



Cognitive Science (2019) e12733

© 2019 The Authors. *Cognitive Science* published by Wiley Periodicals, Inc. on behalf of Cognitive Science Society (CSS).

All rights reserved.

ISSN: 1551-6709 online

DOI: 10.1111/cogs.12733

Recipient Design in Communicative Pointing

Tobias Winner,^a Luc Selen,^a Anke Murillo Oosterwijk,^{a,b} Lennart Verhagen,^c W. Pieter Medendorp,^a Iris van Rooij,^a Ivan Toni^a

^a*Donders Institute for Brain, Cognition and Behaviour, Radboud University*

^b*Erasmus Research Institute of Management, Erasmus University Rotterdam*

^c*Department of Experimental Psychology, University of Oxford*

Received 29 May 2018; received in revised form 5 February 2019; accepted 28 March 2019

Abstract

A long-standing debate in the study of human communication centers on the degree to which communicators tune their communicative signals (e.g., speech, gestures) for specific addressees, as opposed to taking a neutral or egocentric perspective. This tuning, called *recipient design*, is known to occur under special conditions (e.g., when errors in communication need to be corrected), but several researchers have argued that it is not an intrinsic feature of human communication, because that would be computationally too demanding. In this study, we contribute to this debate by studying a simple communicative behavior, communicative pointing, under conditions of successful (error-free) communication. Using an information-theoretic measure, called *legibility*, we present evidence of recipient design in communicative pointing. The legibility effect is present early in the movement, suggesting that it is an intrinsic part of the communicative plan. Moreover, it is reliable only from the viewpoint of the addressee, suggesting that the motor plan is tuned to the addressee. These findings suggest that recipient design is an intrinsic feature of human communication.

Keywords: Pointing; Communication; Recipient design; Perspective taking; Legibility

1. Introduction

Theoretical accounts of human linguistic communication have highlighted how communicators can engage in recipient design,¹ tailoring their messages for their target

Correspondence should be sent to Ivan Toni, Donders Institute for Brain, Cognition and Behaviour, Radboud University, Kapittelweg 29, 6525 EN Nijmegen, The Netherlands. E-mail: ivan.toni@donders.ru.nl

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

audience while taking the latter's specific perspective in mind (Clark, 1996; Clark & Murphy, 1982; Horton & Keysar, 1996; Newman-Norlund et al., 2009). Recipient design is often applied to resolve ambiguities or miscommunications in everyday linguistic communication (Clark & Murphy, 1982; Hanna, Tanenhaus, & Trueswell, 2003). However, it remains unclear whether those computationally demanding operations (Horton & Keysar, 1996; van Rooij et al., 2011) are an intrinsic or an optional feature of human communication. While some scholars argue that recipient design is primarily a repair mechanism (Horton & Keysar, 1996), others see it at the very core of the human communicative system, which should therefore affect communicative signals by default (Blokpoel et al., 2012; Levinson, 2006). Here, we investigate this question by considering one of the simplest instances of human communicative behavior, that is, communicative pointing, in a context of largely error-free communication. Specifically, we test whether and how communicators tune the informational value of their communicative pointing movements to the particular perspective of the addressee.

Pointing in order to spatially disambiguate a referent among a set of potential targets is an intensely studied gesture in language and communication research, offering an important platform for understanding the ontogenetic and phylogenetic emergence of human referential communication (Scott-Phillips, 2014; Tomasello, 2010). In addition, the spatial continuity and temporal extension of pointing signals make them an empirically rich situation for quantifying the degree and the time course of designedness in human communicative behavior.

Communicative gestures can be segmented into a sequence of discrete stroke and holding phases (Kendon, 1980; McNeill, 1992). Previous studies have shown that, in humans, stroke duration (Peeters, Chu, Holler, Hagoort, & Ozyurek, 2015) and holding time (Murillo Oosterwijk et al., 2017) increase with communicative demands. These observations suggest that communicators increase the informational value of the movement by pointing over a longer time, thereby allowing the addressee to accumulate more sensory evidence against the intrinsic noise in the sensorimotor system (Faisal, Selen, & Wolpert, 2008). Beyond differences in timing and end-point, recent studies have shown that the presence of communicative intent during pointing movements also influences movement trajectories in an addressee-dependent fashion (Cleret de Langavant et al., 2011; Murillo Oosterwijk et al., 2017; Ozyurek, 2000). Murillo Oosterwijk et al. (2017) had participants point out referents for an addressee standing either to their left or right side. The main finding was that pointing trajectories were influenced by the spatial location of the addressee. More precisely, trajectories systematically deviated away from the addressee. These addressee-dependent deviations in movement trajectory may be consistent with the interpretation that communicators consider the spatial perspective of their addressee when organizing their pointing movements. However, it is not obvious whether or how those trajectory modulations influence the informative value of the movements. Pointing trajectories curve away from the addressee (Murillo Oosterwijk et al., 2017), and it remains unclear whether this effect increases the information content of the movement. An alternative explanation for this finding could be that communicators avoid their addressee's peripersonal space without a specific concern for informativeness. Furthermore, changes in trajectory variability can potentially cancel

out the informativeness of changes in average pointing trajectories. Here, by conducting novel analyses on the data collected by Murillo Oosterwijk et al. (2017), we test whether addressee-dependent modulations of the movement trajectory are principled in the sense that they maximize the informativeness of pointing gestures from the addressee's point of view. The informativeness of pointing gestures is operationalized with an information-theoretic measure of *legibility*.

The concept of *legibility* is based on an observer making Bayes-optimal inferences about the most probable target referent given the observed pointing movement, and it has been proposed in the field of Human-Robot Interaction to characterize the information present in communicative movements (Dragan, Lee, & Srinivasa, 2013; Holladay, Dragan, & Srinivasa, 2014). The more legible a movement, the more information it contains about its underlying goal. The legibility $\mathcal{L}(Y|G)$ of a movement trajectory Y directed at goal G equals the probability of G given Y :

$$\mathcal{L}(Y|G) = P(G|Y) \quad (1)$$

A legible trajectory is thus informative because it allows for target inferences with high confidence by an ideal observer. In the case of pointing, this means a high probability of inferring the intended referent of the pointing gesture. Note that the legibility of a pointing movement does not guarantee the ability of an actual human addressee to make those inferences in practice. Nevertheless, it is an important metric because it quantifies the presence of relevant information in the movement, an indication that in principle the communicator is making that information available to the addressee. Legibility can be evaluated from different spatial perspectives, with Y representing the observations corresponding to a particular point of view (e.g., the addressee). For instance, it has been shown that if a robot optimizes legibility from the (human) addressee's perspective, this significantly enhances intention recognition by that addressee (Nikolaidis, Dragan, & Srinivasa, 2016).

Note that both modulations of the average trajectory as observed by Murillo Oosterwijk et al. (2017) or Cleret de Langavant et al. (2011), and reductions in variability (Vesper et al., 2011) can independently lead to increased legibility. However, net legibility depends on all moments of the distribution of observations. Murillo Oosterwijk et al. (2017) report a significant increase in trajectory variability, which could have countered the to-be-expected benefits from the increased mean differences. Increased legibility can thus not be taken for granted but must be analyzed explicitly.

In this study, we quantify the time-resolved legibility of human communicative pointing movements. In the experiment, participants played a referential game in which they had to point out targets with their index finger of the right hand for one of two observers. There were two action conditions. In the Communicative condition, the addressee had to infer the correct referent target from the communicator's pointing movement. In an instrumental control condition, the hand movement triggered a computer to display the referent to the addressee, such that the pointing movement itself did not need to be observed by the addressee, while the motoric and attentional demands equaled those evoked in the communicative condition. Under the assumption that communicators

engage in recipient design, the hypothesis is that there will be larger legibility in communicative pointing as compared to instrumental pointing, especially when legibility is evaluated from the perspective of the addressee rather than from the point of view of the passive onlooker or from the egocentric perspective of the communicator.

2. Methods

This study is based on novel and independent analyses of existing data (Murillo Oosterwijk et al., 2017). Below, we report relevant details of the experimental setting in which the data were acquired, together with details of the novel legibility analyses.

2.1. Participants

Data considered for analyses were acquired in 13 triads of right-handed participants (11 females, age 17–29). Each experimental session featured a participant playing the role of the communicator while two observers played as addressee/onlooker. These roles are described in detail below.

2.2. Materials

Communicators were seated at a table in front of two vertically aligned computer screens (each 56 cm). The center of the bottom screen was positioned at eyelevel and just outside of reaching distance of the communicator. Within reaching distance, 17 cm in front of the screen, an infrared position detection frame measured the position of the communicator's index finger in the X-Z plane spanned by the frame. A capacitive sensor ($4.0 \times 3.6 \text{ cm}^2$)—the “home-key”—was used to measure and constrain the communicator's initial hand position on each trial. The sensor was positioned 52–57 cm in front of the bottom screen, depending on the participant's arm length. Two observers sat to the communicator's left and right at the same table, facing the bottom screen at an angle of $\pm 40^\circ$ with respect to the communicator. They had access to a computer mouse to respond to the communicator's signals.

Kinematic data of the communicator were recorded using an electromagnetic tracking system (LIBERTY, Polhemus) at a sampling rate of 250 Hz. Four sensors, providing position and orientation data, were attached to the communicator's hand (at the distal phalanx of the index finger and thumb, at the metacarpophalangeal joint, and at the wrist).

Participants played a referential game involving sets of referent tokens and pointing targets. The pointing targets were distinctly colored square icons ($4 \times 4 \text{ cm}^2$), and the referent tokens were two-colored triangle-circle composites, randomly arranged on a 3×3 grid. On each trial there was one square that matched either one of the referent's two colors, such that the referent could be inferred using rules of pragmatic inference (Frank & Goodman, 2012; see Fig. 1). For example, the blue square would be pointed at by the communicator to indicate that the referent token includes either a blue triangle or

a blue circle. The targets pointed at by the communicator and the referent tokens selected by the addressee were thus related through a logical link rather than through visuospatial regularities.

2.3. *Experimental design*

An experimental session consisted of 240 trials, subdivided into blocks of 10 trials. The experimental task required communicators to inform their addressee about the referent token by pointing at one of three targets. Within the 10 trials of one block, referent tokens and targets were determined at random, with the constraint that conditions could not be repeated for more than three times in a row and all conditions and configurations were presented in equal numbers in both the first and second half of the experiment. Apart from the target position, two factors, *addressee position* (left, right) and *action* (communicative, instrumental), were manipulated.

In the communicative condition, the addressee had to select the referent token based on the pointing movement produced by the communicator. In the instrumental condition, addressees could still observe the movement. However, the computer recognized the indicated token based on the communicator's finger position in the detection frame (within a fixed circular tolerance region in the x-z plane of the frame) and displayed the corresponding referent token to the addressee (Fig. 1). Thus, the two experimental conditions differed only in the addressee's dependency on the relation between communicator's movement and intended referent: Inference of the referent token from the pointing movement was not necessary in the instrumental condition. The otherwise close match between conditions in terms of general task demands and context was designed to minimize global differences in pointing behavior, allowing the detection of modulatory effects of addressee position.

The two observers alternated in their roles as the addressee, who had to identify the referent token, and the onlooker, who had no particular task. The roles of addressee and onlooker switched every block, while the action condition changed every other block. The communicator would thus address each observer in each condition before the action category changed. Before each block of 10 trials, participants were informed of the forthcoming conditions, that is, where the addressee/onlooker was located in the forthcoming block, and whether the trials were part of the Communicative or Instrumental control conditions.

2.4. *Experimental task and procedure*

The experimental task (Fig. 1) required the addressee to select with his computer mouse one out of three possible referent tokens that were displayed on the upper screen. The correct referent token was only known to the communicator who could convey this information by pointing at one of three horizontally aligned targets on the bottom screen. All other forms of communication between the participants were prohibited. Furthermore, the experimental setup (Fig. 1) and continuous visual monitoring of the participants'

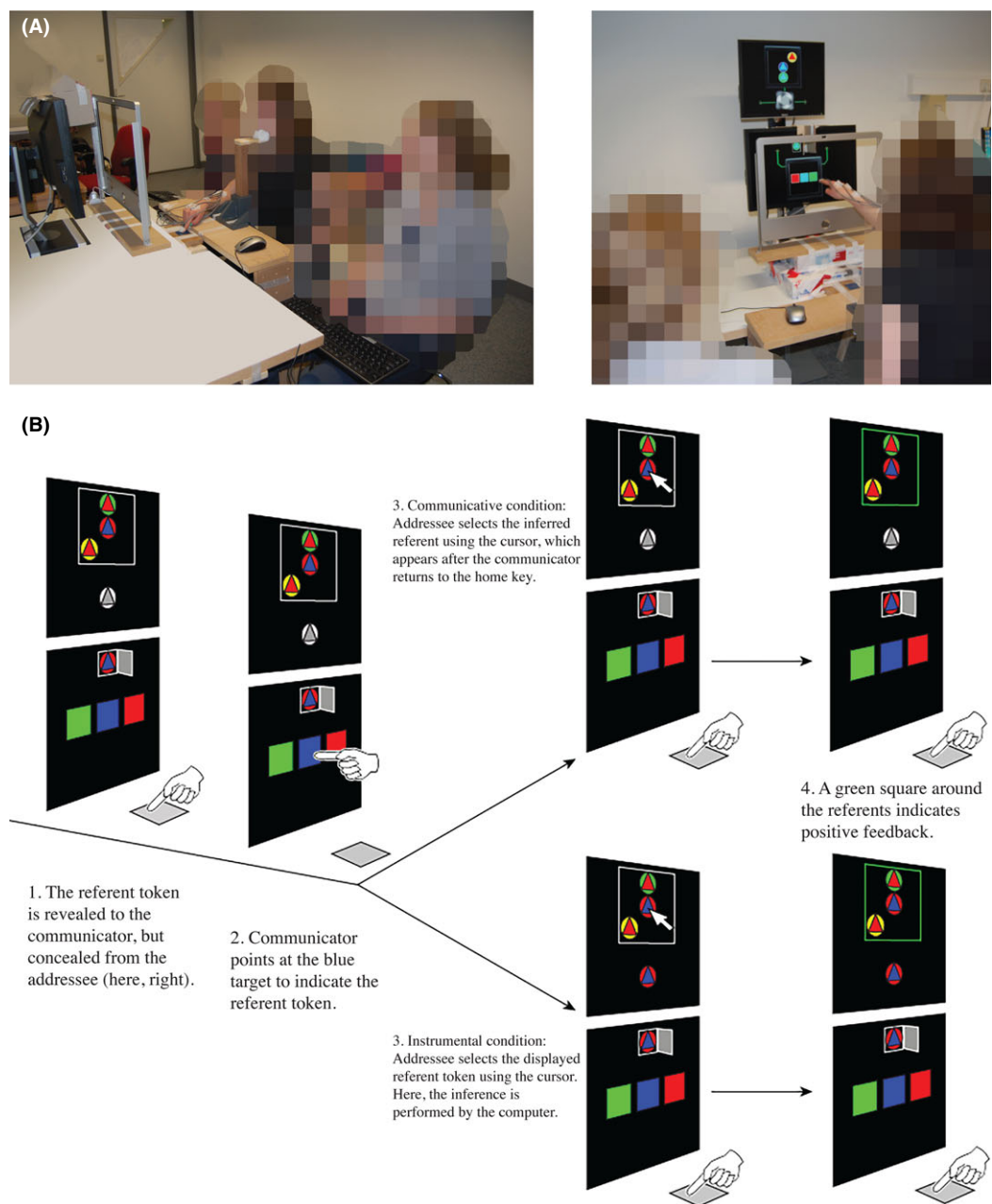


Fig. 1. Experimental configuration and task setup. (A) A communicator (center) and two observers were seated at a table. The communicator's chin rested on a fixed support. A trial started with the communicator resting the right index finger on a home key, triggering the presentation of the visual stimuli on two vertically aligned monitors. On each trial, one of the two observers had full vision of the monitors (the onlooker), whereas the view of the other observer was partially occluded by a cardboard screen (the addressee). The communicator's hand-pointing movements were visible to both addressee and onlookers, and monitored with

an electromagnetic system (gray cables taped to the hand/arm), as well as with an infrared system (gray frame in front of the communicator). On each trial, the communicator was instructed to perform a pointing movement toward the bottom screen, passing through the infrared detection frame, and return to the home key. The addressee's response to the communicator's pointing movement was delivered through a computer mouse controlling a cursor, displayed on the top screen (see panel B) after the communicator's hand returned to the home key. (B) Each triplet of participants played a referential game involving sets of tokens (circle/triangle composites) and of pointing targets (squares), displayed on the monitors in front of the communicator. On each trial, a referent token was displayed on the upper half of the lower monitor, and occluded to the addressee's view by the cardboard screen (in gray, see also panel A). The color of one of the pointing targets matched either one of the referent's two colors, such that the addressee (here, on the right of the communicator) could infer the identity of the referent token from the color of the target pointed at by the communicator, using rules of pragmatic inference (Frank & Goodman, 2012; please note that panels A and B provide different examples of targets/referents combinations). In the example shown in panel B, the blue square is pointed at to indicate that the referent token is uniquely characterized by the presence of a blue component. In the *communicative pointing condition* (top row), the addressee had to infer the referent from the pointing movement itself. In the *instrumental pointing condition* (bottom row), the communicator's pointing movement triggered the display of the pragmatically corresponding token on the upper monitor. The two experimental conditions differed only in the addressee's dependency on the relation between communicator's movement and intended referent: Inference of the referent token from the pointing movement was not necessary in the instrumental condition.

behavior excluded that participants might have communicated task-relevant information during the study through communicative channels other than arm pointing (e.g., gaze, lips, head, or other limb movements).

A trial started with the communicator resting her right hand on the home key, eliciting the presentation of all visual stimuli (i.e., target and reference tokens). The private reference token was displayed visually to the communicator with a cardboard barrier concealing the symbol from the addressee. The communicator was instructed to perform a pointing movement toward the bottom screen, passing through the infrared detection frame, and return to the home key (Fig. 1). No other constraints were put on the movement. Subsequently, the addressee could select the referent token with a mouse click, leading to visual feedback indicating whether the correct referent had been selected. The addressee's mouse cursor appeared at the center of the top screen, only after the communicator returned to the home key. It was thus impossible for the addressee to communicate their understanding of the pointing movement before the movement was completed. This condition was introduced to make movements comparable between conditions. In order to nudge participants toward communicative behavior by means of task contingencies, rather than verbal instructions, we presented the experimental task through a cover story, described at length in an English translation of the instructions to the communicator and addressee (originally in Dutch—see Appendix A).

2.5. Data acquisition and pre-processing

We analyzed kinematic data from the sensor attached to the distal phalanx of the right index finger. These data were first low-pass filtered at 15 Hz (sixth-order bi-directional Butterworth filter). Movement on- and offset were then determined for each trial

based on both position and velocity information: The starting point of a movement was defined as the point in time where the index finger was at a maximum distance from the target after reaching a velocity threshold of 0.1 m/s. Similarly, the end-point was determined as the time after passing the same threshold again, at which the distance from the target was minimum. The resulting trajectories of the index finger were then normalized in time using cubic spline-interpolation over 100 time points (Fig. 2A). Ten time points corresponding to 10% pre-movement time were also considered, for display purposes.

For the subsequent analyses, only trials in which the addressee successfully identified the referent token were included (93.5%). On less than 2% of all trials did the addressee misinterpret a pointing gesture. The remaining 4.5% errors were incorrect movements of the communicator (i.e., pointing at the wrong target) that resulted in an “incorrect” referent selection by the addressee. In addition to incorrect trials according to the task definition, 1% of trials were excluded because, statistically, the end-point of the pointing movement was more likely to come from the distribution associated with the non-targets rather than the target. In other words, judged from the recorded trajectory, the communicator indicated the wrong target in these 1% of trials. On 5% of trials, no well-defined movement was registered according to the on- or offset criteria; 1.5% of trials were outliers on movement duration (with a criterion of 1% likelihood); and 5% of trials were spatial outliers (considering a criterion on the Mahalanobis distance of 0.1% likelihood per time point). This left a total of 80.5% of all trials for analyses.

2.6. View-point modeling and Legibility estimation

To model the viewpoints of the communicator and the addressee/onlooker, the three-dimensional pointing data were projected onto two-dimensional planes, approximating the observer’s view of the screens and of the pointing movements (Nikolaïdis et al., 2016). To approximate the communicator’s perspective, we projected the pointing trajectories onto the coronal plane of the communicator when oriented toward the middle target. The same was done for the addressee and onlooker. The planes were thus orthogonal to a “line of sight” that connects the participant and the middle pointing target. Fig. 2 shows how these transformations resulted in a dimensionality reduction that approximates the viewpoints of the communicator and addressees.

In addition to the three fixed perspectives of the communicator, addressee, and onlooker, we analyzed legibility from a continuum of perspectives by parametrically varying the perspective of a virtual observer along the horizontal axis, which corresponds to placing an observer at the table, either to the left or right of the communicator at different distances. This distance was varied in 31 equal steps between -75 cm and 75 cm (with zero at the level of the central target). The following analysis steps were performed after projecting the data onto planes of interest (i.e., addressee, communicator, onlooker or virtual observer).

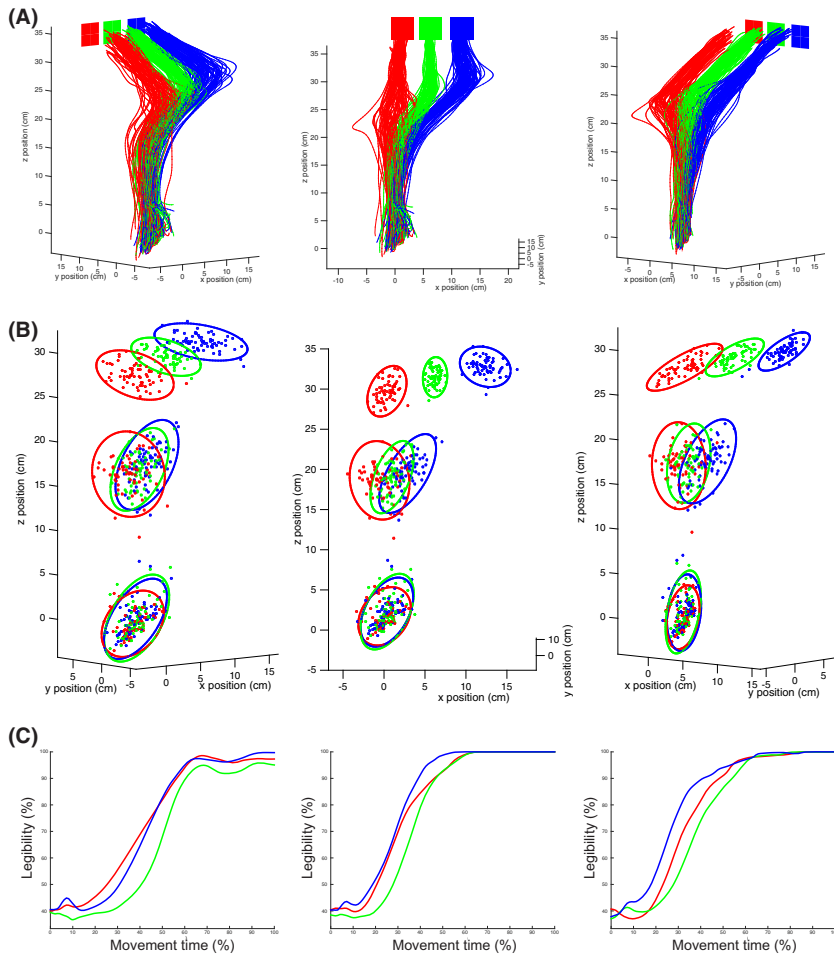


Fig. 2. Spatial distribution of pointing movements and legibility time course. (A) Three-dimensional illustration of the pointing trajectories generated by one participant. In this figure, the red, green, and blue squares indicate the location of the left, central, and right target, respectively (although during the experiment each target had different colors on different trials). Pointing trajectories are colored correspondingly, and projected onto the transverse planes of the communicator (middle panel), and of the left and right observers (left panel, right panel, respectively). (B) Index finger positions (dots) and covariance ellipses (95% volume) for three selected time points ($t = 0\%$, $t = 30\%$ and $t = 70\%$) of the pointing trajectories, as seen from the same perspectives and with the same conventions as in panel A. (C) Legibility time courses as percentage of movement time, evaluated from the viewpoint of the communicator (middle panel) and of the left and right observers (left panel, right panel, respectively). The red/green/blue curves represent legibility toward the right/middle/left target, as in panel A. The largely overlapping ellipses at the onset of the movement, shown in the lower third of panel B, result in low legibility. The spatially divergent ellipses toward the end of the movement, shown in the upper third of panel B, result in greater legibility. It can also be seen that the middle target (in green) has lower legibility, being flanked by the other two targets.

Position mean and covariance were computed separately at each time point for the 12 possible combinations of target, addressee, and condition (left, middle, and right targets; left and right addressee; communicative and instrumental conditions). The time-varying two-dimensional Gaussians defined by these parameters were used as estimates for the conditional probability densities $P(Y|G)$. Based on these densities, we determined the legibility as a function of time for the individual trials using Eqs. 1 and 2. That is, the legibility $L(g)$ of a trial directed at target g when the index finger is at position y_t at time point t was computed using Bayes' rule, assuming a constant uniform prior (i.e., 33.3% for each of the three targets):

$$\mathcal{L}(y_t|g) = P(g|y_t) = \frac{P(y_t|g)}{\sum_{h \in G} P(y_t|h)} \quad (2)$$

Here, G denotes the set of possible referents (left/middle/right). The result is a time course of legibility for every trial (Fig. 2C).

Legibility time courses from trials with different targets and addressee positions were collapsed to obtain two legibility time courses per participant, one for each experimental condition (communicative, instrumental). Finally, we computed the difference between these time courses, resulting in a single time course of *differential legibility* per participant.

2.7. Statistical inference

The whole time series of differential legibility was analyzed using cluster-based permutation tests (Maris & Oostenveld, 2007). Clusters of positive differential legibility were defined as follows: First, t values for the differential legibility time courses were computed at every time point. We then identified the longest sequence of time points for which these values exceeded the threshold corresponding to a two-sided t test. The cluster's area under the average differential legibility curve was used as cluster statistic. Non-parametric permutation tests were performed on that statistic to obtain a conservative estimate of Type I error probability (Maris & Oostenveld, 2007). The permutation test considered the 2^{13} possible subsets from the 13 participants.

3. Results

We considered pointing movements evoked during two tasks in which a communicator signaled a token to an addressee. In one condition (Communicative), the addressee had to extract this information from the movement itself. In the control condition (Instrumental), a computer detected the position of the communicator's finger and presented the corresponding token to the addressee. This experimental manipulation was designed to have conditions closely matched for communicative context and visuospatial demands, allowing us to isolate specific effects of recipient design. This experimental approach was

Table 1
Movement parameters (group averages)

Finger Position at the Movement End-Point (cm)	Instrumental Pointing			Communicative Pointing		
	Left Target	Middle Target	Right Target	Left Target	Middle Target	Right Target
Frontal axis (x)	1.39	5.76	10.26	1.28	5.57	10.44
Sagittal axis (y)	9.02	10.80	12.47	8.68	10.64	12.45
Longitudinal axis (z)	29.27	30.17	30.34	29.26	29.89	30.30
Path length (cm)	40.51	42.35	44.88	40.13	42.01	44.75
Movement time (ms)	933	939	916	968	972	954

successful: There were no statistical differences in average movement speed between conditions (all $p > .05$), making it possible to compare time-normalized trajectories and legibility time courses across the two experimental conditions. Table 1 summarizes a number of temporal and spatial parameters of pointing movements.

Fig. 3A shows the temporal development of average legibility, pooled across the three targets, during Communicative and Instrumental trials, from the addressee's point of view. Starting slightly above chance level (39% rather than 33.33%), the time courses describe a sigmoidal curve approaching almost 98% target certainty toward the end of the movement. Legibility time courses for individual targets are shown in Fig. 2C. Not surprisingly, being flanked to the left and right by other potential targets, the middle target is generally less legible.

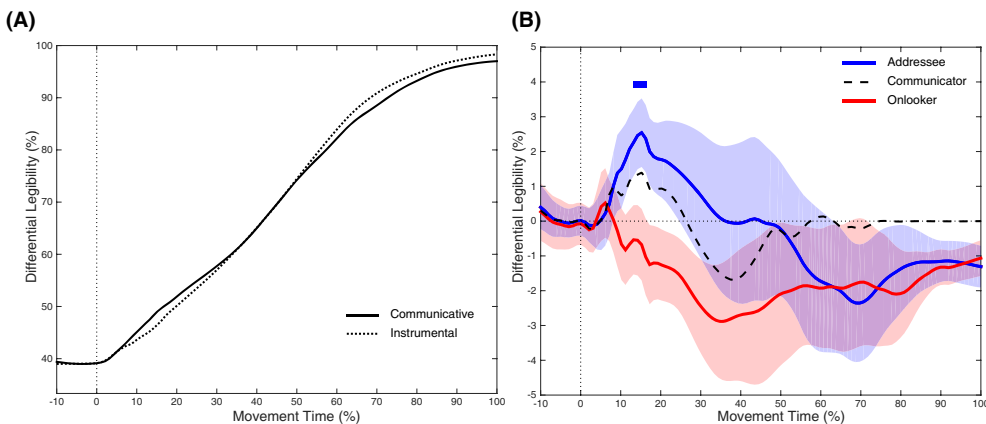


Fig. 3. Differential legibility between communicative and instrumental pointing. (A) Time course of legibility observed during communicative and instrumental conditions (group average) as evaluated from the addressee's perspective. (B) Time course of the difference in legibility between communicative and instrumental conditions (± 1 SE), as evaluated from the addressee's perspective (in blue), and from the perspective of the onlooker (in red). The blue bar marks the cluster (14%–17% of movement time) where this difference is significant for the addressee-projection.

The main finding of this study is that, when evaluated from the perspective of the addressee, there is an increase in legibility for communicative pointing movements relative to instrumental movements (Fig. 3B). This differential effect occurs between 6% and 36% of movement time, being statistically different from zero ($p = .0122$) between 14% and 17% of movement time (peak differential legibility: 2.5%). This effect thus occurs during the initial phase of the pointing movement. From both the communicator's and the onlooker's perspective there was no statistically significant effect in differential legibility. For those comparisons, the t values of differential legibility for all individual time-points failed to cross the threshold and thus no clusters were formed.

While differential legibility was not significant from the communicator's perspective or the onlooker's perspective, Fig. 3B suggests the presence of a gradient of differential legibility depending on the position of an observer relative to the addressee. We tested this possibility by considering how legibility varies gradually as a function of the visual perspective of a virtual observer. Fig. 4 illustrates how the pointing trajectories for the left and right addressee are associated with different patterns of spatially specific legibility profiles. Here, differential legibility was averaged in the significant cluster. These patterns exhibit a mirror symmetry centered around the communicator's position with increasing values of differential legibility in the direction of the addressee.

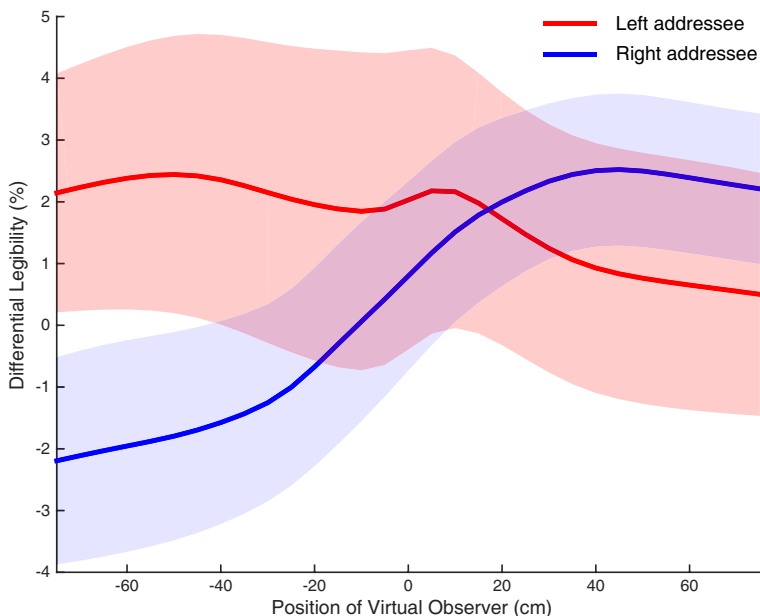


Fig. 4. Differential legibility as a function of addressee's position. Differential legibility between communicative and instrumental conditions, at 14%–17% of movement time, as evaluated at different points of view, for the left addressee (in red) and the right addressee (in blue). Position values on the x -axis indicate the lateral displacement of the virtual observer relative to the communicator.

4. Discussion

In this study, we assessed the legibility of communicative pointing movements and its sensitivity to the spatial perspective of an addressee. We compared the differential legibility of communicative pointing movements (relative to a closely matched control condition) when evaluated from the perspective of the communicator, the addressee, or an onlooker. Differential legibility increases early during pointing, when the movements are seen from the perspective of the addressee, but not from the perspective of the communicator or of the onlooker. We interpret these findings as reflecting an intrinsic human communicative bias toward the production of pointing movements that are informative for a specific addressee, and thus as evidence for recipient design occurring within the communicative process. The intrinsic nature of this bias is supported by two observations. First, the differential legibility effect was observed in a setting in which communication was largely error-free and, hence, there was little to no pressure on communicators to invest additional computational resources in order to engage in recipient design. This suggests that the legibility effect reflects a constituent element of the pointing movements, rather than a consequence of repair or learning mechanisms. Second, the legibility effect occurred early and transiently during the stroke phase of the pointing movement. This temporal profile suggests that communicative pointing movements are produced with a qualitatively different objective than instrumental pointing. This in turn implies a deep integration of recipient design into motor planning and motor control processes.

This study adds to previous work on communicative pointing by quantifying the informative value of those movements. For instance, it is known that communicative pointing results in more emphatic pointing stroke duration (Peeters, Hagoort, & Özyurek, 2015) and longer holding times than those evoked in control conditions (Murillo Oosterwijk et al., 2017). Furthermore, the trajectories of communicative pointing movements are influenced by the spatial location of the addressee (Murillo Oosterwijk et al., 2017). Here, we show that those trajectory modifications increase the informational value of the pointing movement, when seen from the spatial perspective of the addressee, which was not yet implied by the previous observations on this data set (Murillo Oosterwijk et al., 2017). This observation lends itself to be interpreted in terms of recipient design, in line with previous observations suggesting that communicators organize their movement in an inferential perspective centered on the addressee (Blokpoel et al., 2012; Heller, Grodner, & Tanenhaus, 2008; Peeters, Hagoort, & Özyürek, 2015; Surtees, Apperly, & Samson, 2013). By the same token, the current findings are not immediately compatible with accounts that consider recipient design exclusively as a communicative repair operation (Horton & Keysar, 1996). It remains to be seen whether the current findings, obtained in the context of a simple communicative pointing movement, generalize to other gestures, and in particular to co-speech gestures. Several works have shown that speakers' gestures change in response to the presence of an addressee, the addressee's knowledge, and the addressee's spatial position (Bangerter & Mayor, 2013; Bavelas & Healing, 2013; Holler & Beattie, 2003; Özyurek, 2018).

The current study opens a way to quantify whether those changes increase the informativeness of the speaker's gestures for the addressee.

Another important aspect of the present findings concerns the timing of the observed increase in legibility along the pointing trajectories. This effect peaks around 16% of the total movement time, while the end-point distributions show no signs of differential legibility. This observation highlights the importance of considering fine-grained parameters of gesture production and comprehension in the study of communicative behavior (Keysar et al., 2000).

4.1. Interpretational Issues

Legibility has an intuitive interpretation as an information-theoretic quantity, yet legibility estimates obtained from empirical observations can be biased. For instance, Fig. 3 illustrates how the estimated legibility—and therefore classification rate—is numerically above chance even before the target is known to the communicator. Indeed, assuming an idealized Bayesian observer who knows the true probability distributions $P(G|Y)$ and who infers the goal using the maximum a-posteriori estimator, the rate of successful goal classification would always be greater than or equal to the mean legibility. The bias for exaggerated legibility estimates arises because legibility is evaluated on the same dataset used to estimate probability densities. However, this bias does not differ across conditions (Fig. 3), and the main result concerns *differential legibility*, that is, legibility differences across conditions.

Note that legibility differences cannot simply be explained by different demands on spatial accuracy in the communicative and instrumental condition. In fact, instrumental trajectories become less variable (Murillo Oosterwijk et al., 2017) and numerically (though not significantly) more legible than communicative trajectories in the final part of the movement. End-point variance—reflecting more or less “precise” pointing in either condition—therefore cannot provide an alternative explanation for the observed legibility differences early in the trajectory.

This study relies on several simplifications. We focused on the position of the index finger while ignoring its orientation. This simplification might account for the lack of differential legibility effects in the late portion of the pointing movements. It is conceivable that, toward the end of the stroke phase, the orientation of the index finger rather than its position becomes the controlled movement feature. However, this simplification is unlikely to have biased the main finding. The orientation of the index finger does not yet stabilize in the time window that contains the differential legibility effect derived from the index finger position. Further studies combining position and orientation information are needed to test the hypothesis that legibility is modulated along the full trajectory, and to replicate the current observations on differential legibility.

As for the timing and limited temporal extent of the effect, note that legibility of positions was evaluated independently over time. If the observer accumulates evidence over time, then early deviations have a larger “snowball” effect than late modulations. To assess the cumulative effect of the early modulations would require modeling not only

the distribution of positions per individual time step, but the entire joint probability distribution of positions over time. Increased legibility at an early time point would then be reflected in a larger, updated prior probability for the target at the next time point. Here, the estimation of joint trajectory probabilities was not possible given the amount of data available. As a result, the effect of early modulations is not reflected in the later parts of the legibility time course.

In addition, the convergence of trajectories between conditions toward the end of the pointing movements may also be due to universal constraints on the end-posture. For instance, it is well documented that the pointing finger is aligned to the dominant eye on or close to the eye-target line (Bangerter & Oppenheimer, 2006; Henriques & Crawford, 2002). This further highlights the importance of analyzing full gesture trajectories rather than end-points alone in the search for traces of recipient design.

Trajectory modulations and legibility differences between conditions are numerically small. This could be a consequence of the general principle of signal-dependent noise in the nervous system (Faisal et al., 2008; Harris & Wolpert, 1998), penalizing increased amplitudes and complexity of movements with increased variability. In addition, in both conditions, participants were asked to move their finger through the IR-frame. This constrained the end-points of the movements to converge in a similar region. This constraint made conditions motorically comparable, but it could explain why modulations were as subtle as observed. It remains to be seen whether these subtle modulations can have an observable effect on the addressee, for instance in terms of reaction times or gaze. In this study, we cannot assess those effects, given the requirement of returning the communicator's hand to the starting position before the addressee could respond.

Finally, in interpreting the results presented here, we take a perspective on language and communication that emphasizes the role of domain-independent, general socio-cognitive processes in both natural language and gesture (Blokpoel et al., 2012; Levinson, 2006). We take the explanatory success of Bayesian models of pragmatics in natural language (Frank & Goodman, 2012), which employ similar assumptions and reasoning as here, as evidence for a common role of inferential processes across domains. Nonetheless, the possibility remains that our conclusions do not apply to natural language processing. Future work on both gesture and natural language processing can build on the methodology developed here to assess the generality of the findings reported here.

4.2. *Conclusion*

This study uses an information theoretic measure to extract a sensitive index of recipient design from communicative pointing behaviors evoked during successful interactions with an addressee. We present empirical evidence showing that participants increase the information-theoretic value of their communicative pointing gestures. This effect is temporally and spatially specific: It occurs during the early stages of the movement, and it occurs when the movement is processed from the addressee's point of view. These findings provide evidence for recipient design as a constituent element of communicative pointing movements and open the way for quantifying recipient design during human gestural communication.

Acknowledgments

This work is part of the research programme Language in Interaction, financed by the Netherlands Organisation for Scientific Research (NWO). WPM is supported by an NWO Vici grant (no. 453-11-001). The authors thank Mark Blokpoel for editing the photographs.

Note

1. In this paper we adopt the term “recipient design” (Blokpoel et al., 2012; Sacks, Schegloff, & Jefferson, 1974) rather than the term “audience design” (e.g., Clark & Carlson, 1982). The latter refers to the phenomenon that a communicator adapts her signal not only for the addressee but also takes into account other hearers in the audience. Hence, recipient design can occur in the context of audience design, but also without it.

References

- Bangerter, A., & Mayor, E. (2013). Interactional theories of communication. In P. Cobley & P. J. Schulz (Eds.), *Theories and models of communication* (pp. 257–271). Berlin: Walter de Gruyter.
- Bangerter, A., & Oppenheimer, D. M. (2006). Accuracy in detecting referents of pointing gestures unaccompanied by language. *Gesture*, 6(1), 85–102.
- Bavelas, J., & Healing, S. (2013). Reconciling the effects of mutual visibility on gesturing: A review. *Gesture*, 13(1), 63–92.
- Blokpoel, M., van Kesteren, M., Stolk, A., Haselager, P., Toni, I., & van Rooij, I. (2012). Recipient design in human communication: Simple heuristics or perspective taking? *Frontiers in Human Neuroscience*, 6, 253.
- Clark, H. H. (1996). *Using language*. Cambridge, UK: Cambridge University Press.
- Clark, H. H., & Carlson, T. B. (1982). Hearers and speech acts. *Language*, 58(2), 332–373.
- Clark, H. H., & Murphy, G. L. (1982). Audience design in meaning and reference. *Advances in Psychology*, 9, 287–299.
- 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), 1–10.
- Dragan, A. D., Lee, K. C. T., & Srinivasa, S. S. (2013). Legibility and predictability of robot motion. In H. Kuzuoka & V. Evers (Eds.), *ACM/IEEE international conference on human-robot interaction* (pp. 301–308). Piscataway, NJ: IEEE Press.
- Faisal, A. A., Selen, L. P. J., & Wolpert, D. M. (2008). Noise in the nervous system. *Nature Reviews Neuroscience*, 9(4), 292–303.
- Frank, M. C., & Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. *Science*, 336, 998.
- Hanna, J. E., Tanenhaus, M. K., & Trueswell, J. C. (2003). The effects of common ground and perspective on domains of referential interpretation. *Journal of Memory and Language*, 49(1), 43–61.
- Harris, C. M., & Wolpert, D. M. (1998). Signal-dependent noise determines motor planning. *Nature*, 394, 780.
- Heller, D., Grodner, D., & Tanenhaus, M. K. (2008). The role of perspective in identifying domains of reference. *Cognition*, 108(3), 831–836.

- Henriques, D. Y. P., & Crawford, J. D. (2002). Role of eye, head, and shoulder geometry in the planning of accurate arm movements. *Journal of Neurophysiology*, 87(4), 1677–1685.
- Holladay, R. M., Dragan, A. D., & Srinivasa, S. S. (2014). Legible robot pointing. In P. A. Vargas (Ed.), *The 23rd IEEE International Symposium on robot and human interactive communication* (pp. 217–223). Edinburgh: IEEE.
- Holler, J., & Beattie, G. (2003). How iconic gestures and speech interact in the representation of meaning: Are both aspects really integral to the process? *Semiotica*, 146, 81–116.
- Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition*, 59(1), 91–117.
- Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. In M. R. Key (Ed.), *The relationship of verbal and nonverbal communication* (pp. 207–227). The Hague, the Netherlands: Mouton.
- Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (2000). Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science*, 11(32), 38.
- Levinson, S. C. (2006). On the human “interaction engine.” In N. J. Enfield, S. C. Levinson & L. C. Aiello (Eds.), *Roots of human sociality: Culture, cognition and interaction* (pp. 39–69). Duck, NC: Berg.
- Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, 164(1), 177–190.
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago, IL: University of Chicago Press.
- Murillo Oosterwijk, A., de Boer, M., Stolk, A., Hartmann, F., Toni, I., & Verhagen, L. (2017). Communicative knowledge pervasively influences sensorimotor computations. *Scientific Reports*, 7, 4268.
- Newman-Norlund, S. E., Noordzij, M. L., Newman-Norlund, R. D., Volman, I. A., De Ruiter, J. P., Hagoort, P., & Toni, I. (2009). Recipient design in tacit communication. *Cognition*, 111(1), 46–54.
- Nikolaidis, S., Dragan, A., & Srinivasa, S. (2016). Viewpoint-based legibility optimization. In *The Eleventh ACM/IEEE international conference on human–robot interaction* (pp. 271–278). Piscataway, NJ: IEEE Press.
- Ozyurek, A. (2000). The influence of addressee location on spatial language and representational gestures of direction. In D. McNeill (Ed.), *Language and gesture* (pp. 64–83). Cambridge, UK: Cambridge University Press.
- Ozyurek, A. (2018). Role of gesture in language processing: Toward a unified account for production and comprehension. In S.-A. Rueschemeyer, & M. G. Gaskell (Eds.), *Oxford handbook of psycholinguistics* (pp. 592–607). Oxford: Oxford University Press.
- 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*, 12, 2352–2368.
- Peeters, D., Hagoort, P., & Özyürek, A. (2015). Electrophysiological evidence for the role of shared space in online comprehension of spatial demonstratives. *Cognition*, 136, 64–84.
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50(4), 696–735.
- Scott-Phillips, T. (2014). *Speaking our minds: Why human communication is different, and how language evolved to make it special*. London: Palgrave MacMillan.
- Surtees, A., Apperly, I., & Samson, D. (2013). Similarities and differences in visual and spatial perspective-taking processes. *Cognition*, 129(2), 426–438.
- Tomasello, M. (2010). *Origins of human communication*. Cambridge, MA: MIT Press.
- van Rooij, I., Kwisthout, J., Blokpoel, M., Szymanik, J., Wareham, T., & Toni, I. (2011). Intentional communication: Computationally easy or difficult? *Frontiers in Human Neuroscience*, 5(6), 52.
- Vesper, C., van der Wel, R. P., Knoblich, G., & Sebanz, N. (2011). Making oneself predictable: Reduced temporal variability facilitates joint action coordination. *Experimental Brain Research*, 211, 517–530.

Appendix A: Task instructions

Dear Players, you are at the beginning of a secret mission: Project CP.

Your goal: open as many vaults with gold bars as possible. This is not an easy task. The locks of the vaults are extremely well secured. Only by entering a correct code, you will be able to open a safe. Press [THE BUTTON] to continue.

You enter through the frame that you see before you. You move through it with your finger. It works like a touchscreen, but without touch. You can practice this shortly. Then all information about the codes is given. Good luck!

This is an example of a safe that you have to open. You always do that with two people. The person in the middle is always Player 1. The others (left and right) play alternately Player 2. On the screens there are figures: three squares for Player 1 and three circles for Player 2. It is not easy to open the safe. Because of a cardboard barrier, only Player 1 sees the correct code!

Soon the cardboard will be placed on this screen. When the cardboard is placed on the left of the code, then Player 2 on the left must play. When the cardboard is placed on the right of the code, then Player 2 on the right has to play. Soon, Players 2 will have to sit down so that they cannot see the code. The screen of Player 2 is above the screen of Player 1. The outer players have a mouse with which they can select a figure on the top screen.

Normally, the computer can pass the code to Player 2 because there is a cable between the two screens. But this safe is broken! The cable between the two screens is broken and the frame of Player 1 stops responding. There will be squares on the screen, but Player 1 cannot enter a square.

Fortunately, the safe can still be opened, because it is sufficient if Player 2 selects the correct code. Because Player 2 cannot see the code (through the cardboard), Player 1 must help. This can be done by Player 1 using the colors of the squares. The colors of the squares are reflected in the circles. As a hint for Player 2, Player 1 can thus designate a colored square.

Discuss which figure you would enter as Player 2 if Player 1 points to the yellow square. Discuss which figure you would enter when Player 1 points to the red square. Consult which figure you would enter if Player 1 points to the purple square. If Player 2 has selected the correct figure, the lock will recognize the correct code. The safe will then become (temporarily) green in color. When you do this 10 times correctly, then you have the right code and the safe will open.

Some more tips before you start practicing: (a) Player 2 starts by pressing SPACE. Both Players must then also press the home button and ctrl button. (b) Think first and then make the movement to the screen. (c) Immediately press the home button again. Only then Player 2 can (d) release his home button and select the right figure with the mouse. Press [THE BUTTON] to continue practicing.

Now you can practice eight situations together. First, Player 2 (left) must crack the code at the direction of Player 1! You may still consult, but later on you are not supposed to talk. Press [THE BUTTON] to continue practicing.

Now the other Player 2 (right) must crack the code at the direction of Player 1! You may consult again. Press [THE BUTTON] to continue practicing.

There are also safes that are not broken. The frame of Player 1 works correctly and a square can be entered. There can also be a signal from the bottom screen, to the lock of the safe. The computer has each square encoded into a corresponding circle, which fits best in color. Each square thus corresponds to a circle according to a fixed pattern. Player 1 has to look for the square that corresponds to the target figure.

The square that Player 1 indicates is automatically decrypted by the computer and the corresponding circle is forwarded via the cable to the lock. Player 2 then only needs to confirm the code by selecting this figure on the screen. Decipher now the secret color pattern. Why does the dark blue square code this figure (arrow above)? Why does the green square code this figure? Why does the light blue square code this figure?

In this example, the correct square (green) is selected. The signal is converted to the code figure, which appears on the safe for Player 2. Player 2 can then see on his own screen which figure he/she should select. He no longer has to pay attention to which figure Player 1 points to. If Player 2 has entered the correct figure, the safe will become (temporarily) green in color. When you do this several times, then the safe will eventually open.

Now you can practice eight situations again, first four left, then four right. Player 2 does not have to press SPACE. Think carefully, only Player 1 has to crack the code by typing in the right square. So press inside the square, otherwise the computer will not register properly. You will hear a beep now if you have designated the square and you have pressed the home button again. Press [THE BUTTON] to continue practicing.