. https://doi.org/ 10.1037/a0024232
- Wobbrock JO, Morris MR, Wilson AD (2009) User-de ned gestures for surface computing. In: Proceedings of CHI. ACM, New York, pp 1083–1092. https://doi.org/10.1145/1518701.1518866
- Wobbrock JO, Findlater L, Gergle D, Higgins JJ (2011) The aligned rank transform for nonparametric factorial analyses using only anova procedures. In: Proceedings of the CHI. ACM, New York, pp 143–146. https://doi.org/10.1145/1978942.1978963

