Accuracy in Detecting Referents of Pointing Gestures Unaccompanied by Language

Article *in* Gesture · March 2006

DOI: 10.1075/gest.6.1.05ban

Accuracy in detecting referents of pointing gestures unaccompanied by language

Adrian Bangerter and Daniel M. Oppenheimer
Stanford University

It has been claimed that the perceptual accuracy with which people comprehend distal pointing gestures is low. But this claim is at odds with research showing that detection of other indexical signals, e.g., eye gaze, is very accurate. We conducted three experiments to assess people's detection accuracy of targets of distal pointing gestures, using a paradigm adapted from the study of eye gaze. Pairs of people were seated next to each other. One person pointed at targets among 70 points arranged in a horizontal (Experiment 1) or vertical line (Experiment 2). The other person guessed the target. Bias in detection was substantially less than previously shown (approximately 3° in vertical and horizontal conditions), and comparable to levels of accuracy for eye gaze detection. Furthermore, and contrary to previous research, detection accuracy was lower for peripheral targets than for central ones. Experiment 3 replicated Experiments 1 and 2 within one study and further demonstrated that partial occlusion of the pointing arm (from shoulder to elbow) did not adversely affect accuracy.

Keywords: pointing gesture, index, accuracy

How gestures are comprehended is a key issue in research on gesture (Kendon, 1994). Of course, comprehension criteria vary according to the type of gesture studied. For example, in the case of iconic gestures, comprehension is sometimes studied by asking participants to interpret the lexical content of gestures viewed on videotapes without accompanying speech (e.g., Krauss, Morrel-Samuels, & Colasante, 1991). The comprehension criterion for indices (e.g., pointing, eye gaze) is different. The production of an index typically involves extending or pivoting a body part (the hand, eyes, head, lips or whatnot) towards an object of reference or an area in the environment (see the varieties of pointing studied in Kita, 2003). Comprehension of an index is studied by asking participants to identify the object or area or some aspect thereof.

Gesture 6:1 (2006), **85–102.**
ISSN 1568–1475 / E-ISSN 1569–9773

In this respect, the comprehension problem of pointing depends on whether it is proximal or distal (Schmidt, 1999). Proximal pointing occurs when the target is touched by the pointing device. Distal pointing occurs when the target is remotely situated in relation to the pointing device. The comprehension problem for proximal pointing means understanding what aspect of an object is referred to. It can be traced to Wittgenstein's (1971) famous argument that pointing is fundamentally ambiguous without co-occurring language because it is impossible to understand what aspect of an object is being referred to (a part, its color, its shape, and so on). In contrast, the comprehension problem for distal pointing concerns the accuracy with which people are able to spatially locate the target of a pointing gesture (Butterworth, 2003). We call this aspect *detection accuracy*. The issue of detection accuracy has an important psychophysical component, in the sense that it involves visually discriminating where another person is pointing.

How accurate are people in detecting targets of distal pointing gestures? Detection accuracy has been well established for eye gaze. People are very accurate at detecting mutual gaze (i.e., whether someone is looking at them or not; Argyle & Cook, 1976; Cranach & Ellgring, 1973) or the direction of another's gaze (Gale & Monk, 2000). One would expect them to be at least as accurate at detecting referents of pointing gestures. But a recent study (Butterworth & Itakura, 2000) found detection accuracy to be low. Detection accuracy of pointing is an important issue, because it constrains the possible answers to the question of the precise function of pointing in conversation. Indeed, on the basis of their findings of low accuracy, Butterworth and Itakura (2000) claimed that in comprehending pointing, people do not extrapolate a vector from the body part used in the pointing gesture and that pointing functions by directing attention to the visual periphery.

In the present research, we contest this finding with data based on a more precise measurement technique. Thus, the goal of the present research was to determine how accurately the referents of distal pointing gestures can be detected by physically copresent addressees without concurrent linguistic cues. Our data show that detection accuracy is higher than previously estimated, and comparable to accuracy for gaze detection. As a prelude to our research, we discuss methodological aspects of two relevant studies.

Previous research on the accuracy of pointing and eye gaze detection

Butterworth and Itakura (2000, experiments 4 and 5) assessed the accuracy with which a target could be discriminated from among six identical objects on the basis of head and gaze direction and pointing. Participants guessed which of the targets an experimenter was pointing at. In experiment 4, participants (children) sat opposite the experimenter, with the targets located between them. The targets were arranged in a line perpendicular to the experimenter-participant axis (midline) at a distance of 270 cm from the participant. From the participant's point of view, there was a 10° visual angle between targets, and three targets were on each side of the midline. The most accurately identified targets were those directly to each side of the midline and the target on the right-hand periphery. In experiment 5, participants (adults) and experimenter were seated side by side facing six targets, three on each side of their midline. The angular separation of the targets was varied as follows: 4°, 6°, 8°, 10°, 15° and 45°. Results showed that accuracy was greatest for peripheral targets. The authors found that (p. 46) "accuracy to the interior positions requires a minimum separation of 15°". Based on the finding that peripheral targets were better identified than interior ones, they concluded that "comprehension of pointing is unlikely to depend on extrapolating precise linear vectors along the pointing arm" (p. 25), and that "if there is vector extrapolation it is at best approximate and sufficient only to differentiate between widely spaced, identical objects" (p. 48).

The validity of these conclusions may be limited by design problems. One such problem is the use of so few targets. Peripheral targets have only one competitor, whereas interior targets have two. If participants were to guess targets, they would have a higher probability of correctly guessing peripheral targets. The better detection accuracy for peripheral targets might thus be an artifact. The authors mentioned this possibility and tried to correct for it statistically, but it would seem to irrevocably contaminate participants' responses. Furthermore, the use of so few targets limits the precision with which accuracy can be measured. A better method than varying the angular separation between a few targets is using many target units to approximate a continuous metric that allows direct measurement of accuracy.

Gale and Monk (2000) studied eye gaze detection accuracy using such a procedure. In one experiment, one participant (the gazer) looked at different points on a 48-point scale (2° angular separation between each point). Another participant, the estimator, guessed which point the gazer was looking at. Gazer and estimator were seated at a table facing each other with the scale between them. The authors found that the mean error of estimation was 3.8° for horizontal stimulus lines, and 2.6° for vertical stimulus lines. Given this level of accuracy for eye gaze detection, it seems implausible that pointing targets can only be discriminated at a threshold of 15°. If anything, accuracy for pointing detection should be better because the arm and hand provide a longer and thus more precise lever than head-and-eye movement. The setup used by Gale and Monk (2000) seems to be well-suited to measure detection accuracy. Thus, we adapted it to assess the detection accuracy of pointing.

The present experiments

In the three experiments reported here, we sought to measure as precisely as possible the accuracy with which people are able to detect what another person is pointing at, and whether accuracy varies for different target regions (e.g., peripheral regions). Experiment 1 examined accuracy for targets varying in the horizontal dimension. Experiment 2 examined accuracy for targets varying in the vertical dimension. Experiment 3 replicated both dimensions within one experiment. A secondary purpose of Experiment 3 was to determine what kind of information people use to detect targets. We thus introduced a condition where the pointer's arm was hidden by a board between the shoulder and elbow to see if guessers used information from the upper arm to detect targets. If they did, then accuracy could be expected to decrease in this condition.

For all three experiments, two people were seated next to each other facing the targets (Figure 1). One person, the *pointer*, pointed at a target. The other person, the *guesser*, guessed the target identity. (We will refer to pointers as female and guessers as male.) Both participants were naïve in Experiments 1 and 2. In Experiment 3, the pointer was an experimenter and the guesser was a participant. Potential targets were 70 small numbered square boxes arranged adjacently to each other in a line and printed on a sheet of paper. This line was horizontal in Experiment 1 and vertical in Experiment 2. In Experiment 3, both lines were used. A trial consisted of the pointer pointing at a prespecified box and the guesser writing down the number of that box.

Targets were sampled randomly within three areas, a left area (boxes 11–20), a center area (boxes 26–45) and a right area (boxes 51–60) for the horizontal line, and a top area (boxes 11–20), a center area (boxes 26–45) and a bottom area (boxes 51–60) for the vertical line (Figure 1). Pilot experiments revealed that guesser's estimates were often several boxes beyond the actual target, such that if the target was at or near the end of the line, the pointer

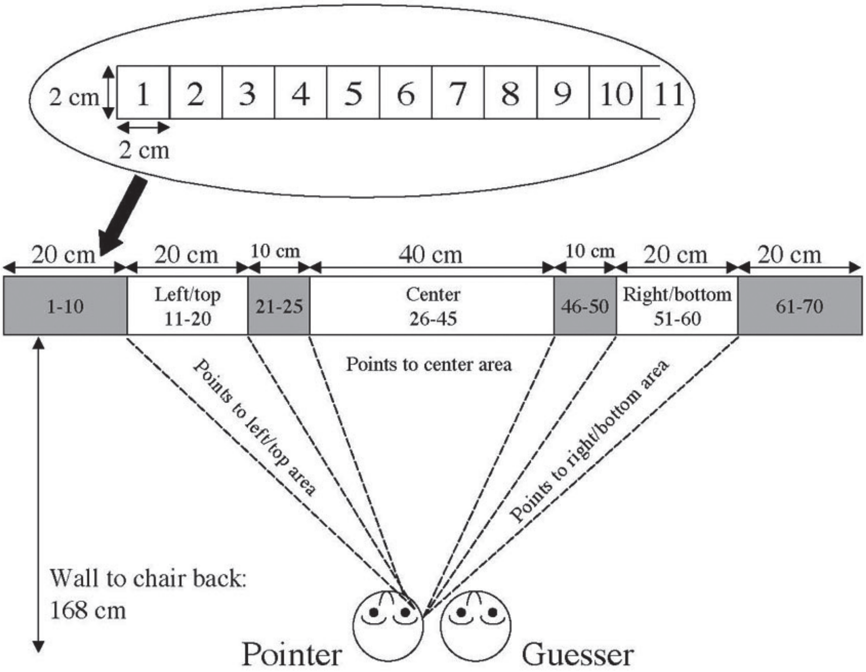

Figure 1. Target areas of the stimulus line, shown in horizontal position. Buffer areas are shaded. Number ranges indicate boxes. Also shown are participant seating positions (below the line) and a blowup view of the boxes as actually seen by the participants (above the line).

sometimes seemed to be pointing at a target beyond the line. This information could be used by guessers to calibrate their estimates. If this is indeed the case, it may explain the higher accuracy for peripheral targets in Butterworth and Itakura's (2000) study. To avoid this problem, the ten boxes at each end of the line (boxes 1–10 and 61–70) served as buffer areas and were never used as targets. Further buffer areas (boxes 21–25 and 46–50) separated the three target areas from each other.

Pointer and guesser were seated next to each other on two chairs facing the wall. The chairs' backs were 168 cm from the wall. In Experiment 1, each chair was on one side of the midpoint of the horizontal line (Figure 1). To refer to symmetrical target areas irrespective of seating, the left area for the left-seated guesser and the right area for the right-seated guesser were designated as *near peripheral areas* and the right area for the left-seated guesser and the left area for the right-seated guesser were designated as *far peripheral areas*.

We computed a measure of bias by subtracting the actual target from the guessed one. Thus, a negative value indicates a guess to the left of or above the actual target, and a positive value a guess to the right of or below the actual target. Note that this is *not* a measure of average *magnitude of error*, because guesses on either side of the actual target compensate for each other when averaged.

In addition to reporting bias, we computed angular measures of the magnitude of error. Bias in cm may not uniformly reflect perceptual bias. Consider the fact that the distance from the pointer's fingertip to the wall varies depending on where the target is. For example, this distance is greater for peripheral targets than for central targets. Thus, a centimeter of bias near the midpoint of the line corresponds to a larger angle of error than a centimeter of bias in the peripheral regions. Reporting bias in cm also makes it difficult to generalize results beyond the setting of the present experiments. However, computing an angle of error requires making assumptions about the origin of the angle. Here we define the origin of the angle as the pointer's shoulder. Angles were calculated as follows: Angle of error = $|\arctan(T/D) - \arctan(G/D)|$, where T is the distance between the target box and the midpoint of the line, G is the distance between the guessed box and the midpoint of the line, and D is the distance from the shoulder of the pointer to the midpoint of the line.

Experiment 1: Horizontal Detection Accuracy

Method

Participants. Six pairs of Stanford University undergraduates participated for credit or pay. All participants had normal or corrected-to-normal vision.

Materials. The horizontal stimulus line was used. It was printed on a large sheet of white paper (152 by 61 cm). This sheet was tacked onto the wall so that the line was horizontal and 112 cm from the floor (approximately shoulder height for a seated person). Two chairs were placed next to each other facing the wall at a distance of 168 cm from it. They were on each side of the midpoint of the line.

Twelve stimulus lists were constructed. Each list specified 30 targets to point at. Targets were chosen randomly from boxes in the target areas of the line (Figure 1). Ten targets were chosen from each target area (left, center and right areas). Other boxes were buffer areas and were never pointed to. The order of these 30 targets was randomized. These lists were printed in a single column on the left side of a sheet of paper (stimulus and answer sheet). On the right side, another column of thirty numbered boxes was printed (answer column).

Procedure. Participants seated themselves in one of the chairs for the duration of the experiment. The person on the left started as the pointer. The person on the right started as the guesser. Each person was given a clipboard with four stimulus and answer sheets and a pencil. The pointer pointed to the box on the wall which corresponded to the first target listed on the stimulus sheet. She held her arm still until the guesser had determined which box she was pointing to, written his answer down in the answer column, and said "okay." She then lowered her arm, crossed out the first number on the stimulus column, and pointed to the box corresponding to the next target on the list. This procedure was repeated until thirty trials were completed.

The participants then switched roles and completed another block of thirty trials. They then flipped over their stimulus and answer sheets and repeated the procedure three times. Thus, in total, each person guessed 120 times in 4 blocks, 40 times for each region of the stimulus line.

Participants were instructed not to speak or otherwise communicate other than saying "okay" to allow the pointer to lower her arm. They were instructed to remain seated and not to move their chairs. Pointers were instructed to always point with their arm closest to the guesser, i.e., pointers seated on the left pointed with their right arm and pointers seated on the right with their left arm. They were allowed to rest their pointing arm if necessary.

Results and Discussion

The mean bias for each participant was computed for each area. Data from one block of one participant had to be discarded because the participant lost track of guesses. Data was analyzed treating each participant as an independent measure, i.e., pair interdependencies were ignored. This is a reasonable assumption because participants did not really interact (i.e., the responses of the guesser did not influence the behavior of the pointer).

Irrespective of seating and target area, guessers' estimates were biased towards the side that the pointer was sitting on (Figure 2). They varied between approximately 4 and 12 cm depending on target area. For statistical analyses, we rendered the data of the right-seated guessers comparable to those of the left-seated guessers by inverting the signs of their individual means and inverting data for left and right target areas. A 2 (guesser seat) by 3 (target area) mixed-model ANOVA revealed a significant main effect of target area, $F(2,20) = 14.1$, $p < .001$. Post hoc tests (Tukey HSD) revealed that bias for the far peripheral

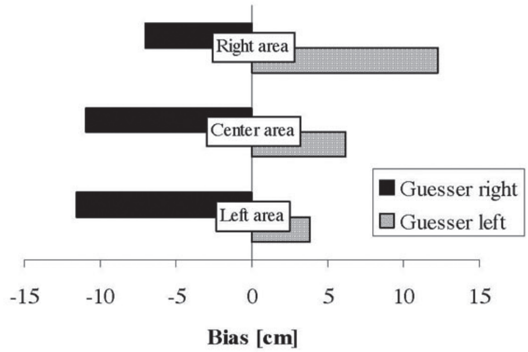

Figure 2. Mean horizontal detection bias in cm as a function of where the guesser was seated and target area in Experiment 1.

area was higher than for the center area, which in turn was higher than for the near peripheral area. The overall mean detection bias (collapsed over target areas) corresponded to an angle of 3.5°.

Further analysis was suggested by the fact that the near peripheral area and the half of the center area closest to it (the *near center area*; boxes 26–35 for left-seated guessers and boxes 3645 for right-seated guessers) are symmetrical with respect to the axis constituted by the guesser's direction of gaze when he looks straight ahead. Thus, the only difference between these two areas is the angle at which the pointer's arm is perceived. We compared bias for three recoded areas, the near center area (as described above), the *far center area* (boxes 36–45 for left-seated guessers and boxes 26–35 for right-seated guessers) and the near periphery. Data were rendered comparable irrespective of seating as before. A 2 (guesser seat) by 3 (target area) mixed-model ANOVA revealed a significant main effect of target area, $F(2,20) = 6.6$, $p < .01$. Post hoc tests (Tukey HSD) revealed that bias was lower for the near peripheral area (*M*: 5.4 cm) than for the far center area (*M*: 9.2 cm) and for the near center area (*M*: 8 cm, but $p < .10$). So comparison of two target areas at equal angular distance from the straight-ahead direction of gaze revealed a difference in bias which can only be attributed to the different angles at which the pointing arm is seen.

Results revealed that detection accuracy of pointing targets is substantially higher than previous accounts would suggest. Moreover, previous reports (Butterworth & Itakura, 2000) of higher accuracy for peripheral targets are contradicted by the present results. Here, bias is greatest when pointing to targets in the far peripheral area. Previous findings are possibly artifactual; the small number of targets used and their placement may have helped guessers to calibrate their estimates. The main factor influencing bias seems to be the orientation of the pointer's arm. Bias is highest when it is pointing away from the guesser's line of sight, and lowest when it crosses the guesser's line of sight. In Experiment 2, we sought to determine guesser bias in the vertical dimension.

Experiment 2: Vertical Detection Accuracy

Method

Participants. Six pairs of Stanford University undergraduates participated for credit or pay. All participants had normal or corrected-to-normal vision.
Materials. The vertical stimulus line was printed on a larger sheet of white paper (61 by 152 cm). This sheet was tacked onto the wall so that the midpoint of the line (between boxes 35 and 36) was 112 cm from the floor (approximately shoulder height for a seated person). Two chairs were placed next to each other facing the wall with their backs at a distance of 168 cm from it. They were on each side of the line. The same twelve stimulus lists as in Experiment 1 were used.
Procedure. The procedure was the same as in Experiment 1.

Results and Discussion

Mean detection bias in cm was computed for each point as in Experiment 1. Data from one block of one participant had to be discarded because the participant lost track of guesses. Mean bias was 6.2 cm above the target, irrespective of seating position and target area (Figure 3). A 2 (guesser seat) by 3 (target area) mixed-model ANOVA revealed no significant main effects or interactions,

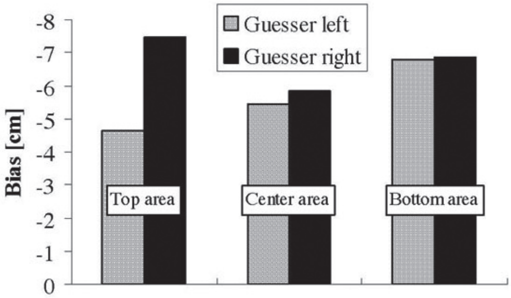

Figure 3. Mean vertical detection bias in cm as a function of where the guesser was seated and target area in Experiment 2.

guesser seat: $F(1,10) = .1$, target area: $F(2,20) = 1.2$, interaction: $F(2,20) = 1.7$, all *ns*. The overall mean detection bias corresponded to an angle of 2.5°.

Bias thus differed from the horizontal dimension. The question also arises whether bias is significantly smaller than for the horizontal dimension. In the case of eye gaze, horizontal error was found to be greater than vertical error (Gale & Monk, 2000, Experiment 1). Here, the two dimensions were assessed in separate experiments, so we could not compare them directly. In Experiment 3, we combined them into a within-subject variable.

Experiment 3: Detection Accuracy Under Partial Occlusion

In Experiment 3, we sought to replicate the findings of Experiments 1 and 2 in a single experiment. We especially wanted to compare accuracy in the horizontal and vertical dimension.

Furthermore, the role of the pointer was played by an experimenter to control one source of variation. Finally, we sought to assess what information people use in guessing targets by occluding the pointer's entire body up to the forearm. This notably excludes cues such as eye gaze and head orientation, as well as directional information from the upper arm. Butterworth and Itakura's (2000) study investigated detection accuracy for different combinations of cues (e.g., head orientation plus pointing or head orientation and eye gaze), but none of their conditions studied pointing detection accuracy without other concurrent cues. Our occlusion manipulation was designed to tease apart these factors.

Method

Participants. Twelve Stanford University undergraduates participated for credit or pay. All participants had normal or corrected-to-normal vision. The data from one participant was discarded when debriefing revealed that he had mentally "corrected" his estimates before writing them down.

Materials. The horizontal and vertical stimulus lines from Experiments 1 and 2 were used under identical conditions as for these experiments. Horizontal and vertical line setups were installed in different areas of the same lab room.

Twelve stimulus lists were constructed using the same method as in Experiment 1. Each list specified 45 targets, 15 from each of the three different areas. Order of occurrence of areas and specific targets within areas was randomized. A white foam board (76 by 159 cm) was placed vertically between seats in the hidden condition. The board obscured the participant's view of the experimenter except for his pointing arm, which was visible from the elbow to the fingertip when he pointed straight ahead.

Procedure. The order in which the different conditions (horizontal — visible, horizontal — hidden, vertical-visible, vertical-hidden) were completed was pre-determined for each participant and counterbalanced. Participants (we will refer to participants as female) were given a clipboard with four answer sheets and a pencil. They always sat on the right chair and the experimenter (the first author) always sat on the left chair. The experimenter always pointed with his right arm.

The experimenter randomly chose four different stimulus lists before each session. He started the first trial by pointing to the box which corresponded to the first number listed on the first stimulus sheet. He held the point until the participant had guessed which box he was pointing to, written her answer down in the answer column, and said "okay." He then lowered his arm, crossed out the first number on the stimulus column, and pointed to the next target on the list. This procedure was repeated until all forty-five boxes had been pointed to. Participants were instructed to remain seated with their backs against the chair back, not to move their chairs, and not to speak or otherwise communicate other than by saying "okay" to allow the experimenter to lower his arm.

Depending on which condition was next, the experimenter then either placed or removed the board, or instructed the participant to move to the other setup, or both. They then completed the pointing and guessing for that condition and the last two conditions. Thus, each person guessed 180 times in 4 blocks, 45 times in each condition.

Results and Discussion

Mean detection bias in cm was computed for each point as in Experiment 1 (Figures 4 and 5). Again, bias was towards the side of the pointer's arm away from the guesser (to the left) for the horizontal line and above the target for the vertical line. A 2 (line orientation) by 2 (visibility) by 3 (target area) repeated-measures ANOVA revealed a significant main effect of target area, $F(2,20) = 5.3$, $p < .05$, a significant interaction between line orientation and target area, $F(2,20) = 20$, $p < .001$, and a significant interaction between visibility and target area, $F(2,20) = 3.7$, $p < .05$. Post hoc tests (Tukey HSD) on the line orientation by target area interaction showed that for the horizontal line, bias was higher for the left target area (M: 6.2 cm) than for the center area (M: 3.5 cm), and higher for the center area than the right area (M: 0.9 cm), whereas for the vertical line,

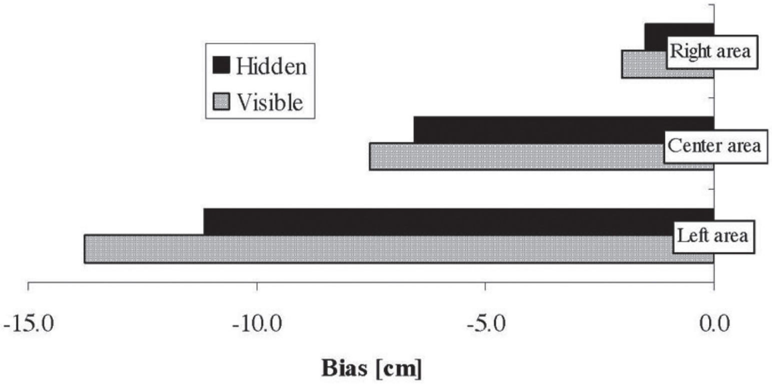

Figure 4. Mean horizontal detection bias in cm as a function of visibility and target area in Experiment 3.

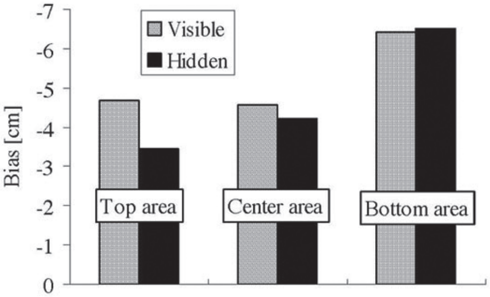

Figure 5. Mean vertical detection bias in cm as a function of visibility and target area in Experiment 3.

there were no differences between areas. This confirmed the results found in Experiments 1 and 2. Post hoc tests on the visibility by target area interaction showed that bias was higher in the visible condition for the left and top target areas. The overall mean detection bias corresponded to an angle of 2.5° for the horizontal line and 1.8° for the vertical line. As shown by the absence of a main effect for line orientation, this difference is not significant.

Occluding the pointer's arm above the elbow had no adverse effects on detection accuracy. In fact, accuracy was higher for the left and top target areas. We have no explanation for this unexpected effect beyond the fact that people do not seem to need visual information from the upper arm to accurately detect targets. In retrospect, it seems possible that the board may actually have facilitated detection by creating a clear marking point by which to gauge the lateral movement of the pointer's arm. The portion of visible arm could potentially serve as a measure of the angular position of the arm, and thus of the target.

This explanation seems implausible for two reasons. First, it was not possible to calibrate this metric at any point during the experiment, because guessers never received feedback as to how accurate their guesses were. Second, the experimenter lowered his arm to his lap between trials. Thus, any measure of the portion of visible arm would have to compare the current position with the preceding one from memory. This would seem to be an extremely error-prone affair. In any case, it is difficult to address this problem, because any occlusion of the pointing arm must necessarily create some kind of a marker, even a setup where the occluding board would move with the arm.

General Discussion

In three experiments, we assessed the accuracy of the detection of pointing referents. Our results show that accuracy is substantially higher than previously estimated (Butterworth & Itakura, 2000). Previous results about the inaccuracy of pointing detection are possibly artifactual and should be qualified accordingly.

Perhaps the most intriguing finding is that estimation of pointing targets in the horizontal dimension was systematically biased to the side of the pointer's arm away from the guesser. Indeed, if guessers had estimated the target by simply aligning their eye with the tip of the pointer's finger, they would have shown a larger bias in the same direction. Such a process may operate automatically and interfere with the participants' (consciously made) estimates in the present experiments. Although this is only a speculation, a number of factors make it plausible.

First, consider the fact that people point by aligning the tip of their pointing finger with their dominant eye. This leads to a *pointer bias*, i.e., the actual target of a pointing gesture is predictably different than the target pointers believe they are pointing at. This was revealed by an informal study we conducted. Participants ($n = 7$, including the authors) were seated in front of the midpoint of the horizontal line used in Experiments 1 and 3. They pointed at 45 different targets (15 from each of the 3 target areas) with each arm. They held a laser pointer in their pointing hand, and aligned a notch at its tip with the target. They then held the point and turned it on. The experimenter noted the actual target indicated by the laser beam. The dominant eye of each participant was determined. The mean bias was 1.2 cm for the arm closest to the dominant eye and 3.3 cm for the arm further from the dominant eye (Figure 6; all participants but one were right-eye dominant; that participant's data was

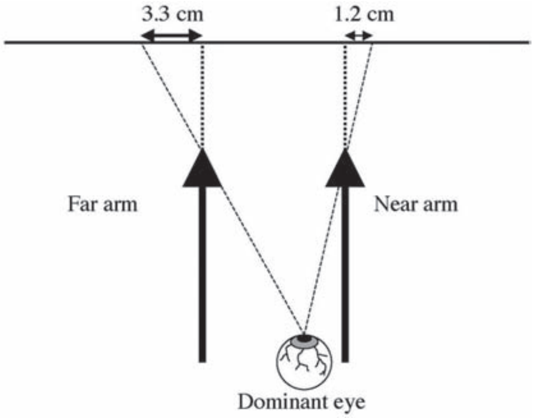

Figure 6. Mean horizontal bias of pointers as a function of ocular dominance.

converted). A paired-sample *t*-test showed that the absolute value of the bias for the arm far from the dominant eye was larger than the bias for the arm closer to it, $t(6) = 3.9$, $p < .01$. Thus, this study shows that pointers align the tip of their pointing finger with their dominant eye when pointing, thereby creating a systematic bias. Guessers may also be subject to this bias when attempting to estimate where another person is pointing. The bias may operate automatically. Studies using a Stroop paradigm (e.g., Langton & Bruce, 2000) have shown that people have difficulty ignoring pointing gestures (e.g., a person pointing up or down) that are incongruent with verbal instructions ("up" or "down"), suggesting that pointing comprehension is at least partly automatic. Furthermore, when people comprehend pointing gestures in everyday interaction, they have little time to consciously estimate the exact location that is being pointed to. Our setup required participants to consciously perform an estimation task that they are seldom confronted with, making interference from automatic processes plausible.

Second, a similar phenomenon has been reported in eye gaze detection (Cline, 1967; Gibson & Pick, 1963), where the estimated direction of eye gaze is affected by the direction of the head. For example, people tended to estimate gaze that was actually directed straight towards them as directed to their right when the gazer's head was angled to their right. Presumably, participants in these studies computed the direction of gaze by integrating information from both the head angle and eye gaze. Likewise, in the present experiments, participants' estimates may have been affected by automatically generated information from the alignment of their dominant eye with the pointer's finger. In other words, we have shown, contrary to Butterworth and Itakura (2000), that people are quite capable of extrapolating a vector from a pointing device to a target, but the possibility remains that they may be perceptually biased by the automatically generated information from the alignment of their eye with the tip of the pointing device. Testing this possibility seems feasible and could be a focus of future studies.

The present research used methodological innovations from previous research (Gale & Monk, 2000) to better estimate detection accuracy by excluding a number of potential designrelated artifacts. We systematically employed a situation where pointer and guesser were seated side-by-side. We chose this alignment because it seems to correspond to a typical situation in everyday conversation, and also to ensure comparability of results with those of Butterworth and Itakura (2000). But the accuracy results reported may vary for different pointer-guesser configurations. This is a potential limitation to the generalizeability of the present research, and we think it is worth pursuing this issue in further studies, using other configurations (e.g., pointer and guesser facing each other or an over-the-shoulder perspective).

The present research has established with precision the accuracy of pointing comprehension. What implications follow for the understanding of pointing gestures in conversation? Although we do agree with Butterworth & Itakura (2000) and others (Bangerter, 2004; Schmauks, 1991) that an important function of pointing in conversation is creating a joint focus of attention, on the basis of our results, we disagree that detection accuracy of pointing is perceptually inaccurate in principle. Arguments about the accuracy or inaccuracy of gesture cannot be used to support an attention-shifting function of gesture. Conversation is a coordinated activity (Clark, 1996), and linguistic and gestural aspects are highly integrated components of that activity (Bavelas & Chovil, 2000; Engle, 1998; McNeill, 1985). That means that participants in conversation will rely differently, and flexibly, on gestural or verbal means to design reference according to their relative costs (Bangerter, 2004). Costs include the collaborative effort necessary to ground reference (Clark & Wilkes-Gibbs, 1986), as well as perceived costs of repairing that message if misunderstood (Clark, 1994; Schegloff, Jefferson, & Sacks, 1977). In other words, participants in conversation are unlikely to rely on pointing gestures to the limit of their accuracy, because they may perceive the risk of misunderstanding (and the necessity of subsequent repair) to be too high. Of course, the question remains of how accurate participants perceive gesture comprehension to be.

Nevertheless, it is important to have a clear understanding of the baseline accuracy of pointing detection. Several recent studies have examined issues where accuracy is relevant. For example, we note an increasing interest in the study of gestures in mediated communication (e.g., Barnard, May, & Salber, 1996; Fussell et al., 2004). Barnard et al. (1996) examined how people comprehended recorded instructions (a video recording of a person talking and/or pointing) to move objects on a computer screen. They compared different views of the instructor (face-to-face, face-to-face reversed, and over-the-shoulder) and instructions given by explicit verbal means and pointing (e.g., "put the big diamond left of the square") or only by deixis ("put this left of the square"). Comprehension was more difficult in deictic conditions, as evidenced by higher error rates, and completion times, and more replays of the instructions. In a video-mediated situation like this, it is difficult to separate the intrinsic ambiguity of gesture comprehension from the possible additional ambiguity resulting from the perception of the gesture over video (which distorts stereoscopic cues). It may be the case that detection accuracy of pointing is not worse over a videolink than face-to-face, as Gale and Monk (2000) have found for eye gaze. Again, this underscores the importance of stringently assessing detection accuracy of pointing in face-to-face conditions.

Note

Adrian Bangerter is now at the University of Neuchâtel, Switzerland. Danny Oppenheimer is now at the Department of Psychology, Princeton University. Adrian Bangerter was supported by an Advanced Researchers Fellowship (No. 8210–061238) from the Swiss National Science Foundation. Danny Oppenheimer was supported by a National Science Foundation Graduate Research Fellowship. Support was also received from the Office of Naval Research (Grant N000140010660 to Herbert H. Clark and Stanley Peters). We thank Paul P. Maglio, Teenie Matlock, members of the User Sciences group at IBM Almaden Research Center, and members of SLUGs at the Psychology Department, Stanford University, for their valuable comments on this work. We especially thank Herb Clark for support and comments.

References

Argyle, Michael & Mark Cook (1976). *Gaze and mutual gaze*. Cambridge: Cambridge University Press.

Bangerter, Adrian (2004). Using pointing and describing to achieve joint focus of attention in dialogue. *Psychological Science, 15*, 415–419.

Bavelas, Janet B. & Nicole Chovil (2000). Visible acts of meaning. An integrated message model of language in face-to-face dialogue. *Journal of Language and Social Psychology, 19*, 163–194.

Barnard, Philip, Jon May, & Daniel Salber (1996). Deixis and points of view in media spaces: an empirical gesture. *Behaviour and Information Technology, 15*, 37–50.

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

Butterworth, George & Shoji Itakura (2000). How the eyes, head and hand serve definite reference. *British Journal of Developmental Psychology, 18*, 25–50.

Clark, Herbert H. (1996). *Using language*. Cambridge: Cambridge University Press.

Clark, Herbert H. (1994). Managing problems in speaking. *Speech Communication*, *15*, 243–250.

Clark, Herbert H. & Deanna Wilkes-Gibbs (1986). Referring as a collaborative process. *Cognition, 22*, 1–39.

Cline, Marvin G. (1967). The perception of where a person is looking. *American Journal of Psychology, 80*, 41–50.

Cranach, Mario von & J. Heinrich Ellgring (1973). The perception of looking behavior. In Mario von Cranach & Ian Vine (Eds.), *Social communication and movement*. London: Academic Press.

Engle, Randi A. (1998). Not channels but composite signals: Speech, gesture, diagrams and object demonstrations are integrated in multimodal explanations. In Morton A. Gernsbacher & Sharon J. Derry (Eds.), *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society* (pp. 321–326). Mahwah, NJ: Erlbaum.

Fussell, Susan R., Leslie D. Setlock, Jie Yang, Jiazhi Ou, Elizabeth M. Mauer, & Adam Kramer (2004). Gestures over video streams to support remote collaboration on physical tasks. *Human-Computer Interaction, 19*, 273–309.

Gale, Caroline & Andrew F. Monk (2000). Where am I looking? The accuracy of video-mediated gaze awareness. *Perception and Psychophysics, 62*, 586–595.

Gibson, James J. & Anne G. Pick (1963). Perception of another person's looking behavior. *American Journal of Psychology, 76*, 386–394.

Kendon, Adam (1994). Do gestures communicate? A review. *Research on Language and Social Interaction, 27*, 175–200.

Kita, Sotaro (Ed.) (2003). *Pointing: Where language, culture and cognition meet*. Mahwah, NJ: Erlbaum.

Krauss, Robert M., Palmer Morrel-Samuels, & Christina Colasante (1991). Do conversational hand gestures communicate? *Journal of Personality and Social Psychology, 61*, 743–754.

Langton, Stephen R. H. & Vicki Bruce (2000). You *must* see the point: Processing of cues to the direction of social attention. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 747–757.

McNeill, David T. (1985). So you think gestures are non-verbal? *Psychological Review, 92*, 350–371.

Schegloff, Emanuel A., Gail Jefferson, & Harvey Sacks. (1977). The preference for self-correction in the organization of repair in conversation. *Language*, *53*, 361–382.

Schmauks, Dagmar (1991). *Deixis in der Mensch-Maschine-Interaktion*. [Deixis in human-computer interaction.] Tübingen: Max Niemeyer.

Schmidt, Chris L. (1999). Adult understanding of spontaneous attention-directing events: What does gesture contribute? *Ecological Psychology, 11,* 139–174.
Wittgenstein, Ludwig (1971). *Philosophical investigations*. New York: MacMillan.

Authors' address

Adrian Bangerter
Institut de Psychologie du Travail et des Organisations
University of Neuchâtel
Fbg. de l'Hôpital 106
2000 Neuchâtel
Switzerland
+41 (0)32 718 13 18

E-mail: adrian.bangerter@unine.ch

About the authors

Adrian Bangerter is a social-organizational psychologist with a variety of interests ranging from interpersonal communication to the diffusion of ideas and resistance to innovation. A common perspective in much of his work is an emphasis on interpersonal, social and cultural aspects of psychological processes.

Danny Oppenheimer is an assistant professor of psychology and public affairs at Princeton University. His work focuses on judgment and decision making, reasoning, and metacognition.