Supplemental Text 1: Experimental Test of the Bayesian Gesture Interpretation Model

The present experiment was conducted to check the assumptions underlying the Bayesian version of the gesture interpretation models, with respect to the parameters and the predicted relationship between different variables. The model is based on the assumption that observer's fuse the result of a linear extrapolation with their a priori assumptions of referent positions, according to the certainty associated with each information source. Here we aimed at identifying A) the a priori assumptions of observers about the referent positions (i.e. the prior); B) whether uncertainty in line extrapolation increases reliance on the prior; C) whether a participant's prior is correlated with her interpretations; D) whether the uncertainty associated with a participant's prior is correlated with reliance on linear extrapolation. While A) tests specific assumption on the model parameters, B-D) are predictions derived from Bayesian theory that are tested to assert the adequacy of the Bayesian approach in the current context.

Three types of trials which differed in the pointer's visibility were interspersed. In one third of trials, participants gave referent estimates to stimuli in which no pointer was shown at all (no-pointer condition). As no image of a pointer was actually shown in these trials, participants needed to base their guess on a priori assumptions on the referent positions. We used this condition to infer the center and the distribution of the prior. To assure compliance, participants were told that the pointer was briefly presented in a stream of flickering random pattern and that they should guess the referent even if they did not consciously perceive the pointer. In one third of trials, a pointer was shown briefly within the stream of random pattern (pointer-masked condition). This was done to strengthen participants' beliefs that an image of a pointer was also presented in the no-pointer condition. Additionally, this condition provided degraded information on the pointer and allowed to test whether uncertainty in line extrapolation increased reliance on the prior. The remaining third of trials required participants to respond while seeing the pointer.

To check whether referent estimates in the current experiment are comparable to those of Experiment 1 and 2a, the pointer-visible condition of the current experiment will be compared to Experiment 2a. Please note that Experiment 2a was identical to the interpretation task of Experiment 1 with the exception that the presented pointers had the same natural combinations of head and arm orientations that were used in the present experiment. Referent estimates of Experiments 1 and 2a were almost identical. Four hypotheses pertaining to the Bayesian version of the interpretation model were tested.

A) It was checked whether the observer's a priori assumption is approximately normal distributed and centered on shoulder height, as assumed in the Bayesian interpretation models.

B) As the processing of the gesture of the masked pointer can be assumed to be less accurate than that of the clearly visible one, Bayesian theory predicts that participants' referent estimates are stronger biased toward their a priori assumption in the pointer-masked condition than in the pointer-visible condition.

C) According to the Bayesian version of the interpretation model, participants form a weighted mean of their a priori assumption and the result of linear extrapolation. This assumption would be supported by a positive correlation between the mean of participants' responses in the no-pointer condition and the average position of their referent estimates in the other conditions.

D) According to the Bayesian version of the interpretation model, participants weigh the result of the linear extrapolation the stronger, the higher the variance of their a priori assumption. This assumption would be supported if the variability of the responses in the no-pointer condition was negatively correlated with the deviance from linear extrapolation.

Methods

Participants

Sixteen participants[1] (15 right-handed, 1 left-handed according to the handedness scale of the Lateral Preference Inventory; Coren, 1993) recruited from the Würzburg area participated as a course requirement or for payment after signing informed consent students (13 women, mean age 28 years).

Design, Procedure, and Data Analysis

We used stimuli comparable to those of Experiment 1, with the difference that the pointer's head orientation corresponded to his arm orientation in a natural way. A detailed description of the stimuli can be found in the method section for Experiment 2a, in which the same stimuli were used (an example is shown in supplemental Figure S1). Additionally, for the pointer-masked and no-pointer-condition the scenes with the pole at 1 m, 2 m, and 3 m were rendered without the pointer. For the masks, 72 different random pattern were generated by splitting the scenes without the pointer in 8x8 pixel segments and reordering these segments randomly. The stimuli were presented on CRT - Monitors (1280 x 1024 pixel, 75 Hz, size of scene 32 cm x 20 cm).

Figure 1a-c show the trial procedures. Each trial began with the presentation of a blank beige screen for 1000 ms. In the no-pointer condition, six random pattern were then presented successively for 40 ms each, followed by a scene without the pointer, which stayed on until the participant marked his or her referent estimate with the mouse. In the

[1] We wanted to have at least as many participants as in Experiment 2a-d, but preferred to retain the additional power that emerged because more participants were eventually available. Given the variability of the mean of the prior in the current experiment, the sample size allowed to detect a deviation of the prior from the value used in the model by 5 cm or 10 pixel (i.e. that the prior would be outside the arm as presented in our stimuli) with a power of $\beta = .9619$.

pointer-masked condition, three random pattern were presented for 40 ms each, followed by a presentation of a scene with a pointer for 120 ms, followed by three more random pattern for 40 ms each, followed by a scene without the pointer, which stayed on until a response was given. In the pointer-visible condition, three different random pattern were presented for 40 ms each, followed by a scene including the pointer that stayed on screen until a response was given. After the response was given, the next trial was started.

Before the experiment, ten randomly selected training trials were administered. Then eight blocks of 45 trials were presented. These included one repetition of each combination of the three distances (1 m, 2 m, 3m) and five arm orientations (-20° , -10°, ... , 20°) for the pointer-visible and pointer-masked condition and 5 repetitions of each of the three distances for the no-pointer condition. Stimulus order was randomized. In half of the blocks, the initial cursor position was at the foot of the pole, in the other it was at the top of the pole. Blocks were presented in random order. Altogether, 360 trials (excluding training trials) were administered, separated by self-paced breaks between the blocks. Trials were excluded if they differed by more than 2 standard deviations from the respective conditions mean (2%).

Results

First, the pointer-visible condition of the present experiment was compared to Experiment 2a (using the same outlier criterion as used here) by means of a split-plot ANoVA with within-subject factors arm orientation and distance and between subject factor experiment. On average, referent estimates were about 4 cm (or 8 pixel) lower in the present experiment than in Experiment 2a, $F(1,20) = 6.2$, $p = .021$, $\eta^2_p = .234$. No significant interaction of the within-subject factors with the factor experiment was found (all Fs ≤ 2.7, all ps $\geq .101$, all η^2_ps $\leq .120$). Thus, despite procedural differences in both experiments, the responses of participants in the current experiment are comparable to those of Experiment 2a and thus also to those of Experiment 1.

A) Does the position and distribution of the referent estimates in the no-pointer condition match the assumptions of the model? In the no-pointer condition, participants guessed that the referent is located on a height of 147 cm (sd = 5cm). This empirical value is very close to the shoulder height of 145 cm used in the Bayesian version of the interpretation models and does not significantly differ from it, $t(15) = 1.8, p = .093, g = 0.449$. Figure 1d shows the distribution of the participant-wise z-transformed referent estimates in the no-pointer condition. Participants' responses were at least approximately normal distributed. Kolmogorow-Smirnow-Lilliefors-Tests conducted on the data of individual participants revealed no significant deviation from normal distribution for 13 of 16 participants (all ps $\geq .160$), for three participants, the test reached significance (all ps $\leq .034$). In sum, these results show that participant's a priori assumption on the referent position can be adequately described as a normal distribution approximately centered on the height of the pointer's shoulder in most cases.

B) Do responses in the pointer-masked condition deviate stronger from linear extrapolation than responses in the pointer-visible condition? Figure 1e-f show the referent estimates in both conditions. A repeated-measures ANOVA with factors arm orientation, distance, and pointer visibility (pointer-masked vs. pointer visible) was conducted. Not surprisingly, the ANOVA revealed the effects of arm orientation, $F(4,60)= 441.5, p < .001, \eta^2_p = .967$, and the interaction of arm orientation and distance, $F(8,120) = 66.8, p < .001, \eta^2_p = .817$. Additionally, referent estimates were higher in the pointer-masked (m = 150 cm) condition than the pointer-visible condition, m = 145 cm, $F(1,15) = 7.7, p < .014, \eta^2_p = .338$. A significant interaction of distance and pointer visibility revealed that the height of referent estimates increased with distance in the pointer-masked condition but not in the pointer-visible condition, $F(2,30) = 17.9, p < .001, \eta^2_p = .544$. Most importantly, the three-way interaction was significant, $F(8,120) = 4.7, p = .002, \eta^2_p = .238$. No other effect reached significance, all Fs ≤ 3.8, all ps $\geq .067$, all η^2_ps $\leq .184$. The three-way interaction indicates that the referent estimates diverged stronger with distance and are thus were less biased toward the mean in the pointer-visible condition than in the pointer-masked condition. This interpretation is confirmed by a follow up repeated-measures ANoVA with factors arm orientation (-20° vs. 20°), distance (1 m vs. 3m), and pointer visibility (pointer-masked vs. pointer visible), which also revealed a significant three-way interaction, $F(1,15) = 10.3$, $p = .006$, $\eta^2_p = .408$. In sum, when the pointer is clearly perceivable, referent estimates deviate less from what would be expected from linear extrapolation than when the pointer is only briefly shown. This corresponds with the Bayesian interpretation, that participants rely the more on a priori information and the less on perceptual information, the less reliable the visual stimulus is.

C) Are the average positions indicated in the no-pointer condition related to the overall height of the referent estimates in the pointer-masked and pointer-visible condition? Figures 1g-h show that participants' average estimates in the no-pointer condition (i.e. the prior) were correlated positively with those in the other conditions. Significant positive correlation were found when the pointer was difficult to perceive ($p < .05$, marked with an asterisk in Fig. 1g-h), positive but non-significant correlation were found when the pointer was clearly visible. Thus, the position of the prior affects the overall position of referent estimates. Additionally, the finding that the effect of the prior is at least numerically stronger in the pointer-masked condition is in line with the Bayesian model.

D) Is the variability of the responses in the no-pointer condition related to the tendency to deviate from linear extrapolation? As a proxy for the deviation from linear extrapolation, the difference between the referent estimates for the upward pointing arm (20°) and the downward pointing arm (-20°) was computed. The lower this value, the more participants bias their response to the mean. Figures 1i-j show that the higher the variability of the responses in the no-pointer condition (i.e. the prior), the less they deviated from linear extrapolation. This result conforms with the Bayesian approach, according to which the high uncertainty in the a priori assumption of the referent estimate (i.e. high variability) should result in an increased reliance on – and hence smaller deviations from – linear extrapolation.

Summary

In sum, the data support the rational of the Bayesian gesture interpretation model in terms of the model parameters and the functional relationship between different variables. A) The center and shape of the prior assumed in the model are in line with the current data. B) The responses on a masked, difficult to perceive gesture and a clearly visible gesture differ according to the prediction of the Bayesian approach. C) In line with Bayesian approach, the center of the a priori assumed distribution of referent positions is related to referent estimates, especially if the gesture is difficult to perceive. D) Participants who did not reveal strong a priori assumption about referent positions were less prone to deviate in their estimates from a linear extrapolation.

References

Coren, S. (1993) The lateral preference inventory for measurement of handedness, footedness, eyedness, and earedness: norms for young adults. *Bul Psychon Soc, 31*(1), 1-3.

Figure 1. Trial procedure and results. The figure shows the trial procedure and results of the control experiment conducted to test the assumptions underlying the Bayesian arm-finger model. a)-c) show stimulus sequences. d) shows the distribution of z-standardized responses in the no-pointer condition. The red line corresponds to a normal distribution. e)-f) show referent estimates in the pointer-masked and pointer-visible condition. Error bars denote ±1 SD. g)-h) plot referent estimates of the no-pointer condition (i.e. the prior) by referent estimates in the other conditions for individual participants. i)-j) plot the standard deviation of referent estimates in the no-pointer condition (i.e. the prior) by the range between the referent estimates for the 20° and -20° arm orientation. Asterisks mark significant correlations ($p < .05$).