I Know You Know I'm Signaling: Novel gestures are designed to guide observers' inferences about communicative goals

Amanda Royka (amanda.royka@yale.edu)[1], Marieke Schouwstra (m.schouwstra@uva.nl)[2], Simon Kirby (simon.kirby@ed.ac.uk)[3], Julian Jara-Ettinger (julian.jara-ettinger@yale.edu)[1]

[1]Department of Psychology, Yale University, [2] Institute for Logic, Language & Computation, University of Amsterdam, [3] School of Philosophy, Psychology & Language Sciences, University of Edinburgh

Abstract

For a gesture to be successful, observers must recognize its communicative purpose. Are communicators sensitive to this problem and do they try to ease their observer's inferential burden? We propose that people shape their gestures to help observers easily infer that their movements are meant to communicate. Using computational models of recursive goal inference, we show that this hypothesis predicts that gestures ought to reveal that the movement is inconsistent with the space of non-communicative goals in the environment. In two gesture-design experiments, we find that people spontaneously shape communicative movements in response to the distribution of potential instrumental goals, ensuring that the movement can be easily differentiated from instrumental action. Our results show that people are sensitive to the inferential demands that observers face. As a result, people actively work to help ensure that the goal of their communicative movement is understood.

Keywords: Communication; theory of mind; goal inference

Introduction

In one episode of the sitcom Seinfeld, a waitress scratches her cheek with her middle finger while talking to the main characters. The group of friends are then forced to debate whether the action was a gesture of derision or a purely goal-directed scratch. Even though the waitress was making eye contact and performed an action that looked like a conventional gesture, the characters still had to infer the goal of her movement. This example nicely illustrates that gestural communication is an inferential process. In the same way that observers need to infer whether the goal of a movement is to grab object A or object B, observers also need to infer when the goal of a movement is to communicate (Royka, Aboody & Jara-Ettinger, 2018; Royka, Chen, Aboody & Jara-Ettinger, in review; Scott-Phillips, Kirby & Ritchie, 2009; Trujillo et al., 2018). While researchers have long appreciated that people have to infer what a communicative signal means (Akhtar, Carpenter & Tomasello, 1996; Frank & Goodman, 2012; Galantucci, 2005; Sperber & Wilson, 1986), there has been significantly less work exploring how communicators signal that a signal is a signal in the first place.

The important role of communicative inference is neglected partly because there are simpler answers to the question, how do you let someone know that you are moving to communicate? Communicative movements are often conventionalized (Ekman & Friesen, 1969; Goldin-Meadow, 1999; Kendon, 1997) and accompanied by ostensive cues, such as eye contact (Behne, Carpenter & Tomasello, 2005; Csibra & Gergely, 2009; Senju & Csibra, 2008), possibly allowing people to recognize gestures through a superficial analysis of action alone. However, this cannot account for the full flexibility of human gesture. Indeed, in novel experimental contexts, humans can both create new communicative gestures and attribute communicative goals to unfamiliar movements in the absence of ostensive cues or pre-established conventions (De Ruiter et al., 2010; Royka et al., 2018; Scott-Phillips et al., 2009). Therefore, ostensive cues and the use of conventional gestures must function in tandem with a more dynamic means of signaling that a movement is meant to communicate: inference.

Communicative goals are unique because they can only be fulfilled when the intended observer correctly infers the communicative goal. Here, we propose that people shape their communicative movements to help observers correctly infer that their goal is communicative. In other words, the very features of a gesture should signal that the movement is meant to communicate—even in the absence of ostensive cues or prior experience with the signal. Thus, under this account, signalers should choose their communicative movements based on which goals their observers are likely to infer. As such, signalers should change their communicative movements in response to changes in the distribution of non-communicative goals.

Some initial evidence shows that, when creating novel signals, people move in ways that are spatially and temporally inefficient relative to the pursuit of instrumental goals in the environment (e.g., making pauses or repetitive oscillations; De Ruiter et al., 2010; Scott-Phillips et al., 2009). This is consistent with the idea that gestures ought to be differentiated from non-communicative goal-directed movements since observers expect instrumental action to be efficient (see Jara-Ettinger, Gweon, Schulz & Tenenbaum, 2016 for review; Csibra, Bíró, Koós & Gergely, 2003; Gergely, Nádasdy, Csibra & Bíró, 1995; Skerry, Carey & Spelke, 2013). However, it is unclear if people created these spatiotemporally inefficient gestures because those movements were readily distinguishable from instrumental action. Additionally, these studies focused on interactive situations where communicators received feedback from their observer. Therefore, this differentiation between instrumental and communicative movements may be driven by observers' reactions and not reflect a spontaneous
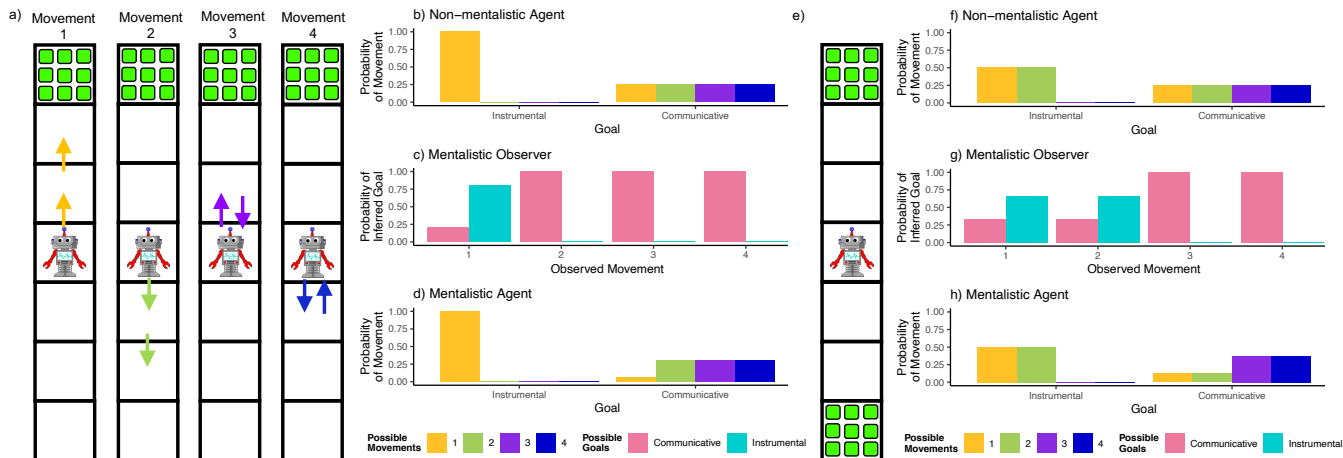
Figure 1: (a) A simple depiction of the possible movements considered by our model test. The bright green squares in the top of the grid represent instrumental goals. (b) The probability of selecting different movements (y-axis) as a function of goal type (x-axis) for a non-mentalistic agent. Bar color represents movement types as shown in panel a. (c) The probability that a mentalistic observer infers different goals when watching the non-mentalistic agent. Pink bars represent the probability of inferring a communicative goal and cyan bars represent the probability of inferring an instrumental goal. (d) The probability that a mentalistic agent will select different movements (y-axis; see panel a) as a function of goal type (x-axis). (e) A depiction of an alternative environment in which there are instrumental goals in both possible directions. The addition of a second area of instrumental goals leads to changes in (f) the probability that the non-mentalistic agent will select different movements, (g) the probability that the mentalistic observer will infer different goals, and (h) the probability that the mentalistic agent will select different movements.

motivation on the part of the communicator to ensure that their behavior was legible. We draw inspiration from these earlier signaling games to introduce a new paradigm that requires communicators to create signals without observer feedback and allows us to manipulate the distribution of non-communicative goals in the environment to test if signalers react to this distribution accordingly.

To determine whether people account for their observers' goal inferences when creating novel signals, we first formalize this inferential account of communication with a computational model of social reasoning. This framework forms the basis of our predictions about how agents should create communicative movements as a function of possible instrumental goals in the environment. We then present two experiments that support this inferential account of communicative signaling. Participants had to communicate through movement while also using movement to accomplish instrumental goals. In Experiment 1, we examined whether people take advantage of the skewed distribution of instrumental goals to spatially differentiate their communicative movements from instrumental movements. In Experiment 2, we remove the spatial signaling solution to test whether people use repetitive sequences to signal when instrumental goals are uniformly distributed.

Computational Framework

The objective of this computational framework is to derive qualitative predictions about how communicators should signal if they are choosing their movements to help observers infer their communicative goal. Crucially, this paper's main contribution is not the model itself, but rather the experiments that provide evidence of the predicted signaling strategies. The goal of this model is therefore to generate qualitative signaling strategies, rather than specific quantitative predictions (although we return to the importance of further quantitative work in the discussion).

Our computational framework (model code available at https://osf.io/mqk3d/) is structured around recursive social inference, which has been shown to be successful in modeling goal-inference and communication (Frank & Goodman, 2012; Shafto, Goodman & Griffiths, 2014; Yoshida, Dolan & Friston, 2008). Under this framework, agents choose their communicative movements and observers infer goals for those movements as determined by the recursive relation:

$$P_{agent}(M|G_c) \propto P_{observer}(G_c|M)P(M)$$

and

$$P_{observer}(G|M) \propto P_{agent}(M|G)P(G)$$

where $P_{agent}(M|G_c)$ is the probability that the agent will perform movement, M, given their communicative goal G_c, and $P_{observer}(G|M)$ is the probability that the observer will infer goal, G, after observing movement M.

We ground this recursive model at two levels with a non-mentalistic agent. This non-mentalistic agent is assumed to take the shortest path with likelihood 1 (and all other movements have likelihood 0) when acting towards an instrumental goal, and is equally likely to choose any movement when acting towards a communicative goal. The mentalistic observer then uses these likelihood functions to

reason about the non-mentalistic agent, and to infer the goal of the observed movement. Finally, the mentalistic agent reasons over this model of the mentalistic observer to decide how to act. Throughout, we use a uniform prior over goal type (instrumental or communicative).

Model Results

Simulations using this computational framework reveal how our account predicts changes in signaler strategies based on the spatial layout of non-communicative goals in the environment. To test our framework, we considered a simple two-dimensional world and all movements comprised of exactly two steps (n=4; Figure 1a). These four movements capture different types of actions that an agent can generate. We began by first considering a situation where all the potential goals are clustered in a single location (Figure 1a). In a case like this, a non-mentalistic agent is equally likely to use any of the four movements to communicate (Figure 1b). A mentalistic observer would therefore identify movement towards the goal area as likely to be instrumental, and the remaining movements as communicative (Figure 1c). A mentalistic agent, knowing the inferences that a mentalistic observer would make, then decides to shift their communicative movements to the ones that are not efficiently directed at the instrumental goals (Figure 1d).

Thus, recursively reasoning about the inferences of their observer should drive communicators to use movements that are unlikely to be produced while pursuing instrumental goals. While this category of solutions was derived from a simple goal space, it is easy to see how this principle would generalize to both real-world movements (consider, e.g., a body position that is unlikely to ever be generated when acting towards the world, such as a thumbs up) and more complex situations in a grid-world.

In some cases, the space of instrumental goals may not be clustered in a way that allows agents to signal by moving away from those goals. To explore the effects of this goal configuration, we modified the environment so that all directions led towards instrumental goals (Figure 1e). Again, in this scenario, a non-mentalistic agent is equally likely to use any of the four movements to communicate (Figure 1f). However, since moving in either direction is now consistent with instrumental goals, a mentalistic observer would identify the movements that directly travel towards either of the goal regions as likely to be instrumental. Therefore, the two movements that retrace themselves back to their origin are likely to be inferred as communicative (Figure 1g). As a result, a mentalistic agent prefers to communicate through the movements that are repetitive since those movements are inconsistent with efficient instrumental action (Figure 1h).

These results show how agents who are motivated to make their communicative action understood should change their behavior in response to the distribution of goals in the environment. We test these predictions in two experiments using goal distributions that are structurally similar to the ones used in our computational framework.

Experiment 1

In Experiment 1, we test whether the spatial distribution of instrumental goals leads people to gesture in regions that are unambiguously non-instrumental. To begin to answer this question, we presented participants with an environment where the instrumental goals have a skewed distribution (qualitatively similar to the set-up in Figure 1a).

Methods

Participants 80 participants from the US and UK (as indicated by their IP addresses) were recruited through the Prolific research platform. Two participants were excluded for failing an attention check, giving us a final sample size of 78 participants.

Stimuli The study was implemented using the p5.js (https://p5js.org) and jsPsych (de Leeuw, 2015) libraries for JavaScript. Experiment 1's stimuli consisted of a virtual environment where participants could move a robot's arm to touch different buttons (Figure 2a). The arm of the robot continuously followed the participant's mouse with a scaled delay proportional to seven percent of the distance between the arm and the mouse.

Buttons were arranged in a grid on the top half of the screen. During gameplay, different buttons that needed to be pressed would blink (one at a time) while a progressively faster beeping sound played. Once the participant moved the robot's arm over the button, another randomly-selected button was activated and the beeping noise reset. To enable participants to move the arm fluidly, they did not need to click on the buttons to score points; participants only had to touch the button with the arm.

Procedure First, participants read a cover story in which they learned that they would be working together with a partner (who they were led to believe was another online participant) to operate a spaceship. To score points, the participant had to move the robot's arm over the buttons to make the spaceship fly, while their partner had to defend the spaceship from aliens whenever an alien alert occurred (i.e., a warning sound played and the outside of the window flashed red). However, their partner could not hear or see the alert. The participant was told that their partner could see their robot's arm movements and were instructed to signal their partner whenever an alert occurred. In the directions, participants were told that their partner knew that their buttons were located in the top half of their screen, but could not see them. To further clarify the situation, the participant was shown that their "partner's" screen would show their robot's movements but not the buttons. This set-up the expectation that the partner knew approximately where instrumental goals were located.
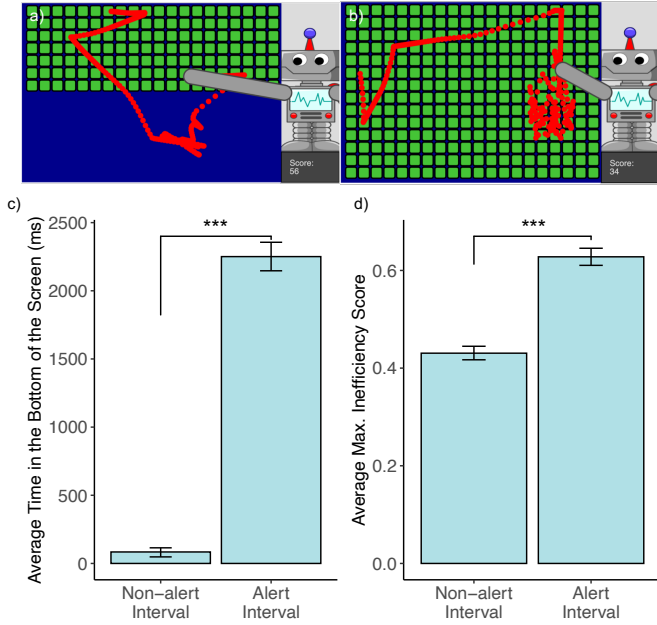
Figure 2: Images of the game set-ups for (a) Experiment 1 and (b) Experiment 2. The green squares are the buttons. The red lines have been added in as examples to show how two participants moved during alert intervals. (c) Results from Experiment 1. The bars depict the average number of milliseconds spent in the bottom half of the screen. (d) The bars in Figure 2d depict the average maximum inefficiency scores for movements. Error bars show 95% bootstrapped confidence intervals.

For each button pressed by the participant, they received one point. Participants were told that at the end of the round, their points would be multiplied by the number of times that their partner correctly responded to an alien alert. Since the participants were led to believe that their partner's success was contingent upon their signaling behavior, this scoring system incentivized participants to both push buttons and signal their partner.

Each participant played the game for five minutes and heard 12 alien warnings. To ensure that there was sufficient time between each warning message, the five-minute game was broken down into ten-second intervals and the warnings were played at the beginning of a random subset of the intervals. To allow participants to adjust to the movement of the robot's arm and the task, there were never alerts during the first two intervals. Participants were told in advance that there would be no alerts for the first 20 seconds and that their partner knew this as well.

After the game, participants were asked whether or not they signaled to their partner. If participants said no, then they were asked to indicate why they did not signal to their partner from a preset list of explanations (e.g., "I didn't know how", "I didn't think it mattered"). Finally, all participants were asked to explain their strategy.

Results and Discussion

Our account predicts that participants should choose to signal on the bottom half of the screen. Therefore, participants should have spent more time in the bottom half of the screen during the intervals when there were alien alerts (*alert intervals*) relative to baseline intervals when there were no alien alerts (*non-alert intervals*). We analyzed the time spent in the bottom half of the screen using a linear mixed effects model predicting time based on interval type (alert vs. non-alert) with random slopes and intercepts for participant (using the maximal model that converged throughout all analyses; Barr, Levy, Scheepers & Tily, 2013). As predicted, participants spent significantly more time in the bottom half of the screen during alert intervals compared to non-alert intervals ($\beta_{\text{Interval Type}}= 2166.53$, $p<.001$; Figure 2c).

In addition to signaling in an area unconfounded with instrumental goals, it is possible that people signaled by changing the dynamics of their movements relative to their instrumentally goal-directed motion. To test this, we also quantified the inefficiency of the movements during alert intervals and non-alert intervals. When participants were not signaling their partner, we assumed that they would be moving as directly as possible from button to button. Therefore, we created an *inefficiency score* to quantify the amount of unnecessary movement that participants made while moving from button to button:

$$InEff(p) = 1 - \frac{d^\star(p)}{d(p)}$$

where $d(p)$ is the actual distance travelled and $d^\star(p)$ is the shortest distance between the start- and end-points. Therefore, movements that form a straight, maximally efficient path between the start and end buttons will have inefficiency=0 and the inefficiency score will increase as more unnecessary distance is added to the movement. If participants signal their partner through additional movements, then they should stop pushing buttons, move inefficiently to signal, and then resume pushing buttons, thus yielding one movement per alert interval with a high inefficiency score.

We selected the maximum inefficiency score per interval and compared the maximum inefficiency scores between alert and non-alert intervals using a mixed effects model with random slopes and intercepts for participant and random intercepts for interval. As predicted, participants made movements with significantly higher maximum inefficiency scores during alert intervals compared to non-alert intervals ($\beta_{\text{Interval Type}}= 0.196$, $p<.001$; Figure 2d). The average maximum inefficiency score was also greater than zero in the non-alert intervals ($M = 0.43$), suggesting that participants were not perfectly efficient while traveling from button to button. However, this is unsurprising since the analyses above were conducted using the position of the robot's arm rather than the participant's mouse. Because the arm followed the participant's mouse with a scaled delay,

participants often produced unintentional deviations in their movements, especially when moving quickly. Crucially though, the maximum inefficiency scores for alert intervals were significantly higher than for non-alert intervals, suggesting that participants were incorporating additional inefficient deviations into their movements while signaling to their partner.

To further explore the strategies of participants, a coder categorized each player's signaling strategy. Consistent with our time-based analysis, 69.3% of all signals were performed exclusively in the bottom half of the screen and 8.9% of signals moved through both halves of the screen. Of the signals in the bottom half of the screen, 84.6% moved into the bottom half and then made repetitive movements with the arm (most frequently using oscillations or small circular movements), while 6.6% moved into the bottom half and held the arm still—creating a pause in their movement.

These results suggest that people can spontaneously account for the goals that may be inferred by their communicative partners when gesturing.

Experiment 2

Experiment 1 provides evidence that people are sensitive to the spatial distribution of other goals that an observer could infer. As such, people prefer to signal where there are no goals. In Experiment 2, we remove this strategy such that movement in all areas on the screen is confounded with instrumental goals. If people can flexibly account for the inferences that will be made by an observer, then they should differentiate the manner of their movement when communicating from their instrumentally-directed movements. Specifically, we predict that participants will use repetitive movements to create novel communicative signals.

Methods

Participants 80 participants from the US and UK (as indicated by their IP addresses) were recruited through the Prolific research platform. Two participants were excluded for failing to reach a pre-registered minimum game score and one participant was excluded for a data-saving failure, giving us a final sample size of 77 participants.

Stimuli The study design was nearly identical to Experiment 1, except the buttons were positioned in a grid-like formation throughout the player's entire screen (Figure 2b).

Procedure The cover story, game phase, and post-test questions were nearly the same as in Experiment 1. However, the participant was told that their partner knew that the buttons were distributed throughout their entire screen, instead of only in the top half of their screen.

Results and Discussion

If participants are signaling through repetition, then they should be moving more during alert intervals, thus creating more inefficient movements. Figure 3a shows the average maximum inefficiency scores in alert and non-alert intervals. Participants made movements with significantly higher inefficiency scores during alert intervals compared to non-alert intervals ($\beta_{Interval\ Type}$= 0.212, p<.001; Figure 3a) as analyzed using a linear mixed effects model with random slopes and intercepts for participant.

Blind coders watched replays of participants' movements without the buttons or alerts and counted the number of repetitions made during each interval. Agreement between the coders was high (Pearson's correlation r = 0.967, p < 0.001), so we averaged together the two coders' repetition counts for each interval. We analyzed this averaged number of repetitions per interval using a linear mixed effects model predicting repetitions based on interval type (alert vs. non-alert) with random slopes and intercepts for participant and interval. Consistent with our predictions, participants performed significantly more repetitions during alert intervals compared to non-alert intervals ($\beta_{Interval\ Type}$= 9.57, p<.001; Figure 3b).
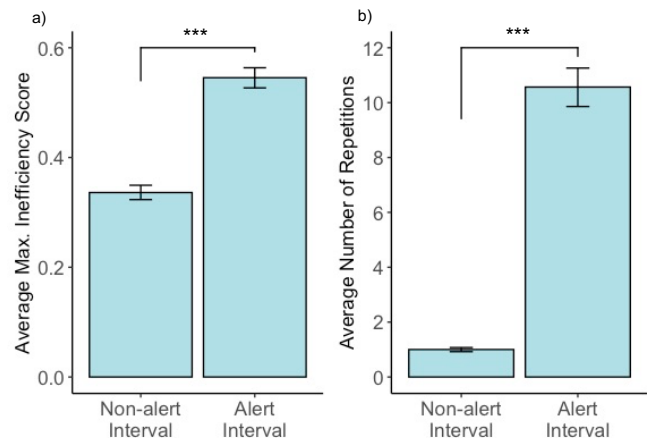

Figure 3: (a) The bars depict the average maximum inefficiency scores for movements. (b) The bars depict the average number of repetitions performed. Error bars show 95% bootstrapped confidence intervals.

General Discussion

Humans communicate with a wide range of gestures that observers must quickly recognize in order to engage in successful seamless social interactions. Our work demonstrates that people are sensitive to this pressure when gesturing and shape their gestures to help observers detect their communicative goal. In particular, we show that people generate gestures that are distinct from instrumentally goal-directed movements. When instrumental goals are confined to one area, people prefer to signal by moving in an area that is unconfounded with those other possible goals (Experiment 1). However, when an environment is saturated with instrumental goals such that

no one area is unconfounded, people change the manner of their movement to create communicative signals (Experiment 2). Together, these studies show that people flexibly account for the goal inferences that are likely to be made by observers and subsequently create communicative movements to guide those inferences accordingly.

Our findings are aligned with the qualitative predictions made by our computational framework which formalized gesture as a social process where actors are motivated to ensure that their behavior is understood. This framework is equivalent to the Rational Speech Act (RSA) framework and the cooperative communication framework developed to understand communication and pedagogical interactions (Frank & Goodman, 2012; Shafto et al., 2014; Wang et al., 2019). A key difference, however, is that these models are typically embedded in a situation where communicative intent is assumed to be known. Our work instead focused on the process of inferring whether an agent is communicating or not. As such, our work suggests that recursive social reasoning as formalized in RSA not only underpins how we reason about communicative meaning, but also how we distinguish communicative movement from non-communicative movement. Here we derived qualitative predictions from our framework; our future work will quantitatively test the model's predictions against participant behavior in a more controlled signaling game.

One limitation of our model is that it assumed that the observer watched the entire movement before making inferences about the underlying goal. If the observer was making inferences as the movement unfolded, then our predictions for both observers and agents might shift. For example, our model deemed repetition to be a good signaling strategy, but observers could understand your communicative goal faster if the initial movement is away from any instrumental goals (e.g., Movement 4 in Figure 1a) rather than towards (e.g., Movement 3 in Figure 1a). Indeed, related research has found that observers are more confident inferring communicative goals for movements that quickly reveal that they are not world-directed (Royka et al., 2018, Royka et al., in review).

A further limitation of our model was that it had no access to other features of movement such as speed or the sharpness of directional changes. The increase in the inefficiency of movement in Experiment 1 suggests that people change other features of their movements in addition to those predicted by our model. It is possible that these features further reflect a motivation to ensure that observers recognize the movement's goal. Future work will investigate this possibility.

The experiments presented here are consistent with work showing that people virtually communicating with a real partner use inefficient movements to signal (De Ruiter et al., 2010; Scott-Phillips et al., 2009). One critical difference from this earlier work, however, is that our signalers did not effectively receive any feedback from their "partner." Our results therefore show that, even in the complete absence of an observable observer or feedback, people still create communicative movements that are tailored for the goal inferences of the implied observer. Thus, in signaling interactions, communicators spontaneously consider the likely goal inferences of observers when deciding how to communicate.

Physical movements are the basic ingredients not only for signs in conventional sign languages, but also for gestural communication (Goldin-Meadow & Brentari, 2017). Although gesture is a term used for a variety of non-verbal behaviors, many of which are involuntary and arguably non-linguistic (Goldin-Meadow & Brentari, 2017), here we tested our hypotheses by asking participants to create *emblematic* gestures. *Emblems* are intentionally produced gestures that do not systematically combine with other gestures (Coppola & Senghas, 2017) and can be produced with or without accompanying speech (Ekman & Friesen, 1969). Although emblems are typically conventionalized (Kendon, 1997), our findings suggest that, in the absence of convention, the forms of emblems will emerge as a function of the constraints imposed by goals in the environment. As such, real-world conventional gestures may share a common structure that allows them to quickly reveal their lack of instrumental purpose.

Other types of gestures, however, may have different features that impact the extent to which they are shaped to help observer inferences. For example, illustrators are gestures that accompany speech and serve to depict parts of the speaker's message (Ekman & Friesen, 1969). Thus, because illustrators are already occurring in the context of a communicative interaction, illustrators may not need to reveal their communicative goal as much as other gestures, such as emblems, which can occur in isolation.

Moreover, in more naturalistic contexts, there are likely additional factors that shape gestures, but fall outside of the scope of the framework presented here. For example, social and pedagogical interactions often involve representational gestures that depict either an object or an action (Brand et al., 2002; Kelly, Byrne & Holler, 2011; Özyürek, 2002) as well as deictic (i.e., "pointing"-like) gestures (Bangerter & Louwerse, 2005; Gliga & Csibra, 2009). We look forward to investigating in future work how pressures for iconicity and spatial reference may compete with the pressure to be distinct from other types of goal-directed movement when creating novel communicative gestures.

As social agents moving in a highly complex world, humans must make inferences about the intentions of other agents. Here we show that—perhaps even more impressively—humans are able to shape their own actions to make their intentions more obvious to observers thus easing observers' inferential burden. As such, signaling and inferring communicative intent can be understood as a set of recursive inferences based on the assumption that agents create signals that will quickly and unambiguously convey their communicative goal.

Acknowledgements

References

Akhtar, N., Carpenter, M., & Tomasello, M. (1996). The role of discourse novelty in early word learning. *Child Development.*

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language.*

Bangerter, A., & Louwerse, M. M. (2005). Focusing attention with deictic gestures and linguistic expressions. In *CogSci Proceedings.*

Behne, T., Carpenter, M., & Tomasello, M. (2005). One-year-olds comprehend the communicative intentions behind gestures in a hiding game. *Developmental Science.*

Brand, R. J., Baldwin, D. A., & Ashburn, L. A. (2002). Evidence for 'motionese': modifications in mothers' infant-directed action. *Developmental Science.*

Coppola, M., & Senghas, A. (2017). Is it language (yet)? The allure of the gesture-language binary. *Behavioral and Brain Sciences.*

Csibra, G., Bíró, S., Koós, O., & Gergely, G. (2003). One-year-old infants use teleological representations of actions productively. *Cognitive Science.*

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *TiCS.*

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a web browser. *Behavior Research Methods.*

De Ruiter, J. P., Noordzij, M. L., Newman-Norlund, S., Newman-Norlund, R., Hagoort, P., Levinson, S. C., & Toni, I. (2010). Exploring the cognitive infrastructure of communication. *Interaction Studies.*

Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica.*

Frank, M. C., & Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. *Science.*

Galantucci, B. (2005). An experimental study of the emergence of human communication systems. *Cognitive Science.*

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition.*

Gliga, T., & Csibra, G. (2009). One-year-old infants appreciate the referential nature of deictic gestures and words. *Psychological Science.*

Goldin-Meadow, S. (1999). The role of gesture in communication and thinking. *TiCS.*

Goldin-Meadow, S., & Brentari, D. (2017). Gesture, sign, and language: The coming of age of sign language and gesture studies. *Behavioral and Brain Sciences.*

Jara-Ettinger, J., Gweon, H., Schulz, L. E., & Tenenbaum, J. B. (2016). The naïve utility calculus: Computational principles underlying commonsense psychology. *TiCS.*

Kelly, S., Byrne, K., & Holler, J. (2011). Raising the ante of communication: evidence for enhanced gesture use in high stakes situations. *Information.*

Kendon, A. (1997). Gesture. *Annual Review of Anthropology.*

Özyürek, A. (2002). Do speakers design their cospeech gestures for their addressees? The effects of addressee location on representational gestures. *Journal of Memory and Language.*

Royka, A., Aboody, R., & Jara-Ettinger, J. (2018). Movement as a message: inferring communicative intent from actions. In *CogSci Proceedings.*

Royka, A., Aboody, R., Chen, A., & Jara-Ettinger, J. (in review). Shaped to Communicate: Movements that efficiently reveal their lack of external goals are seen as communicative.

Scott-Phillips, T. C., Kirby, S., & Ritchie, G. R. (2009). Signalling signalhood and the emergence of communication. *Cognition.*

Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology.*

Shafto, P., Goodman, N. D., & Griffiths, T. L. (2014). A rational account of pedagogical reasoning: Teaching by, and learning from, examples. *Cognitive Psychology.*

Skerry, A. E., Carey, S. E., & Spelke, E. S. (2013). First-person action experience reveals sensitivity to action efficiency in prereaching infants. *PNAS.*

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition.* Harvard University Press.

Trujillo, J. P., Simanova, I., Bekkering, H., & Özyürek, A. (2018). Communicative intent modulates production and comprehension of actions and gestures: A Kinect study. *Cognition.*

Wang, P., Wang, J., Paranamana, P., & Shafto, P. (2019). A mathematical theory of cooperative communication. *arXiv*

Yoshida, W., Dolan, R. J., & Friston, K. J. (2008). Game theory of mind. *PLoS Comput Biol.*