Psychological Research
DOI 10.1007/s00426-016-0824-8

ORIGINAL ARTICLE

# How to point and to interpret pointing gestures? Instructions can reduce pointer–observer misunderstandings

**Oliver Herbort**[1] · **Wilfried Kunde**[1]

Received: 11 May 2016 / Accepted: 1 November 2016


**Abstract** In everyday communication, people often point. However, a pointing act is often misinterpreted as indicating a different spatial referent position than intended by the pointer. It has been suggested that this happens because pointers put the tip of the index finger close to the line joining the eye to the referent. However, the person interpreting the pointing act extrapolates the vector defined by the arm and index finger. As this line crosses the eye-referent line, it suggests a different referent position than the one that was meant. In this paper, we test this hypothesis by manipulating the geometry underlying the production and interpretation of pointing gestures. In Experiment 1, we compared naïve pointer-observed dyads with dyads in which the discrepancy between the vectors defining the production and interpretation of pointing acts has been reduced. As predicted, this reduced pointer–observer misunderstandings compared to the naïve control group. In Experiment 2, we tested whether pointers elevate their arms steeper than necessary to orient it toward the referent, because they visually steer their index finger tips onto the referents in their visual field. Misunderstandings between pointers and observers were smaller when pointers pointed without visual feedback. In sum, the results support the hypothesis that misunderstandings between (naïve) pointers and observers result from different spatial rules describing the production and interpretation of pointing gestures. Furthermore, we suggest that instructions that reduce the discrepancy between these spatial rules can improve communicating with pointing gestures.

## Introduction

Gesturing is an elementary part of human communication. One of its various functions is to refer to specific objects or events in the environment. This is typically accomplished by pointing, that is, by extending the arm and the index finger toward a referent. Pointing is pervasive in human communication (Butterworth, 2003; Roth, 2001) and already used and comprehended by 1-year-olds (Behne, Liszkowski, Carpenter, & Tomasello, 2012; Carpenter, Nagell, Tomasello, Butterworth, & Moore, 1998; Leavens & Hopkins, 1999). While pointing may facilitate communication in many situations, pointing has proven to be less useful when a high degree of spatial acuity is required and when the referent cannot be directly touched. In such situations, observers of pointing gestures frequently misidentify the referent implied by a pointer (Herbort & Kunde, 2016; Lücking, Pfeiffer, & Rieser, 2015). Even though misunderstandings of pointing gestures can often be resolved by verbal communication, also the understanding of pointing per se is essential. As an example, consider the following situation. A person walking on a beach has spotted a drowning swimmer and now wants to indicate the swimmer's position to a lifeguard. As the ocean lacks landmarks, it is very difficult to communicate the location verbally. Hence, the person has to rely on pointing. As the swimmer's head is not easily spotted and occasionally occluded by waves, the lifeguard needs the information conveyed by the pointing gesture to narrow his search for

**Electronic supplementary material** The online version of this article (doi:10.1007/s00426-016-0824-8) contains supplementary material, which is available to authorized users.

✉ Oliver Herbort
oliver.herbort@psychologie.uni-wuerzburg.de

[1] Department of Psychology, Julius-Maximilians-Universität Würzburg, Röntgenring 11, 97070 Würzburg, Germany

Published online: 10 November 2016

the swimmer. A less dramatic but common scenario is the interaction with pre-verbal infants. Infants point frequently and caregivers often provide verbal labels for the pointed at objects—thus contributing to language acquisition (Butterworth, 2003; Zhen & Gros-Louis, 2015). Again, the caregivers need to rely on the gestures alone. In both situations, the effective communication with pointing gestures is essential because verbal descriptions are not possible. However, even when pointing is accompanied by speech, the understanding of pointing gestures facilitates interactions. For example, the number of words necessary to agree on a referent is considerably lower when pointing is unambiguous (Bangerter, 2004). Consequently, pointing is used less frequently when it is likely to fail (Bangerter, 2004; Pechmann & Deutsch, 1982). In sum, the ability to point accurately and comprehend pointing gestures is important to facilitate communication in many situations.

Recent research has shown that despite the widespread use in everyday communication, pointing gestures are systematically misunderstood, at least when studied in the laboratory. For example, observers judged pointing gestures to be directed to a systematically higher position than was implied by the pointer (Herbort & Kunde, 2016). These systematic misunderstandings have been attributed to the different geometric rules guiding the production and interpretation of pointing gestures (Herbort & Kunde, 2016; Wnuczko & Kennedy, 2011). For example, in a recent study, we asked participants to point at various vertical positions and recorded the positions of their shoulders, eyes and index fingers (Herbort & Kunde, 2016). Most participants pointed by extending the arm and putting the index finger between their eyes and the referent. When the same participants were asked where (computer-generated) pointers were pointing, they appeared to extrapolate the vector defined by the extended arm and index finger. These extrapolations were increasingly biased toward a horizontal axis with increasing distance between pointer and referent. Figure 1 illustrates how these methods of pointing and interpreting pointing gestures cause systematic misunderstandings. When the pointer elevates the arm to put the index finger (F) between the eye (E) and the referent (B), the arm is actually oriented toward a higher position (A). As the observer extrapolates the elevation of the arm–finger line (S–F–A in Fig. 1), they judge the pointed at location to be higher than implied by the pointer. For example, when we asked ten persons to indicate where the person in Fig. 1 was pointing, all marked a location close to position A (black bars), even though the pointer has been instructed to point to position B. Moreover, observers use this method for a number of observer perspectives, including looking over the pointer's shoulder (Herbort & Kunde, 2016).

The above explanation for misunderstandings in pointer–observer communication has been derived from recording the kinematics of isolated pointing acts and the interpretation of static pointing gestures (Herbort & Kunde, 2016; Wnuczko & Kennedy, 2011). We now want to extend this work in two ways. First, the above explanation has been derived by relating geometric features of unconstrained, naive pointing gestures to the actual or estimated referent location. Here, we want to further test the hypothesis by actively manipulating the geometry underlying the production and interpretation of pointing acts (e.g., by manipulating the origin of the eye–finger vector relative to the shoulder). Second, we want to derive means that could help to overcome misunderstandings of pointing gestures. Hence, we tried to examine pointing in a relatively natural, dyadic setting and did not encumber pointers by a motion-tracking system.

Given humans' general ability to follow instructions, it may seem trivial that instructions to point in specific ways will actually change pointing behavior. However, on closer inspection, the extent to which such instructions make it to actual behavior is not so trivial (as our results eventually reveal). For example, explicit instructions do not necessarily improve perceptual (Poulter, Jackson, Wann, & Berry, 2005) or perceptual–motor skill learning (Sanchez & Reber, 2013). Moreover, in the case of pointing, the instructions have to overcome habitual processes that have been shaped in an almost lifelong experience.

## Experiment 1

In Experiment 1, we compared three groups of pointer–observer dyads. The task of the pointer was to communicate a number on a vertically oriented number line to an observer. The task of the observer was to identify this number. We focused on the vertical component, because systematic misunderstandings between pointers and observers occur very consistently (Bangerter & Oppenheimer, 2006; Herbort & Kunde. 2016; c.f. Wnuczko & Kennedy, 2011) and because it is relevant in many situations (e.g., pointing at which height a painting should be fixed at a wall, or at a specific star in the night sky). The generality of the results with respect to other spatial dimensions will be discussed in the "General discussion".

According to the hypothesis outlined above, observers misunderstand pointers because the vector extrapolated by an observer (arm–finger line) differs from the vector that actually connects a pointer's posture with the referent (eye–finger line). To test the hypothesis, we reduced the discrepancy between both vectors and evaluated whether this also reduced misunderstandings between the pointers and observers. In Experiment 1, we tested three

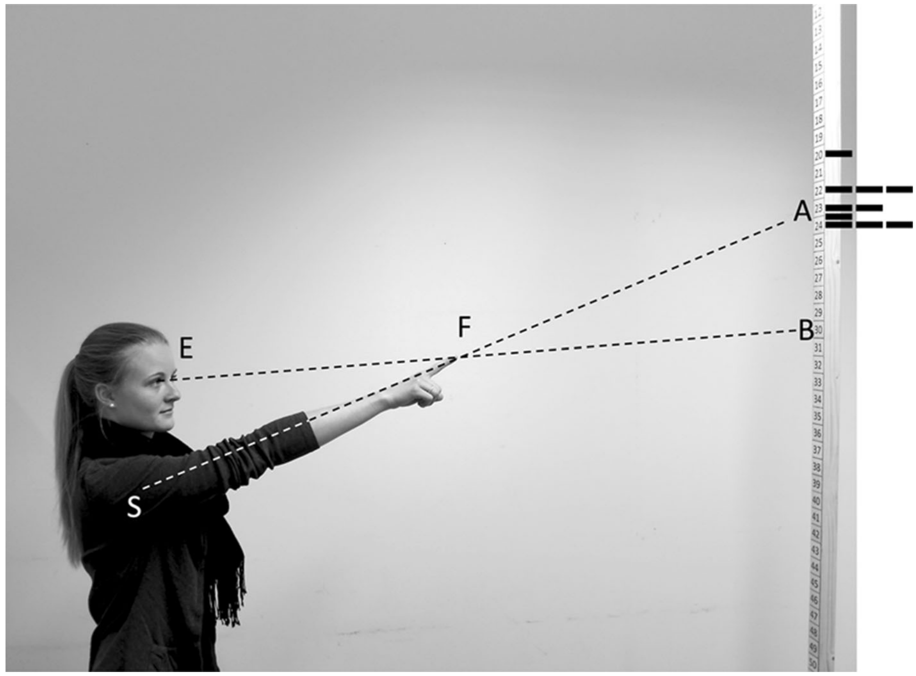


**Fig. 1** A simple model of misunderstandings of pointing-based communication. When a pointer intends to point somewhere (position B in the figure), she puts the index finger (F) between the eyes (E) and the referent (B). By contrast, the pointing person appears to be pointing at position A for an observer, because the arm (i.e., the vector from shoulder S to finger F) aims in this direction. The *black bars* show the responses of ten participants pointing (all women, mean age 26 years) who were asked where the depicted person was pointing. The pointer in the photo was instructed to point to position B. The photo that was handed to the participants had no annotations

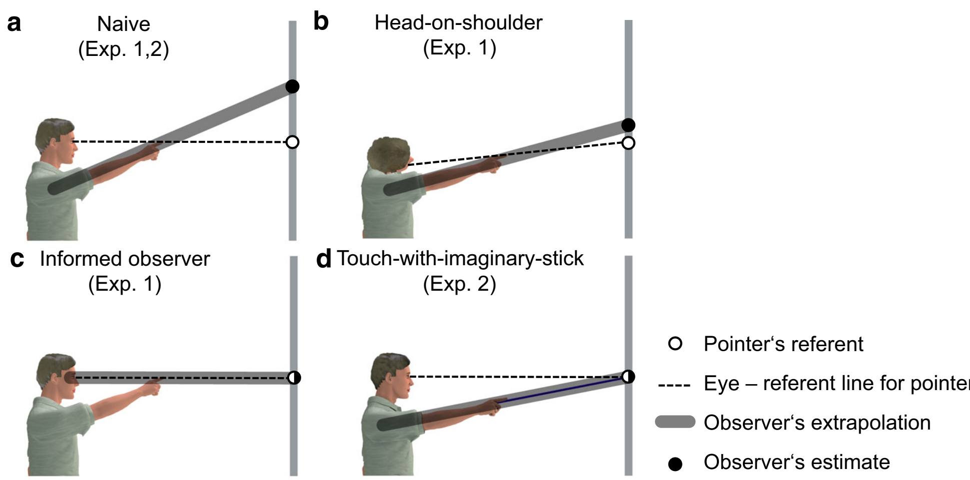


**Fig. 2** The figure shows the predictions for Experiment 1 (**a**–**c**) and Experiment 2 (**a**, **d**)

independent groups of dyads. One group of naïve dyads served as a control group. In the other two groups, the discrepancy between vectors was reduced. In one of those groups, the eye–finger line was brought closer to the arm–finger line by reducing the distance between the eye and the shoulder. In the other group, observers were instructed to extrapolate the eye–finger vector, thus equating the hypothetical rules describing pointing production and interpretation.

In the following, the manipulations in each group are described in more detail. Figure 2a–c shows the geometric rationale underlying the experimental conditions. If pointers align eye, index finger, and referent but observers align shoulder, index finger, and referent, systematic misunderstandings should be eliminated, if eye and shoulder would occupy the same position. This may be anatomically impossible, but can be approximated by moving the eyes as closely as possible to the arm (Fig. 2b). Hence, pointers in the *head-on-shoulder* group were instructed to rest the head on the right shoulder and look through the right eye. The observers in the head-on-shoulder group were asked to identify the referent, but did not receive any special instructions.

Likewise, if the hypothesis above was correct, misunderstandings should be reduced when observers tried to extrapolate the eye–finger line instead of the arm–shoulder line. Hence, we informed observers that most pointers put the index finger between the eye and the referent and that hence the eye–finger line should be used to identify the referent (*informed observer* group, Fig. 2c). The pointers in the informed observer group were told to point as accurately as possible without specifying how to point.

The accuracy of the referent estimates of the head-on-shoulder group (with naïve observers) and informed observer group (with naïve pointers) was compared to a third group in which pointers and observers were naïve (*naive* group, Fig. 2a). Neither pointers nor observers received any special instruction other than to point as accurately as possible and to identify the referent, respectively.

The accuracy of the pointer–observer communication was operationalized as the mean signed error between the referent implied by the pointer and the referent identified by the observer. As a reduction in signed error could be undone by a concurrent increase of the estimates' variability, we also calculated absolute errors and the percentage of trials in which the observer correctly identified the referent. These variables are reported in Online Resource 1. In both experiments, absolute errors closely resembled the signed errors and correct identifications of the referent were relatively rare.

If misunderstandings of pointing gestures resulted from a discrepancy between the geometry underlying the production and interpretation of pointing acts, the signed error should be lower in the head-on-shoulder group and informed observer group than in the naïve group. Furthermore, if the discrepancy between the vectors defining the production and interpretation was smaller in the head-on-shoulder condition and informed observer condition than in the naïve condition, the errors should increase faster with distance in the naïve condition than in the other two conditions (Fig. 2). That is, we expect the errors in the naïve condition to be only a little larger than those in the other conditions when the distances between the pointer and referent is small. However, when the pointer–referent distance is larger, we expect the errors in the naïve condition to be considerably higher than those in the other conditions. Finally, we expect that observers in the naïve group overestimate the referent positions and that the extent of the overestimation increases with the distance between the pointer and referent.

## Method

### *Participants*

Seventy-two students and staff of the University of Würzburg gave informed consent and were compensated for participation (35 women, mean age 22 years). According to the handedness scale of the Lateral Preference Inventory (Coren, 1993), one pointer was left-handed (but also pointed with the right hand) and 34 were right-handed (one participant did not fill out the handedness questionnaire). The sample size allowed to detect an elimination of systematic misunderstandings (effect size estimated from Herbort & Kunde, 2016, Experiment 4) with a probability of $p > 0.95$ and to counterbalance the order of the levels of the factor distance.

### *Stimuli and procedure*

Figure 3a shows the setup of the experiment. Pointers pointed at a vertical number line that was attached to a wooden pole. The number line consisted of white squares with black borders (4 cm × 4 cm), numbered from 1 (284 cm above floor) to 64 (32 cm above floor). The vertical distance between the center of adjacent squares was 4 cm. The pole was positioned in front of the pointer. The observer stood 2 m to the right of the pointer. Both participants received a list on a clipboard. The pointer's list consisted of referent numbers that were labeled with letters. It was placed on a music stand so that it could only be seen by the pointer. The observer received a list of empty lines that were also labeled with letters.

A trial began when the experimenter named a letter (e.g., "A" in the first trial, "B" in the second trial, and so on). The pointer then pointed to the referent number on the list that was labeled with that letter. Then, the observer noted her guess of the pointer's referent number in her list next to the announced letter. Once the observer said "yes", the pointer lowered the arm and the next trial started.

Odd referent numbers from 27 to 47 were used, corresponding to heights of 180, 172,…, and 100 cm above the floor. The experiment consisted of three blocks of 11 trials, each, which were presented in random order. The distance between the pole and pointer differed between blocks (1, 2, 3 m). The order of distances was counterbalanced over dyads and randomly assigned. Dyads were randomly assigned to one of the three groups and the members of each dyad were randomly assigned to the roles of the pointer and observer.

In the naïve and head-on-shoulder condition, observers were only instructed to identify the referent. In the informed observer condition, the observers got the following information, which was originally provided in German: "Most persons look at the referent during pointing and put the index finger on the line between eye and referent. The referent can thus be found by considering the eye–index finger line. Do not consider the orientation of the arm, because it is often misleading". The instruction was complemented by a diagram of a pointing person, which included a line from the eyes, through the index finger tip, to a referent.

In the naïve condition and informed observer conditions, pointers did not receive specific instructions other than to indicate the position of the referent to the observer by

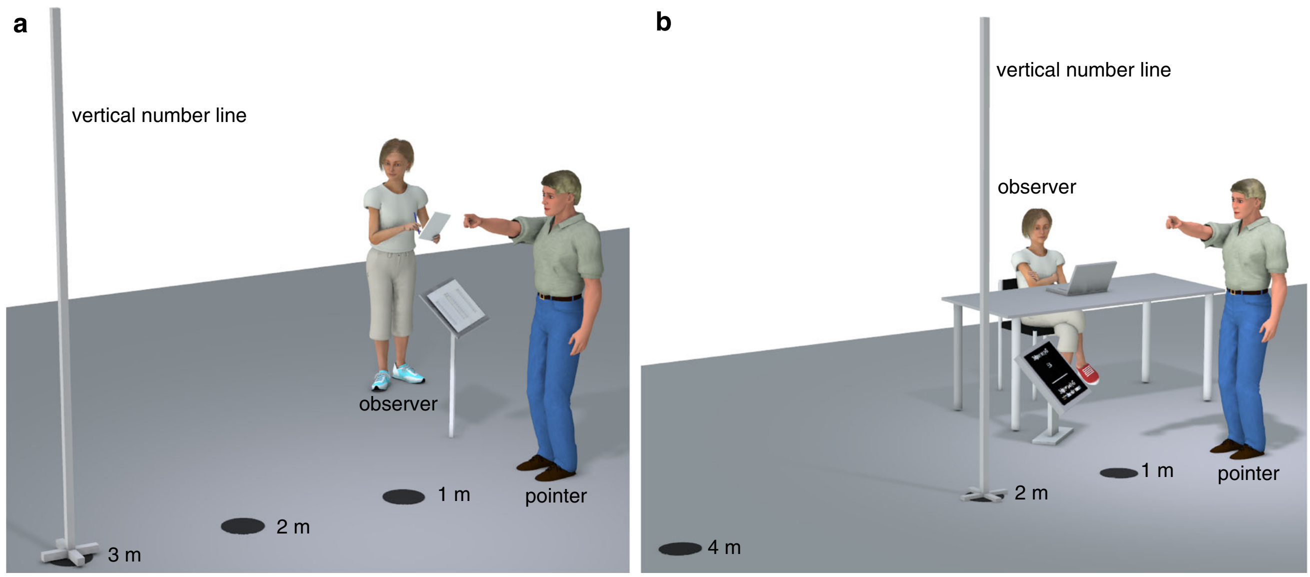


**Fig. 3** The figure shows the setup of Experiments 1 (**a**) and 2 (**b**)

pointing with the right arm. In the head-on-shoulder condition, the pointer's instruction was: "Rest your head on your right shoulder. Close the left eye and look through the right eye. Extend arm and index finger. Your index finger should be on the target from your point of view".

The data of all trials of all participants were included in the analysis, except one illegible response (0.1% of all trials) and three outliers (0.3% of all trials), in which responses deviated from what could be expected from a linear regression of the observer's estimate on the actual referent position for the specific dyad and distance (absolute standardized residual >2.5).

## Results

For the analysis, we converted the referent numbers given to the pointers and estimated by the observers into their vertical position in centimeters. The signed difference between the observers' estimates and pointers' referents was averaged for each distance and dyad. Positive signed errors occurred when the observer's estimate was higher than the pointer's referent. Table S2 in Online Resource 1 reports signed errors, absolute errors and the percentage of correctly identified referents, as well as the results of $t$ tests comparing the different instructions.

Figure 4a shows signed errors (c.f. Table S1). For all distances and all instructions, the observers' estimates were on average too high, all $t(11)\text{s} \geq 4.4$, all $p\text{s} \leq 0.001$, all $g\text{s} \geq 1.27$. Hence, systematic misunderstandings persisted in every experimental condition. A split-plot ANOVA[1] with within-participant factor distance and between-participant factor instruction revealed that signed errors depended on the distance, $F(2,66) = 189.8$, $p < 0.001$, $\eta_p^2 = 0.852$. There was a main effect of instruction, $F(2,33) = 3.7$, $p = 0.036$, $\eta_p^2 = 0.183$. Additionally, the interaction was significant, $F(4,66) = 3.9$, $p = 0.013$, $\eta_p^2 = 0.190$. Pairwise $t$ tests showed that the errors increased from 1 to 2 m and from 2 m to 3 m in each instruction group, all $t(11)\text{s} \geq 4.1$, all $p\text{s} \leq 0.002$, all $g\text{s} \geq 1.192$. The informed observer group made smaller errors than naïve group at the group at the 2 m distance ($t[22] = 2.7$, $p = 0.013$, $g = 1.100$) and marginally smaller errors at the 3 m distance ($t[22] = 1.8$, $p = 0.085$, $g = 0.736$). No significant difference was found at 1 m distance, $t(22) = 1.1$, $p = 0.300$, $g = 0.434$. Likewise, the head-on-shoulder group made smaller errors than the naïve group at 2 m distance ($t[22] = 3.1$, $p = 0.006$, $g = 1.253$) and 3 m distance, $t(22) = 2.6$, $p = 0.015$, $g = 1.076$. No significant difference was found at 1 m distance, $t(22) = -0.5$, $p = 0.633$, $g = -0.204$. The errors of the informed observer group did not differ significantly from those of the head-on-shoulder group at either distance, all $t(22)\text{s} \leq 1.5$, all $p\text{s} \geq 0.141$, all $g\text{s} \leq 0.615$.

## Discussion

The aim of Experiment 1 was to test whether reducing the discrepancy between the hypothesized vectors characterizing pointing production and interpretation also reduces pointer–observer misunderstandings. When the eye–finger and arm–finger vector were aligned more closely in the head-on-shoulder condition, misunderstandings were indeed reduced. Likewise, when observers were instructed

[1] We report Greenhouse Geisser corrected $p$ values but uncorrected dfs throughout the article.

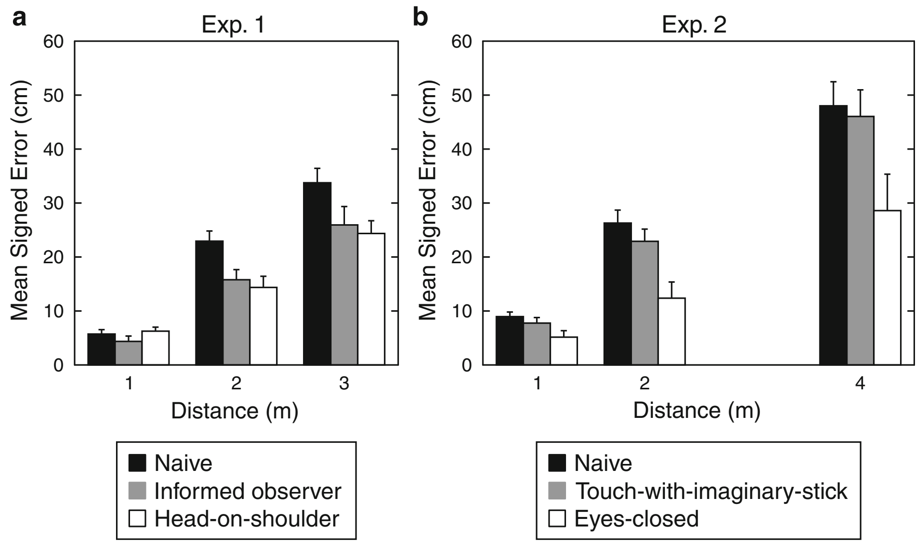


**Fig. 4** The figure shows the mean signed errors of Experiments 1 (**a**) and 2 (**b**) by pointer–referent distance and instruction. Positive errors indicate that the observers' estimates were higher than the pointers' referents. *Error bars* show 1 SEM

to use the geometric rules that describe naive pointing production (alignment of eye, index finger, and referent) misunderstandings were reduced. Moreover, observers of the naive group systematically overestimated the referent position and overestimations increased with pointer–referent distance. This pattern is in line with the hypothesis that the observer extrapolates a vector from the pointer's arm, but that the arm is not extended in the direction of the referent, replicating an earlier report (Herbort & Kunde, 2016).

Misunderstandings in the head-on-shoulder group were smaller than in the naïve group. However, systematic misunderstandings remained. This was expected, because even moving the open eye close to the arm would not completely align the eye of the pointer with the arm–finger line of the observer of the pointing gesture (Fig. 2b). Moreover, as bringing the head as close as possible to the arm or shoulder might have been cumbersome, the distance between the eye and the arm might not have been reduced as much as anatomically possible. Finally, it might be possible that naïve pointers maneuvered their index fingers slightly below the referent in their visual field (Wnuczko & Kennedy, 2011). By contrast, our instruction for pointers in the head-on-shoulder group was to put the index finger on the referent. This instruction might thus have counteracted the aim to reduce the pointers arm elevation.

The informed observer group could have been expected to show essentially no systematic misunderstandings, because the vectors characterizing pointing production and interpretation should have been equated. Nevertheless, the instruction to extrapolate the eye–finger line reduced the misunderstandings only by about 25%. We can only speculate about this finding, but suggest that two factors have limited the effect of the instruction. First, it is conceivable that not all participants followed instructions. Even though this is hard to access in hindsight, a post hoc analysis of questionnaire responses provides some hints in this direction. In a questionnaire issued after the experiment, participants were asked, "how they tried to do the task". Seven out of 12 observers mentioned the eye–finger line spontaneously. Please note, however, that this question was formulated intentionally rather open to reveal unexpected strategies, and participants were not required to mention any geometric rules at all. Hence, not mentioning the eye–finger line did not imply that instructions were not followed. Nevertheless, signed errors (averaged over all trials) of those who mentioned the eye–finger line were considerably smaller than the errors of those who did not mention it, 11.6 cm (sd = 5.1 cm) vs. 20.6 (sd = 4.9 cm), $t(10) = 3.0$, $p = 0.013$, $g = 1.77$. Moreover, those who did not mention the eye–finger line performed virtually identical to those naïve dyads who did not mention the eye–finger line either (average error over all trials 20.6 cm, sd = 5.8).[2] Second, even the apparently compliant informed observers kept overestimating the referent, suggesting that they still processed the pointer's arm. This corresponds to the finding that perceptual judgments pertaining to one stimulus are affected by the presence of other, comparable stimuli (Armstrong & Marks, 1997; Hollingworth, 1910). In sum, it seems that the pointer's outstretched arm and finger are salient visual cues that cannot be easily ignored, even when instructed to do so.

One possible explanation for the persistent misunderstandings in all conditions might be that pointers generally

[2] One observer of a naive dyad and one observer of a head-on-shoulder dyad claimed to base referent estimates on the eye-finger line. The respective naïve dyad performed actually worse than the average naïve dyads. The errors of the respective head-on-shoulder dyad were about half the size of the errors of the remaining head-on-shoulder dyads.

put the index finger above the eye–referent line. This possibility cannot be ruled out, because the pointers' movements were not recorded. However, as participants consistently put the index finger in or slightly below the eye–referent line in previous, similar experiments, this explanation seems unlikely (Herbort & Kunde, 2016; Wnuczko & Kennedy, 2011).

## Experiment 2

Experiment 1 provided further evidence that pointer–observer misunderstandings result from the different geometric rules guiding the production and interpretation of these gestures. Moreover, when the geometric rules of production and interpretation were made more similar, errors were reduced. An inherent constituent of our hypothesis is that the production and interpretation of pointing are both based on visual information. This is obviously true for the interpreter of pointing gestures: All that an observer has is the visual information from inspecting the pointer. However, the matter is not that clear with respect to the pointer herself. We assume that she points by bringing the finger onto the eye–referent line, or in other words, by visually aligning finger and referent from her ego perspective (cf. Fig. 1). However, the pointer has also proprioceptive information about the arm and finger and can thus point with reasonable accuracy without vision during the pointing action itself (Taylor & McCloskey, 1988; Wnuczko & Kennedy, 2011). Interestingly, if it is true that pointers point too high (from the perspective of observers) because they visually co-align finger and referent, then removing visual feedback for the pointer might eventually reduce this systematic "too high" discrepancy. Thus, removing visual feedback might be a means to improve communication by pointing. Testing this was the first aim in the *eyes-closed condition* of Experiment 2.

Second, instead of removing feedback, we tried to change the object that is visually controlled. To this aim, we asked pointers to imagine that they held a stick that extended the vector defined by the arm and could be used to touch the referent (touch-with-imaginary-stick condition, Fig. 2d). This is conceptually similar to holding a laser pointer. As pointers are expected to reduce the discrepancy between the referent and the imagined tip of the stick—and not the index finger—in their visual field, they should elevate the arm less without this instruction.

For comparison, we included a condition in which we asked pointers to point as usual (naïve condition). As in Experiment 1, we focus on signed errors. Absolute errors and the percentage of trials in which the observer correctly identified the pointed at number are reported in Online Resource 1. If the arm elevation of the pointer is determined by a visual control process, signed errors in pointer–observer communication should be smaller in the eyes-closed and touch-with-imaginary-stick condition.

### Methods

#### *Participants*

Thirty-six students of the University of Würzburg signed informed consent and were compensated for participation (27 women, mean age 27 years). According to the handedness scale of the Lateral Preference Inventory (Coren, 1993), all pointers were right-handed. This sample size allowed to detect an effect comparable to that of the head-on-shoulder group in Experiment 1 with a probability of $p > 0.90$ and to counterbalance the order in which instructions were presented.

#### *Stimuli and procedure*

Figure 3b shows the setup of the experiment. The pointer stood in front of the pole used in Experiment 1. The observer was seated at a desk 2 m to the right of the pointer. An LCD below the desk and out of the observer's view showed instructions to the pointer. A laptop on the desk was used to control the LCD and to collect the observer's responses.

At the onset of a trial, the pointer pointed at the referent with the number indicated on the LCD, together with a short reminder of the current pointing instruction. When the observer identified the referent, she entered the corresponding number in the laptop. Then the instruction to lower the arm was shown on the pointer's LCD and a beep was played. After 2 s, the next trial began.

Experiment 2 had a within-participant design. The experiment consisted of nine blocks. One pointing instruction was used in the first three blocks, another in the second three blocks, and a third in the last three blocks. The order of the instructions was counterbalanced. Again, three pointer–pole distances were used. Besides 1 m and 2 m, we presented referents at a distance of 4 m to allow for larger effects. Pointer–pole distances were presented for each instruction in a pseudo-random order. In each block, each referent number between 27 (180 cm above the floor) and 47 (100 cm) was presented once in a pseudo-random order. Thus, each dyad performed 189 trials (9 blocks à 21 trials). The role of the pointer and observer was determined by the toss of a coin.

Observers were instructed to report the referent that was indicated by the pointer. Pointers were told that the experiment tested methods for improving the understandability of the pointing gestures. In the naïve condition, the pointer's instruction was the German translation of: "Try to point naturally, as you would normally". In the eyes-closed condition, the pointer's instruction was: "Look at the number and remember its position. Close your eyes and point to the number with your right arm. Once you hear a beep, lower your arm and open your eyes". The pointer's instruction in the touch-with-imaginary-stick condition was: "When you point, imagine that you're holding a stick that extends your arm straight to the referent".

*Analysis*

Thirty trials (0.9% of all trials) were excluded, because the responses strongly deviated from the response that could be expected from a linear regression of the observer's estimate on the actual referent position for the specific dyad, instruction, and distance (absolute standardized residual >2.5).

## Results

Figure 4b shows the mean signed errors of the observers. Table S2 reports the signed errors, absolute errors and the percentage of correctly identified referents, as well as the results of *t* tests comparing the different instructions. As in Experiment 1, the observers' estimates were on average too high, as revealed by *t* tests conducted for each instruction and distance, all $t(17)$s $\geq 4.1$, all $p$s $\leq 0.001$, all $g$s $\geq 0.97$. Hence, neither instruction eliminated systematic overestimations of the referent.

A repeated-measures ANOVA with within-participant factors distance and instruction revealed that signed errors depended on the distance, $F(2,34) = 64.3$, $p < 0.001$, $\eta_p^2 = 0.791$. There was a main effect of instruction, $F(2,34) = 11.5$, $p = 0.001$, $\eta_p^2 = 0.403$. Instruction and distance interacted, $F(4,68) = 7.1$, $p = 0.003$, $\eta_p^2 = 0.296$. Pairwise *t* tests revealed an increase of errors from 1 to 2 m and from 2 to 4 m for each instruction, all $t(17)$s $\geq 3.4$, all $p$s $\leq 0.004$, all $g$s $\geq 0.795$. Signed errors were smaller in the eyes-closed condition than in the naïve condition at each distance, all $t(17)$s $\geq 3.4$, all $p$s $\leq 0.003$, all $g$s $\geq 0.799$. *t* tests comparing the naïve condition and the touch-with-imaginary-stick condition for each distance revealed a marginal significant effect at the 2 m distance ($t[17] = 1.9$, $p = 0.080$. $g = 0.440$) but no effect at the 1 m ($t[17] = 1.6$, $p = 0.142$. $g = 0.363$) or 4 m distance, $t(17) = 0.6$, $p = 0.542$, $g = 0.147$. Signed errors were lower in the eyes-closed condition than in the touch-with-imaginary-stick condition at 1 m and 2 m distance ($t[17]$s $\geq 2.2$, all $p$s $\leq 0.038$, all $g$s $\geq 0.512$) but not at 4 m distance, $t(17) = 1.4$, $p = 0.189$, $g = 0.340$.

## Discussion

In Experiment 2, we examined whether manipulating visual information available to or used by the pointer affects misunderstandings. With respect to the hypothesis that the pointer's arm elevation is the result of visual control of the pointer's index finger, the experiment yielded mixed results. Pointing with closed eyes reduced the systematic error almost to 50%. This is in line with reports that eye, index finger and referent are no more aligned, once vision is removed (Taylor & McCloskey, 1988; Wnuczko & Kennedy, 2011). It suggests that pointers elevated the arm less when they controlled the gesture proprioceptively rather than visually. This raises the question whether the lower arm elevation is specific to pointing at distant objects or the result of more general control strategies for manual aiming without vision. As experiment 2 is mute with respect to these possibilities, we checked how well participants can control the arm without vision in the used task. Six participants (four female, mean age 25 years, all right-handed) were asked to touch various numbers (27, 29,…, 45, in random order) on the vertical number line with or without vision. With closed eyes, participants touched the number line on average 0.3 cm (sd = 1.26 cm) below the announced target. With open eyes, participants were always on target. These results are in line with a vertical aiming experiment that showed only a very small undershooting bias for upward movements without visual feedback (Elliott et al., 2014). Even when considering that the number line was necessarily closer in this experiment (60 cm), the tiny numerical bias toward lower positions of 0.3 cm can hardly explain the decrease in systematic errors in the eyes-closed condition of Experiment 2 (3.9 cm at 100 cm). This suggests that the reduction of misunderstanding in the eyes-closed condition most likely did not result because proprioceptively controlled arm movements generally undershoot the target in our experiment. Rather, naïve pointers seemed to elevate the arm above the shoulder–referent line because they relied on visual cues provided by the tip of the index finger and the referent. These cues are highly salient for the pointer, but are not used by observers.

However, participants with closed eyes still did not align the arm with the shoulder–referent line in our experiment. By contrast, in previous reports, blindfolded participants pointed with the arm oriented along the shoulder–referent line (Taylor & McCloskey, 1988) or even below this line (Wnuczko & Kennedy, 2011). The reason for the difference is unclear right now. Notable differences between the previous reports and the experiment reported herein concern the spatial layout of the experiment and the communicative situation. Whereas referents were arranged on the floor in the previous experiments, they were mounted vertically in the present ones. Moreover, the present experiment emphasized the communicative intent of the pointing gestures more strongly than the previous reports.

Asking participants to imagine holding a stick did not result in a significant reduction of pointer–observer misunderstandings. As pointing movements are affected when the arm is extended by an actual stick (Wnuczko & Kennedy, 2011), we attribute these findings to pointers' inability to control their movement based on an imagined tool. This is in line with the finding that also movements with real and imagined tools differ (Hermsdörfer, Li, Randerath, Goldenberg, & Johannsen, 2012). One possible reason for this inability might be the high salience of the index finger as opposed to an imagined object. Additionally, it might have been difficult for participants to use visual or proprioceptive information about their arm elevation to determine the end point of the imagined stick.

## General discussion

We reported two experiments that tested the hypothesis that observers systematically misunderstand pointers, because the geometric rules describing the production and interpretation of pointing gestures differ (Herbort & Kunde, 2016; Wnuczko & Kennedy, 2011). That is, pointers insert the index finger between their eyes and the referent, whereas observers extrapolate the arm–finger vector. To this end, we tested how manipulating the geometry of pointing and its interpretation affected pointer–observer misunderstandings. Additionally, we wanted to evaluate whether the instructions we used might help to reduce misunderstanding. In the following, we will first evaluate the theoretical implications, then turn to the more pragmatic side, and finally speculate on the persistence of the misunderstandings to recalibration.

### The geometry of pointing

In previous experiments on naïve pointing, in which the movements of pointers were recorded, it was found that the pointer's eye, index finger, and shoulder were aligned and that observers interpreted pointing gestures by extrapolating the pointer's arm–finger line (e.g., Herbort & Kunde, 2016, Wnuczko & Kennedy, 2011). It was suggested that this causes systematic misunderstanding between pointers and observers. Here, we complemented these results by actively manipulating the geometric rules underlying pointing production and interpretation.

The results of the naïve conditions of both experiments replicated earlier reports and already support the hypothesis (Herbort & Kunde, 2016). Observers generally judged the referent location as higher than implied by the pointer and this bias increased with the distance of the referent. Furthermore, misunderstandings were numerically small—even though statistically significant—for the 1 m distance, suggesting that the vectors defining the production and interpretation of pointing gestures intersect at a position that is relatively close to the 1 m distance These results are in line with the assumption that pointers insert the index finger between the eyes and the referent, but observers extrapolate the arm–finger vector.

To test this hypothesis more rigorously, we altered the geometry of pointing by giving specific instructions to pointers and observers. In Experiment 1, we reduced the discrepancy between the eye–finger and arm–finger vector by instructing pointers to bring the eye as close as possible to their shoulders. As expected, misunderstandings decreased, suggesting that misunderstandings result from the discrepancy between these vectors. In Experiment 1, we also asked observers to extrapolate the eye–finger line, which should point directly at the referent. As predicted, this instruction reduced misunderstandings.

In Experiment 2, we further elaborated on the reason why pointers elevate the arm higher than would be required to orient it directly toward the referent. We tested the hypothesis that participants visually control the tip of their index finger, most likely because it is a prominent visual cue from their point of view. When visual feedback was removed during pointing, misunderstandings were indeed reduced. This suggests that the arm and finger are directed more toward the referent when controlled based on proprioceptive feedback.

Additionally, pointers were instructed to imagine holding a stick that prolongs the vector defined by the arm and allows touching the referent. If participants were able to control the imagined tip of the stick, they should orient their arm directly toward the referent, thus reducing misunderstandings. However, no significant effect of the instruction was found. We attributed the limited effect of this instruction on participants' difficulty to correctly predicting the end point of the imagined stick.

In all conditions in which instructions affected misunderstandings, both the reduction of misunderstandings as compared to the naïve conditions and the magnitude of the remaining misunderstandings increased with distance. This suggests that our instructions altered the orientation of the vectors involved in the production and interpretation of pointing. In turn, it strongly suggests that pointing can indeed be construed as a vector extrapolation process (Bangerter & Oppenheim, 2006; Herbort & Kunde, 2016; Wnuczko & Kennedy, 2011). In sum, the data largely confirmed the initial assumption. However, the effect of our instructions was smaller than expected. Specific reasons for the limited effectiveness of the individual instructions were discussed in the respective sections.

## Application

In the experiments presented herein, we took the approach of altering participants' behavior by means of explicit instructions. In the following section, we will discuss potential limitations of such instructions and the generalizability of the results to more everyday-like situations. In perceptual tasks, providing simple instruction can have a considerable effect on performance (Biederman & Shiffrar, 1987). Likewise, action plans can be easily overridden by explicit strategies (Mazzoni & Krakauer, 2006). The finding that systematic pointer–observer misunderstandings remained in all conditions raises the question to which extent pointer–observer communication could be improved with other instructions or training. One intriguing reason for a limited effect of such interventions might be that gestures are not only used for deixis, but are also beneficial for the gesturer. For example, gesturing facilitates lexical access (Rauscher, Krauss, & Chen, 1996), planning of speech (Alibali et al. 2000), or spatial reasoning (Alibali 2005; Alibali, Spencer, Knox, & Kita, 2011) of the gesturer. Due to such benefits of self-oriented gesturing, pointers might have been reluctant to adapt their pointing gestures fully to our instructions.

Our experiments can only reflect a small subset of the situations, in which people communicate with pointing gestures. This raises the question whether our results can be applied in other situations and more naturalistic settings. In the following, we discuss the spatial layout of the task, the observer's point of view, and the role of speech.

First, in our experiments, we focused on the vertical dimension. Reducing misunderstandings in this dimension is helpful in many situations. In the lifeguard example in "Introduction", the elevation of the pointer's arm indicates the distance between the distressed swimmer and the pointer, which is already a valuable information for the lifeguard. Likewise, the faithful interpretation of the arm elevation is important in many other settings (e.g., pointing at a star in the night sky or an item on a blackboard). Furthermore, we think that our conclusions generalize to situations with other spatial layouts, because other experiments have reported misunderstandings that can also be explained by a mismatch between gesture production and interpretation. When referents were arranged on the floor at different distances from a pointer, observers overestimate the distance of the pointed at object (Wnuczko and Kennedy 2011). When objects were arranged horizontally, observers have been reported to exhibit a bias to the left (Butterworth and Itakura 2000). As these misunderstandings can be explained by the same bias that drives the overestimation of vertical positions in our naïve dyads, these misunderstandings might also be reduced using the instructions evaluated herein.

Second, in our experiments, the observers watched the pointers from the side. Even though this situation is quite common, it can be questioned whether similar misunderstandings would arise when the observer's viewpoint changed. We would argue that our findings generalize to a range of perspectives. Previous research suggested that the relative positions of the pointer and observer affect interpretations only slightly or not at all, at least in the vertical domain. This was even the case when the observer was positioned right next to the pointer or looked over his shoulder (Bangerter & Oppenheim, 2006, Herbort & Kunde, 2016).

Third, unlike many real-world situations, pointing gestures were not accompanied by speech in our experiments. As speech could have resolved the misunderstandings, one could ask whether improving deixis per se is worthwhile at all. It should be noted that one of the very functions of pointing is to complement speech, when verbal deixis is difficult or impossible. Pointing replaces speech, if possible, and ambiguous points require considerably longer verbal descriptions than unambiguous points (Bangerter, 2004). Hence, increasing the spatial accuracy of pointing can be expected to facilitate communication or joint activities, even if speech is permitted. Additionally, we mentioned several situations in "Introduction", in which verbal exchange between the pointer and observer is not possible at all (pointing pre-verbal infants) or very difficult (lifeguard example). In such situations, keeping the geometry of naïve pointing acts and their interpretations in mind might reduce misunderstandings and facilitate communication.

In sum, our findings have a fair chance of generalizing to a variety of situations, including the everyday communication with pointing gestures. Nevertheless, further research is necessary to assess the nature of misunderstandings and the impact of instructions in other, more realistic situations.

## Recalibration of pointing gestures

It is interesting to note that observers systematically misunderstand pointers even though both have an almost lifelong experience with pointing gestures. This is even more surprising because pointing is beneficial (Bangerter, 2004), many people do experience problems when using pointing gestures (70% of 89 participants said so in an unpublished study of our group), and pointers even adapt their gestures to their observers in other ways (Cleret de Langavant et al., 2011; Peeters, Chu, Holler, Hagoort, & Özyürek, 2015). In the following section, we want to speculate why pointing production and interpretation is not recalibrated.

In our opinion, three factors play a role. First, naïve pointing and its interpretation rely on highly salient stimuli. It would be difficult to imagine how the arm elevation is perceived from another perspective, but putting the index finger on the target is easy. Likewise, extrapolating the eye–finger vector would require extrapolating an imagined line instead of the actually present arm and finger. Second, misunderstandings can be often compensated. In many cases, the required acuity for successful communication is low, because possible candidate referents can be easily singled out by speech (e.g., "There is the Eiffel tower"). Such situations might then reinforce naïve pointing and the naïve interpretation of pointing gestures. Third, when pointing fails, people might not attribute misunderstandings to the geometric rules underlying pointing production and interpretation, but to other factors, such as the limited acuity of pointing per se, the different perspectives of the interlocutors, or the inability of the other person. At least anecdotally, people react to failed pointing by using verbal descriptions, by approaching each other's position, or by blaming the other. In sum, in most situations, naïve pointing is an easy and efficient form of deixis. However, when misunderstandings become apparent, people fail to attribute them to the differences of the geometric rules underlying pointing production and interpretation.

**Acknowledgements** We thank Georg Schüssler for technical support, Hanna Mathew and Felicitas Muth for assistance with the data collection.

**Compliance with ethical standards**

**Ethical approval** All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

## References

Alibali, M. W. (2005). Gesture in spatial cognition: expressing, communicating, and thinking about spatial information. *Spatial Cognition and Computation, 5*(4), 307–331. doi:10.1207/s15427633scc0504_2.

Alibali, M. W., Kita, S., & Young, A. J. (2000). Gesture and the process of speech production: we think, therefore we gesture. *Language and Cognitive Processes, 15*(6), 593–613. doi:10.1080/016909600750040571.

Alibali, M. W., Spencer, R. C., Knox, L., & Kita, S. (2011). Spontaneous gestures influence strategy choices in problem solving. *Psychological Science, 22*(9), 1138–1144. doi:10.1177/0956797611417722.

Armstrong, L., & Marks, L. E. (1997). Differential effects of stimulus context on perceived length: implications for the horizontal-vertical illusion. *Perception and Psychophysics, 59*(8), 1200–1213. doi:10.3758/BF03214208.

Bangerter, A. (2004). Using pointing and describing to achieve joint focus of attention in dialogue. *Psychological Science, 15*(6), 415–419. doi:10.1111/j.0956-7976.2004.00694.x.

Bangerter, A., & Oppenheimer, D. M. (2006). Accuracy in detecting referents of pointing gestures unaccompanied by language. *Gesture, 6*(1), 85–102. doi:10.1075/gest.6.1.05ban.

Behne, T., Liszkowski, U., Carpenter, M., & Tomasello, M. (2012). Twelve-month-olds' comprehension and production of pointing. *British Journal of Developmental Psychology, 30*(3), 359–375. doi:10.1111/j.2044-835X.2011.02043.x.

Biederman, I., & Shiffrar, M. M. (1987). Sexing day-old chicks: A case study and expert systems analysis of a difficult perceptual-learning task. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 13*(4), 640–645. doi:10.1037/0278-7393.13.4.640.

Butterworth, G. (2003). Pointing is the royal road to language for babies. In S. Kita (Ed.), *Pointing: Where language, culture, and cognition meet* (pp. 9–33). Mahwah, NJ: Lawrence Erlbaum Associates.

Butterworth, G., & Itakura, S. (2000). How the eyes, head and hand serve definite reference. *British Journal of Developmental Psychology, 18*(1), 25–50. doi:10.1348/026151000165553.

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G. & Moore, C. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development, 63*(4), i–vi, 1–143. doi:10.2307/1166214.

Cleret de Langavant, L., Remy, P., Trinkler, I., McIntyre, J., Dupoux, E., Berthoz, A., & Bachoud-Lévi, A.-C. (2011). Behavioral and neural correlates of communication via pointing. *PLoS ONE, 6*(3), e17719. doi:10.1371/journal.pone.0017719.

Coren, S. (1993). The lateral preference inventory for measurement of handedness, footedness, eyedness, and earedness: norms for young adults. *Bulletin of the Psychonomic Society, 31*(1), 1–3.

Elliott, D., Dutoy, C., Andrew, M., Burkitt, J. J., Grierson, L. E., Lyons, J. L., … Bennett, S. J. (2014). The influence of visual feedback and prior knowledge about feedback on vertical aiming strategies. *Journal of Motor Behavior, 46*(6), 433–443. doi:10.1080/00222895.2014.933767.

Herbort, O., & Kunde, W. (2016). Spatial (mis-)interpretation of pointing gestures to distal referents. *Journal of Experimental Psychology: Human Perception and Performance, 42*(1), 78–89. doi:10.1037/xhp0000126.

Hermsdörfer, J., Li, Y., Randerath, J., Goldenberg, G., & Johannsen, L. (2012). Tool use without a tool: kinematic characteristics of pantomiming as compared to actual use and the effect of brain damage. *Experimental Brain Research, 218*(2), 201–214. doi:10.1007/s00221-012-3021-z.

Hollingworth, H. L. (1910). The central tendency of judgment. *The Journal of Philosophy, Psychology and Scientific Methods, 7*(17), 461–469. doi:10.2307/2012819.

Leavens, D. A., & Hopkins, W. D. (1999). The whole-hand point: the structure and function of pointing from a comparative perspective. *Journal of Comparative Psychology, 113*(4), 417–425. doi:10.1037/0735-7036.113.4.417.

Lücking, A., Pfeiffer, T., & Rieser, H. (2015). Pointing and reference reconsidered. *Journal of Pragmatics, 77*, 56–79. doi:10.1016/j.pragma.2014.12.013.

Mazzoni, P., & Krakauer, J. W. (2006). An implicit plan overrides an explicit strategy during visuomotor adaptation. *The Journal of Neuroscience, 26*(14), 3642–3645. doi:10.1523/JNEUROSCI.5317-05.2006.

Pechmann, T., & Deutsch, W. (1982). The development of verbal and nonverbal devices for reference. *Journal of Experimental Child Psychology, 34*(2), 330–341. doi:10.1016/0022-0965(82)90050-9.

Peeters, D., Chu, M., Holler, J., Hagoort, P., & Özyürek, A. (2015). Electrophysiological and kinematic correlates of communicative intent in the planning and production of pointing gestures and speech. *Journal of Cognitive Neuroscience, 27*(12), 2352–2368. doi:10.1162/jocn_a_00865.

Poulter, D., Jackson, R., Wann, J., & Berry, D. (2005). The effect of learning condition on perceptual anticipation, awareness, and visual search. *Human Movement Science, 24*(3), 345–361. doi:10.1016/j.humov.2005.06.005.

Rauscher, F. H., Krauss, R. M., & Chen, Y. (1996). Gesture, speech, and lexical access: The role of lexical movements in speech production. *Psychological Science, 7*(4), 226–231. doi:10.1111/j.1467-9280.1996.tb00364.x.

Roth, W.-M. (2001). Gestures: their role in teaching and learning. *Review of Educational Research, 71*(3), 365–392. doi:10.3102/00346543071003365.

Sanchez, D. J., & Reber, P. J. (2013). Explicit pre-training instruction does not improve implicit perceptual-motor sequence learning. *Cognition, 126*(3), 341–351. doi:10.1016/j.cognition.2012.11.006.

Taylor, J. L., & McCloskey, D. (1988). Pointing. *Behavioural Brain Research, 29*(1–2), 1–5. doi:10.1016/0166-4328(88)90046-0.

Wnuczko, M., & Kennedy, J. M. (2011). Pivots for pointing: visually-monitored pointing has higher arm elevations than pointing blindfolded. *Journal of Experimental Psychology: Human Perception and Performance, 37*(5), 1485–1491. doi:10.1037/a0024232.

Zhen, W., & Gros-Louis, J. (2015). Caregivers provide more labeling responses to infants' pointing than to infants' object-directed vocalizations. *Journal of Child Language, 42*(3), 538–561. doi:10.1017/S0305000914000221.