Communication with Surprise – Computational and Neural Mechanisms for Non-Verbal Human Interactions

Tatia Buidze[1,*], Tobias Sommer[1], Ke Zhao[2,3], Xiaolan Fu[2,3], Jan Gläscher[1]

[1] Institute of Systems Neuroscience, University Medical Center Eppendorf, Hamburg University, Germany
[2] State Key Laboratory of Brain and Cognitive Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, 100101, China
[3] University of Chinese Academy of Sciences, Beijing, 100049, China
[*] Correspondence: ta.buidze@uke.de

Abstract

Communication, often grounded in shared expectations, faces challenges without common linguistic backgrounds. Our study explores how people instinctively turn to the fundamental principles of the physical world to overcome communication barriers. Specifically, through the Tacit Communication Game, we investigate how participants develop novel strategies for conveying messages without relying on common linguistic signals. We developed a new computational model built from the principle of expectancy violations of a set of common universal priors derived from movement kinetics. The model closely resembles the Sender's messages, with its core variable – the information-theoretic surprise – explaining the Receiver's physiological and neural responses. This is evidenced by a significant correlation with the pupil diameter, indicating cognitive effort, and neural activity in brain areas related to expectancy violations. This research highlights the adaptability of human communication, showing how surprise can be a powerful tool in forming new communicative strategies without relying on common language.

Introduction

Expectations and their subsequent violations play a fundamental role in shaping human perception and experience. While this principle has its place in learning, where prediction errors serve as a teaching signal improving future expectations[1], it also manifests distinctively in communication: a sender directs the receiver's attention to the crucial parts of a message by diverging from established expectations, thus creating a psychological state of surprise. Consider humor as a prime example of this mechanism in action. A joke's effectiveness hinges on its setup, aligning the audience with a certain trajectory of expectation[2]. These expectations are grounded in a broader, often unspoken social contract of common knowledge (sometimes called schemas[3]) – a collective understanding that streamlines cognitive processing and facilitates collaboration[4]. The punchline

subverts this expectation by presenting an unexpected resolution that, ideally, induces laughter[5,6]. The surprise that fuels humor arises when a mismatch between these collective schemas providing a canvas of expectations (a so-called "prior") and new, unexpected information occurs. Thus, the effective use of surprise as a communicative tool relies on prior expectations, which – in the case of humor – are grounded in the shared cultural and linguistic frameworks.

When people lack a shared language, they often revert to using gestures and facial expressions, as might be the case when interacting with individuals from different cultures who understand basic non-verbal cues but do not share a common spoken language. But what happens when neither verbal nor these basic non-verbal communication methods are available?

Here, we propose that in such scenarios, humans instinctively revert to even more basic, universal properties of the physical world (e.g., movement direction or velocity), to form a new kind of shared understanding. This is not about discovering new universal principles; rather, it is about repurposing universally understood properties and principles of the physical world in unique ways to build a new common communicative system through repeated social interactions. Once these common grounds are established, they can serve as the prior from which senders can deviate to create surprise and direct the receiver's attention to essential information. Consider a fundamental principle from movement kinetics that a moving object will continue its path if no other external force is applied[7]. When this principle forms the core of a new common language, the trajectory along a straight line not only signifies a direction but also embodies a shared expectation. Deviations from this path thus transform into communicative acts, signals endowed with meaning during social interactions. These shifts can signal anything from approaching threats to hidden rewards, challenging individuals to interpret and adapt to these evolving communicative cues.

In this study, we explored the formation of novel common communicative systems within the framework of the Tacit Communication Game (TCG) and examined how creating surprising events by strategically deviating from these newly established norms are used for effective communication between players. The TCG is an experimental task designed to elicit novel communicative messages through spatial trajectories[8–11]. The task is conducted on a square game board featuring a different goal position for each player, namely, the Sender (she/her) and the Receiver (he/his). However, only the Sender is aware of both goal positions (see Fig.1a). The Sender's objective is to create a trajectory (the "message") from her starting position to the two goal states that conveys to the Receiver where his goal state is. Here, we hypothesized that in the TCG, senders who are prohibited to use linguistic or gestural cues will turn to the universally shared, basic knowledge about the physical properties of the world and communicate by deviating from these concepts. To test this, we developed a novel computational model, the Surprise Model, for the Sender's message creation in the TCG. The model utilizes intuitive priors based on principles of movement kinetics and goal orientation, and constructs messages based on

maximizing the information-theoretic surprise (a measure quantifying the psychological state of surprise) at the Receiver's goal state[12].

We hypothesized that our Surprise Model would accurately replicate the behavioral data of participants and that the core element of the model - the information-theoretic surprise - would be traceable in the participant's physiological and neural responses. Specifically, considering the role of the Locus coeruleus in detecting unexpected changes in environmental dynamics and its impact on pupillary dilation responses (PDRs)[13–15], we expected a correlation between the model's surprise and the corresponding PDRs. Additionally, we hypothesized that if the Receivers used the surprise signal in the message to decode their goal state, this would elicit a neural representation of surprise. We therefore anticipated a positive correlation between model-derived surprise and neural activity in fronto-central EEG sensors located above the anterior cingulate cortex (ACC), a brain region conventionally associated with the detection of expectancy violations[16,17].

Our findings offer compelling support for these hypotheses in two independent samples of human participants. Messages generated by the Surprise Model closely resemble human Sender-created messages. Furthermore, we identified a notable impact of model-derived step-by-step surprise on both PDR and EEG data across all participants. Specifically, we observed a positive correlation between model-derived surprise values and neural activity in fronto-central sensors over the ACC and the Frontopolar Cortex.

Results

In this section, we first provide a brief overview of the Tacit Communication Game, followed by the a model-free analysis of behavioral data. Subsequently, we introduce our computational models. Building upon this foundation, we proceed to validate the computational models and explore physiological observations pertinent to non-verbal communication within the TCG framework.

Experimental Task

Tacit Communication Game. The Tacit Communication Game (TCG) is a cooperative game designed for two players to establish a non-verbal communication. The game involves a Sender who knows both players' intended destinations and must create a spatial path (the "message") to indicate the Receiver's target location. The game unfolds over five primary steps (see Fig.1a). Initially, the complete goal configuration, including the starting location, the Sender's goal, and the Receiver's goal, is fully visible to the Sender but not to the Receiver. Specifically, the Sender has access to all three elements of the configuration, whereas the Receiver is only presented with the starting location and the Sender's goal, leaving the Receiver's own goal undisclosed. Next, the Sender not only navigates her token towards her destination but also encodes the location of the Receiver's goal within her movements, thus creating a 'message' that informs the Receiver of his

goal location while simultaneously advancing towards her own. Following this, both participants observe the trajectory created by Sender. The Receiver then interprets the path and moves his token where he thinks his goal state is. A successful trial occurs when the Receiver reaches the correct location, granting both players an equal reward. The reward is quantified as 10 points reduced by the number of steps the Sender utilized to transmit the message.

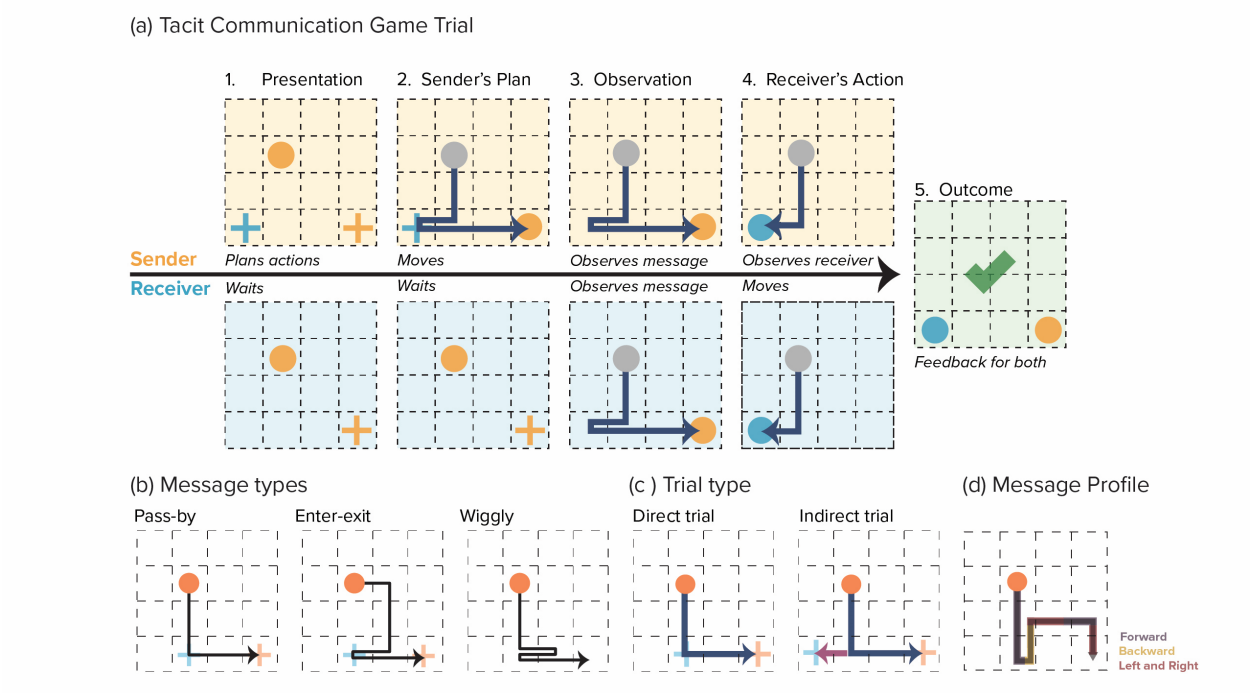


Figure 1 | Trial Steps and Goal Configurations in the Modified TCG
(a.1) Presentation: Our version of TCG begins with players being presented with the relevant goal configuration. The presented goal configuration for the Sender always includes three elements: the starting location, the Receiver's goal location, and the Sender's own goal location. For the Receiver, it comprises the starting location and the Sender's goal location. Both the Sender and the Receiver start at the same location on the grid. **(a.2) Sender's Plan:** Sender designs a message by moving her token first to the Receiver's goal and eventually to her goal. **(a.3) Observation:** The Receiver and the Sender both examine the path crafted by the Sender. This step was incorporated to facilitate the simultaneous recording of EEG data from both participants under uniform visual stimulus conditions and to equalize the step length, preventing a "pause" strategy to indicate the Receiver's goal location (see supplementary text and Figure S1 for more details). **(a.4) Receiver's Action:** Receiver moves towards the inferred goal. **(a.5) Outcome:** Success is displayed if the Receiver's choice is correct, indicated by a green checkmark. **(b) Message Types:** Three message types created by the participants. **(c) Trial Types**: An illustration of two different trial types. **(d) Message Profile:** An illustration of a message profile featuring three characteristics.

Trial Types. The TCG consists of two different types of goal configurations: *'Direct'* and *'Indirect'* (Fig.1c). Direct goal configurations can be communicated by the shortest path from the starting location to the Sender's goal location, while indirect goal configurations require a deviation from the direct path. Direct Trials are inherently more difficult for the Receiver because his goal state

is not “standing out” from the direct path to the Sender’s goal state. In contrast, when the Receiver’s goal is not on the direct path (i.e., and Indirect Trial) the Sender must step sideways to enter the Receiver’s goal state. This difference also prompts the creation of different Message Types (Fig.1b).

Datasets. We collected behavioral data from two distinct samples (subsequently referred to as dataset 1 (D1) and dataset 2 (D2) consisting of 29 and 31 pairs of participants, respectively). Eye-tracking and EEG hyperscanning data were exclusively obtained in D2 experiments. All pairs played 120 trials of the TCG. The two datasets differed in the proportion of direct and indirect trials: D1 had 66% indirect and 34% direct trials, while D2 had 76% indirect and 24% direct trials.

Model-Free Analysis of Behavioral Data

Using our experimental data, we developed different indices to describe the behavior of the Sender to gain a more detailed understanding of the cognitive processes involved in planning new communicative messages (For detailed results on D1 and D2, see Figures 3a and 4a, respectively). These indices are also used in posterior predictive checks to test whether the fitted model can generate simulated data that closely resembles the observed experimental data.

Message Types. We identified three most-common and distinct message types that Senders used to communicate their intentions to Receivers. These message types are illustrated in Figure 1b (see Methods for details).

Enter-Exit Message. Predominantly used by participants (*D1: M=63.82%, SE=5.00%; D2: M=54.95%, SE=6.6%*), the Enter-Exit Message involves Senders momentarily diverging from the direct route to the Receiver's goal before proceeding to their own goal, noted for its high success rates of 87% in D1 and 88% in D2.

Wiggly Message. Ranking as the second most common (*D1: M=21.47%, SE=4.67%; D2: M=26.65%, SE=5.91%*), this message features Senders navigating on the direct path, shifting between the Receiver's goal and the next state to create a zigzag pattern, achieving accuracy rates of 63% in D1 and 76% in D2.

Pass-By Message. The least frequently employed but still notable (*D1: M=14.70%, SE=3.66%; D2: M=18.41%, SE=4.95%*), Pass-By Messages involve Senders briefly passing through the Receiver's goal area on their way to their own goal, with the lowest observed accuracy rates of 42% in D1 and 29% in D2.

Our findings indicate that participants utilized all message types in both trial contexts, but with different frequencies. Notably, during the Direct Trial, there was a marked increase in the use of

Wiggly and Pass-By Messages, suggesting a preference among participants to remain on the direct route during these trials. Conversely, in the Indirect trial, the most used message type was the Enter-Exit message (see the second rows of Figures 3a and 4a for D1 and D2 results, respectively). A comprehensive breakdown of these findings is provided in Supplementary Tables S1b and S2b.

Message Profiles. The message profile details the number of movements in each direction, relative to the last movement into the current state. For instance, the Enter-Exit message profile depicted in Figure 1d includes movements Forward, Left, Right, and Backward. Each message type exhibited a specific profile. *Enter-Exit Messages* are characterized by frequent left/right movements, followed by forward and backward movements. *Wiggly Messages* have a higher frequency of backward movements (due to the zigzag pattern), followed by left/right and then forward movements. *Pass-By Messages* differ from the other message types by exclusively relying on left/right and forward movements, without incorporating any backward movement. Exact frequencies can be found in the Supplementary Material (Table S1c for dataset 1 and Table S2c for dataset 2).

Computational Models

In this section, we provide a narrative overview of our Surprise Model. The mathematical details can be found in the Methods section. The Surprise model consists of 3 distinct components: (i) the Movement model, detailing the step-by-step priors over all available actions, (ii) the State model, providing a prior guiding message creation toward the Sender's goal, and (iii) the Reward, providing an orientation to the Receiver's goal and an urgency toward shorter messages. The details of these components are described below. We tested the necessity of these components in a formal model comparison, in which we compared degenerate versions (Movement and State models only) against the Surprise model.

The *Movement model* defines priors for all possible movements in each state. These priors are based on a fundamental idea from movement kinetics: without an external force, a moving object is expected to continue its motion in the same direction where it came from. Any deviation from this direction has a lower prior probability and is therefore more surprising. This leads to the rule illustrated in Figure 2a, where $p\ (forward) > p(left) = p(right) > p\ (backward)$. The calculation of these subject-specific priors is governed by two estimable parameters, γ - the probability of a backward movement, and λ, an "odds-ratio" parameter that describes the ratio of priors for forward movement and leftward or rightward movement. Movement priors are dynamic and can change with every step along the message path.

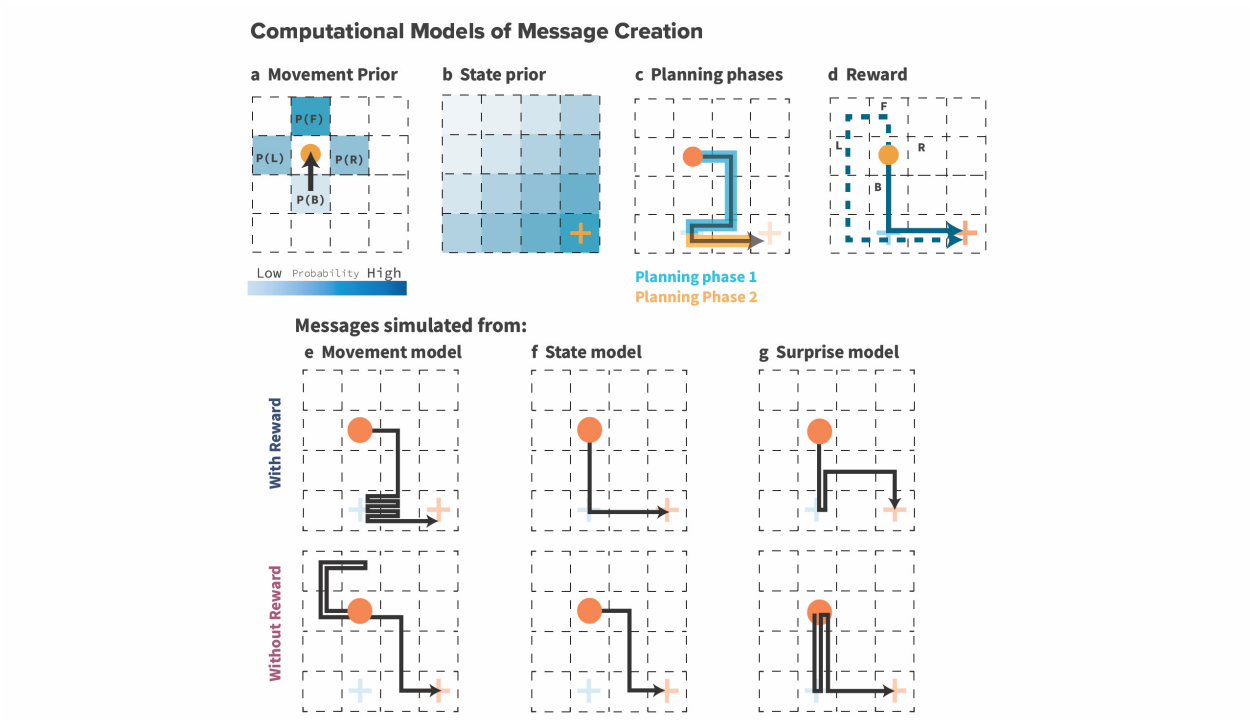


Figure 2 | Computational Models of Message Creation
(a) Movement Prior assigns highest probability to forward movement, then left/right, and least to backward. **(b) State Prior** conveys the sender's goal by reducing probabilities at each state with increased distance. **(c) Planning Phases**: During Phase 1, the Sender communicates the Receiver's goal; in Phase 2, the Sender proceeds to her own goal. **(d) Reward**: Illustrates higher rewards for backward movements, facilitating shorter paths to goals. **(e-g) Model Components**: Simulations with each model component alone reveal that the Movement model (e) emphasizes the Receiver's goal with surprising events, while the State model (f) focuses solely on efficient sender's goal achievement. The Surprise model (g) combines the efficiencies of the sender's goal achievement (State priors) with the emphasis on signaling for the Receiver's goals (Movement priors). Comparing messages with reward information (first row) to those without (second row) demonstrates that incorporating reward minimizes cost and efficiently signals the Receiver's goal. These messages also match human messages most closely.

The *State model* defines state priors, which provide a general sense of the goal orientation for the Sender and decrease in concentric circles from her goal according to an inverse power rule (Fig. 2b). Thus, moving along an increasing state prior will move the Sender straight toward her goal. The state prior is static for the duration of the trial, but changes with different goal configurations. Its shape is determined by a single slope α parameter (see Methods for details).

In the *Surprise model*, the Sender selects the step-by-step actions by intentionally violating expectations that are defined from both types of priors: *Movement* and *state* priors. The process of creating a message involves two distinct phases: (i) In the first phase of the message, Movement and State priors govern action selection from the start location to the Receiver's goal state. (ii) In the second phase of the message from the Receiver's goal state to the Sender's goal state, the

Sender aims to reach her goal state as quickly as possible (Fig. 2c). Here, in the second phase, the principle of selecting a surprising action is abandoned in favor of a policy that is solely governed by the state prior.

Rewards for each potential next action are calculated as the discounted cost (i.e., number of steps) to reach the Sender's goal state via the shortest path through the Receiver's goal state. Rewards provide an urgency to reach the Sender's goal quickly to retain the maximum points from the initial endowment. The discount rate is estimated with a single free parameter ε (see Methods for details). For example, Figure 2d showcases a scenario where the action of moving backward (B) is linked to a greater reward due to its ability to reach goal locations via the shortest path (depicted by the blue solid line), while other actions (F, L, R) incur higher costs (indicated by dashed line for action F).

The combination of the surprise from the Surprise model and the rewards yields an expected value over all available next actions given a specific planning horizon (see Methods for details). These expected values are filtered through a softmax function (parameterized with a choice stochasticity parameter τ) to render the final action probabilities for the next action.

Figure 2e-g illustrates the effect of each model component on message planning in an intuitive way. Simulations with only the Movement Model (Fig. 2e, top) reveal that the Sender will perform many zigzagging moves, because of the high surprise of the backward action. In contrast, the pure State model (Fig. 2f, top) creates a direct path message without surprising moves. The combination of both models (Surprise model) (Fig. 2g, top) generates messages comparable to the human-created Enter-Exit messages. The second row illustrates the effect of the different model components without the reward. The pure Movement and State models (Fig. 2e, f, bottom) generate similar message patterns, but fail to communicate the Receiver's goal. The Surprise model (Fig. 2g, bottom) generates 2-step wiggles, but without the reward component, this model is not advancing to the final Sender's goal efficiently. In conclusion, by combining Movement and State priors along with the surprise values and discounted future rewards, the Sender can effectively communicate the Receiver's goal location. Once this objective is accomplished, the state priors then guide the Sender towards taking the shortest path to reach her own goal location.

Model Fitting and Model Comparison

To determine which model variant most accurately reflects the observed behavioral data, we fitted the free parameters for each model by minimizing the negative log-likelihood of the model for the collected trial-by-trial data across subjects. The resulting model likelihoods, model frequencies[18] and protected exceedance probabilities (PEP)[19] are outlined in Table 1a for dataset 1 and table 1b for the dataset 2. Here, log-likelihood serves as a cost function for model fitting, with higher values indicating a more precise fit. Model frequency reflects the extent of a model's suitability across

different subjects, and PEP evaluates the probability of a model being the most appropriate, while also considering the potential that no model may ideally fit the data.

The Surprise model fitted the observed data best. Its corresponding model frequencies — 0.80 for Dataset 1 and 0.78 for Dataset 2 — suggest strong consistency in explaining the behavior across the majority of participants. This is further substantiated by the PEP; High PEP, close to 1 for both datasets, signals a strong statistical confidence that the Surprise model is the best model among the three variants tested. This measure of confidence also considers the null hypothesis that no model in the model space is likely across the population. In contrast, the Movement and State models exhibit significantly lower PEP, which translates to a lower chance of these models being the most fitting explanation for the datasets. This convergence of likelihood, frequency, and PEP distinctly positions the Surprise model as the most compelling model for capturing the nuances of the observed behavioral data.

Predictive Accuracy of the Surprise Model across Different Samples. Next, our objective was to evaluate the ability of the surprise model to accurately replicate participant behaviors across different samples. We fitted the Surprise model to the initial sample using hierarchical Bayesian estimation[18] and used the estimated group parameter values to simulate the message types and profiles for the subsequent sample. To quantitatively assess the congruence between actual and predicted frequencies, we employed Bayes factor analysis[20]. The majority of Bayes factors obtained underscored a notable similarity across samples, as evidenced by BF_{01} values ranging from 1.2 to 11.6 (for details, see Table S3). This analysis demonstrates that the surprise model is not merely the most suitable fit for the behavioral data, but also emphasizes its capability to accurately mirror behaviors across varied datasets.

a) Data set 1

Model	Log Likelihood	Model Frequency	PEP	Posterior Group Parameters				
				τ	ε	λ	γ	α
Surprise model	-60.06	0.80	1.00E+00	6.93	0.39	0.10	2.14	1.84
State model	-70.38	0.16	1.57E-04	2.90	0.39	-	-	1.11
Movement model	-96.57	0.05	5.00E-06	1.92	0.62	0.15	1.50	-

b) Data set 2

Model	Log Likelihood	Model Frequency	PEP	Posterior Group Parameters				
				τ	ε	λ	γ	α
Surprise model	-60.46	0.78	9.99E-01	6.80	0.31	0.11	2.03	1.83
State model	-69.48	0.22	7.45E-04	2.77	0.45	-	-	1.15
Movement model	-94.26	0.00	2.44E-140	2.50	0.70	0.15	1.98	-

Table 1 | Comparative Evaluation of Model Performances Across Two Datasets: Likelihoods, Frequencies, PEP (protected exceedance probabilities) and Posterior Group Parameters.

Model: Refers to the three tested computational models: Surprise, State, and Movement models. **Log Likelihood:** Represents the log-transformed probability of the data given the model, where higher values indicate a better fit of the model to the observed data. **Model Frequency:** Reflects the proportion of subjects for which a given model is most likely the best fit, with values closer to 1 indicating higher generalizability across the sample population. **PEP:** is the probability that a model is more likely than any other model, while also considering the possibility that no single model is a good fit for the data population. **Posterior Group Parameters:** estimated group values for parameters: ε (discount rate for planning horizon), λ (backward movement), γ (forward and left/right movement), α (slope of state prior), and τ (choice stochasticity).

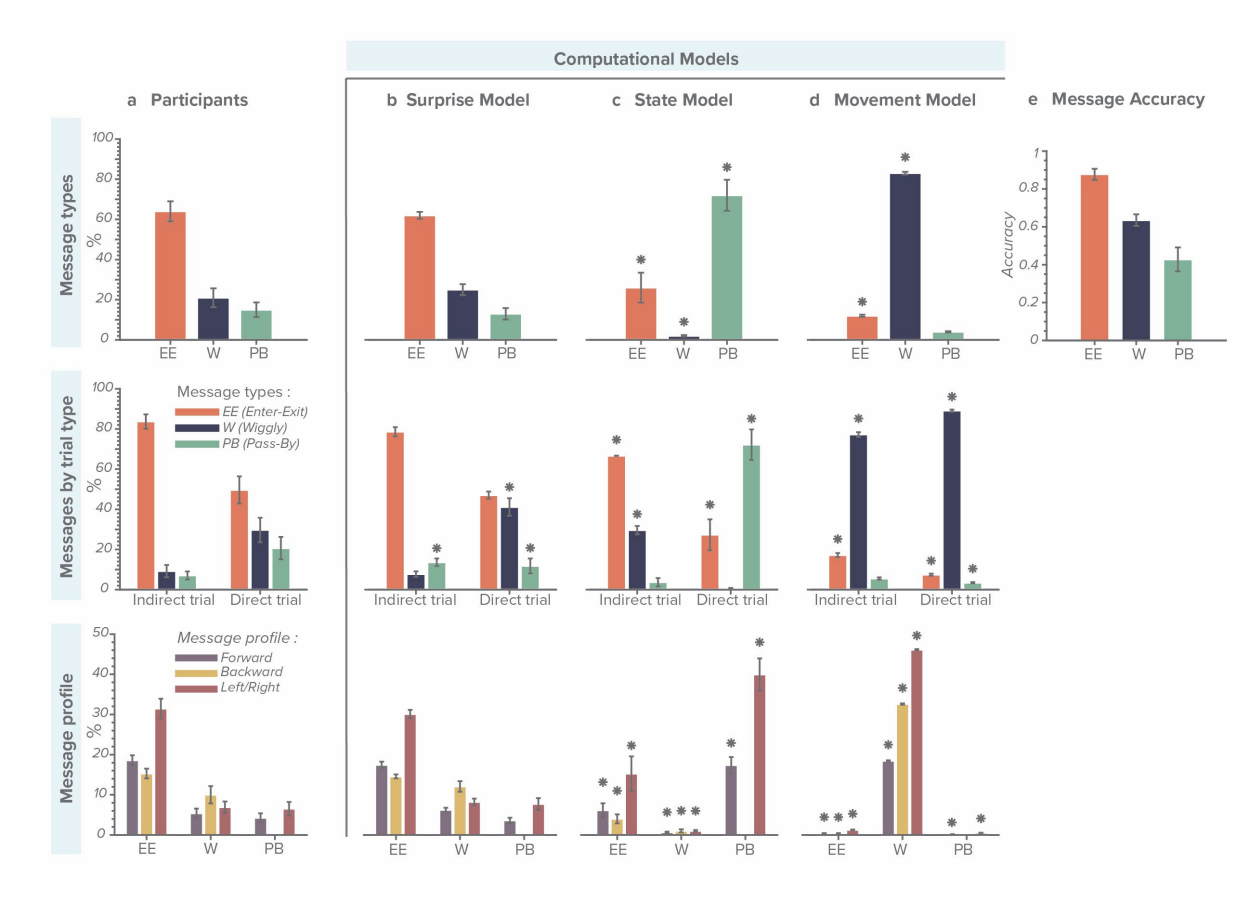


Figure 3 | Comparing Participant-Generated and Model-Simulated Messages for the Dataset 1
(a) Participants: Participants behavioral data categorized by message types, trial type, and message profile. **(b) Surprise Model** simulated data after model fitting exhibits similarity to participants data (*$BF_{01}>1$).* **(c) State Model** and **(d) Movement Model** simulated data reveal differences from participant's data **(**asterisks indicate *$BF_{01}<1$).* All error bars represent means +/- SEM. **(e) Message Accuracy:** Accuracy as the Receiver's success in finding his goal state. Enter-Exit message has the highest accuracy rate, followed by Wiggly and Pass-By messages.

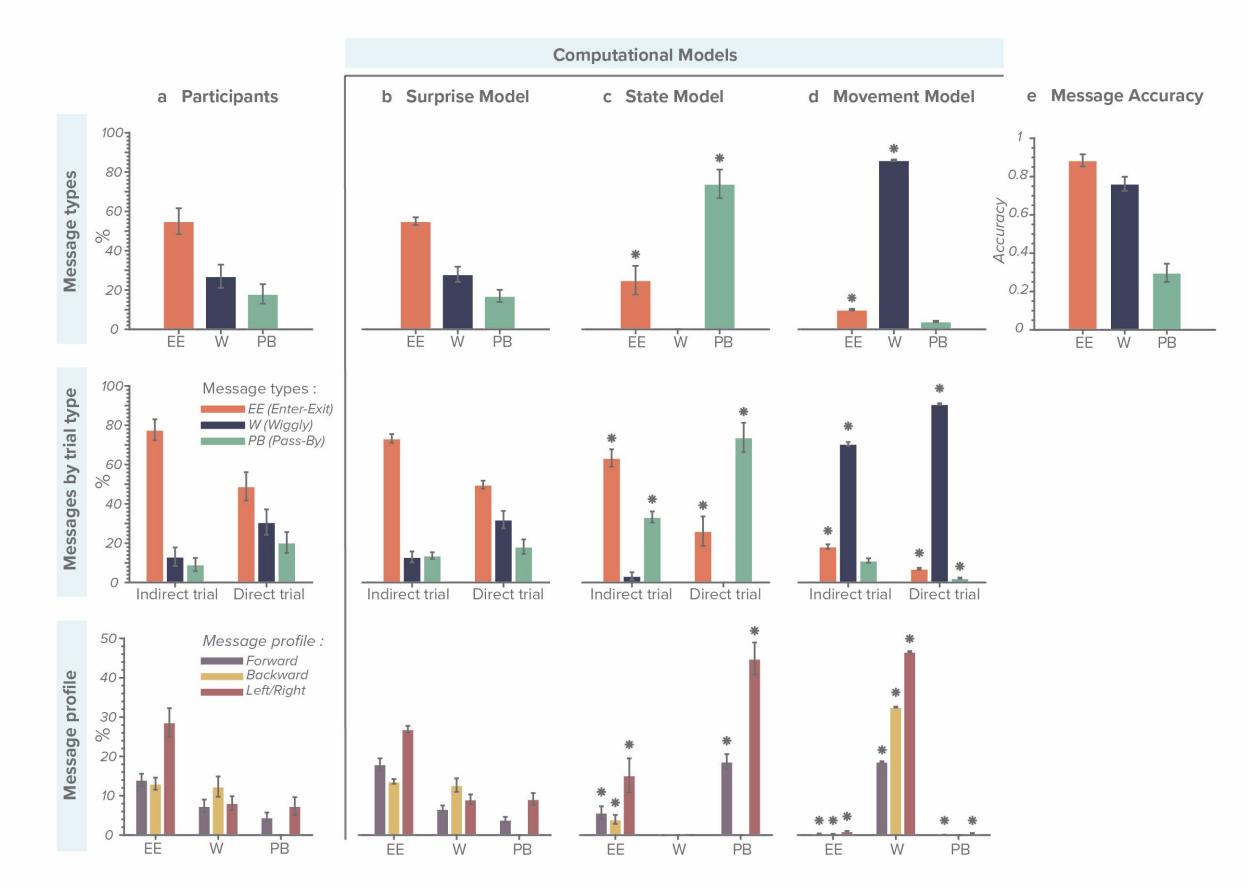


Figure 4 | Comparing Participant-Generated and Model-Simulated Messages for the Dataset 2
Note the close correspondence of the participant's behavioral data and the modeling results of Dataset 2 to Dataset 1 (Figure 3). **(a) Participants:** Participants' behavioral data categorized by Message types, Messages elicited in different types of trials, and Message profile. **(b) Surprise Model** simulated data after model fitting exhibits similarity to participants data **(**$BF_{01}>1$). **(c) State Model and (d) Movement Model** simulated data reveal differences from participant's data **(**asterisks indicate $BF_{01}<1$) **e) Message Accuracy:** Accuracy is calculated as the Receiver's success in finding his goal state. Enter-Exit messages have the highest accuracy rate, followed by Wiggly and Pass-By messages. All error bars represent means +/- SEM.

Posterior Predictive Check

Here we compare message generation by the Surprise model and its degenerate versions (i.e., State, and Movement models) with the messages generated by the participants. We employed a Bayesian approach[20] to assess the support for two hypotheses: the null hypothesis (H_0), which suggests no difference between the messages generated by each model and the empirical data, and the alternative hypothesis (H_1), which posits substantial differences between the model and the empirical data. We quantified the strength of this evidence for or against these hypotheses using a Bayes Factor (BF_{01}) for each behavioral index. Deviations of the model-generated from the behavioral data ($BF_{01} < 1$) are marked with asterisks above the model-generated indices in Figure 3 and 4 (for details, see Table S1 and S2).

Evidence for the Surprise Model. When comparing the means of human-created message types with those of the Surprise Model , the resulting BF_{01} (*D1: BF_{01}=10.20 (Enter-Exit), BF_{01}=6.48 (Wiggly), BF_{01}=10.36 (Pass-By); D2: BF_{01}=11.12 (Enter-Exit), BF_{01}=10.95 (Wiggly), BF_{01}=10.95 (Pass-By)*) strongly support the null hypothesis that human-created message types are similar to the ones generated by the Surprise model. We conducted a similar analysis for message profiles, revealing moderate to strong evidence supporting H_0. This is indicated by BF_{01} values for various profile elements, ranging from 1.31 to 10.42 (details in Table S1c and S2c for D1 and D2, respectively). These findings suggest that messages generated by the Surprise model are largely similar to human message profiles, though there are notable exceptions regarding Wiggly and Pass-by messages in different trial contexts of dataset D1. For more detailed information, please see Table S2b.

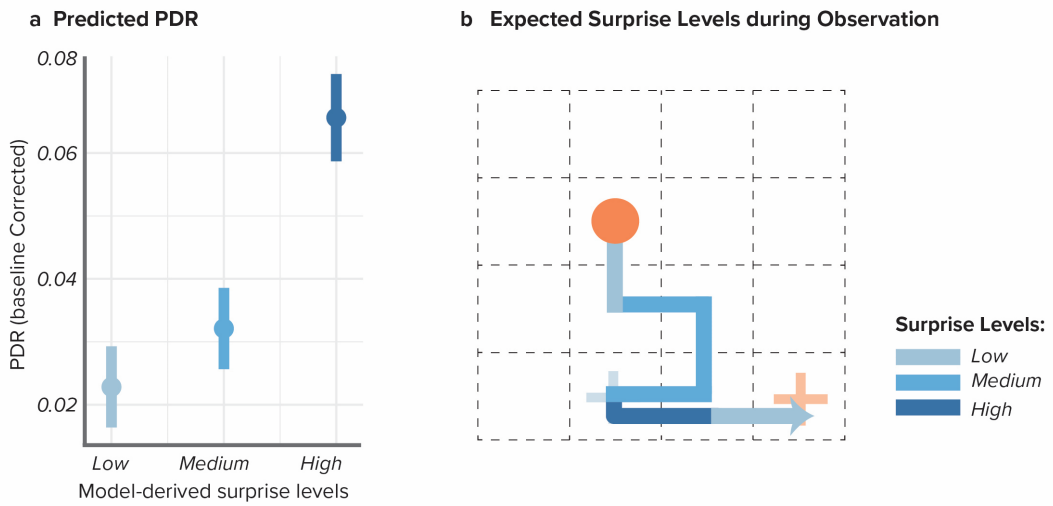
Evidence against State and Movement Models. In contrast to the Surprise model, when comparing the means of human-created message types to the means of the State and Movement models, all the calculated BF_{01} favor the alternative hypothesis, indicating that messages generated by these models are different from the messages generated by our participants, suggesting that these models are not capturing human message generation accurately. In summary, these results provide strong evidence in favor of the Surprise model and its principle of surprised-based action selection, as it can replicate the behavior of human senders.

Comparison with an Existing Computational Account. Due to its absence of estimable parameters, the Belief-Based Model (BBM) was excluded from the comparative analysis and parameter estimation within the context of TCG, as it's impossible to compute a subject-specific model likelihood necessary for model comparison, as previously discussed. Nonetheless, we simulated messages with BBM and performed a Bayes factor analysis to compare it against participant-created messages. For an in-depth examination of BBM's structure and the outcomes of its data simulation, please refer to the supplementary material.

Model-Based Analysis of PDR Data

Previous studies have demonstrated a close link between pupillary dilation responses (PDR) and violations of expectations (i.e., a surprise)[13]. We therefore hypothesized that the step-by-step surprise values calculated by our model would correlate significantly with the Receiver's PDR while observing the Sender's messages. To test this, we devised a model-informed analysis, in which we used the step-by-step surprise values from the Surprise model as a predictor for the participant's step-by-step PDR and used a mixed-effects model as implemented in the lmer package in R to test this correlation[22]. In the model, we incorporated a fixed effect for task condition—a categorical variable with three predictors (low, medium, and high surprise, split at the 33rd and

66th percentile of the range of surprise values), along with random intercepts for each participant. The mean PDR change of message steps served as the predicted variable (see Figure 5 for details). Our results revealed that the baseline-corrected PDR (w.r.t. to a pre-stimulus baseline of -200ms to onset) exhibited a significant effect for surprise (*F (3,41)=33.89), p<.001, eta=0.71*). Post-hoc tests revealed significant differences in PDR among the different levels of surprise. The high surprise level demonstrated a significantly larger PDR compared to the Medium surprise level with a mean increase of 0.03 mm *(SE=0.00845, df=38.50, t=3.41, p=0.0046 (Bonf)*) and compared to the Low surprise level *(mean increase 0.04 mm, SE=0.00845, df=38.50, t=4.45,* and *p=0.0002 (Bonf)*). However, the contrast between the Medium and Low surprise levels was not statistically significant, showing a mean difference of 0.01 mm (*SE=0.00627, df=31.33, t=1.40,* and *p=0.52 (Bonf)*). These findings indicate that the Receiver can detect the surprise conveyed by the Sender, with a greater arousal-related PDR occurring at the High surprise level.



c Pairwise Contrast between High, Medium and Low Suprise Levels

Contrast	Est.	S.E.	t val.	d.f.	p
High vs. Medium	0.03	8.45E-03	3.41	38.50	0.00
High vs. Low	0.04	8.45E-03	4.45	38.50	0.00
Medium vs. Low	0.01	6.27E-03	1.40	31.33	0.52

Figure 5 | Model-Based Analysis of PDR Data
(a) Predicted PDR: High surprise levels exhibit a significantly greater PDR when compared to Medium and Low surprise levels. **(b) Surprise Levels during Observation:** As the message steps are observed by the Receiver, varying levels of surprise associated with different steps are expected. Certain steps are more surprising than others. **(c) Pairwise Contrast between High, Medium and Low Surprise Levels.** All error bars represent means +/- standard error (SE).

Model-Based analysis of EEG Data

We hypothesized that if Receivers identify their goal location based on the deviation from their expectations, then we should observe a neuronal encoding of surprise in brain regions associated with error detection and conflict resolution[23,24].

We used the step-by-step surprise values from the Surprise model again as a predictor for EEG power in a model-informed EEG analysis. Specifically, we followed a trial-by-trial model-based approach[25–28] in which for each Receiver, we performed a regression at each electrode and time point within 0–1200 ms of each message step. Subsequently, we analyzed the regression weights by testing them against 0 across subjects for all electrodes and time-points.

First, we identified clusters of electrodes and time points that showed significant sensitivity to the predictor surprise irrespective of message type, after correcting for the multiple comparison (The cluster formation threshold was set at P=0.005 for t-tests)[29]. Specifically, we observed two distinct time-space clusters correlating significantly with surprise across electrodes and time. The first cluster displayed a positive effect at the frontal electrodes between 400 and 800 milliseconds of each message step. This cluster is visually depicted in Figure 6a, with yellow markers ($p < 0.05$ cluster corrected). The second cluster showed a positive effect at the frontal-central electrodes between 800 and 1200 milliseconds. This cluster is highlighted in Figure 6a with pink markers, also denoting significance ($p < 0.05$ cluster corrected). This analysis suggests that information-theoretic surprise is encoded in the neural activity levels in the brain regions identified by the significant clusters.

However, we also observed that different message types (Enter-Exit, Wiggly, and Pass-By) are associated with different levels of surprise. Is this difference also reflected in the neural activity in these areas? To answer this question, we repeated the above analysis for each message type separately.

Initially, we computed the average of the EEG data for each message type, followed by applying the same regression analysis at every time point and electrode as previously described. Subsequently, we examined the average regression weights from the clusters identified as significant in the earlier analysis (frontal and central clusters). The bar plots in Figure 6b next to the regression weight time course show the average surprise levels associated with each message type as calculated by the Surprise Model. We observed that message types characterized by high surprise, like Enter-Exit and Wiggly, exhibited a stronger correlation with EEG signals from fronto-central electrodes compared to less surprising messages such as Pass-by.

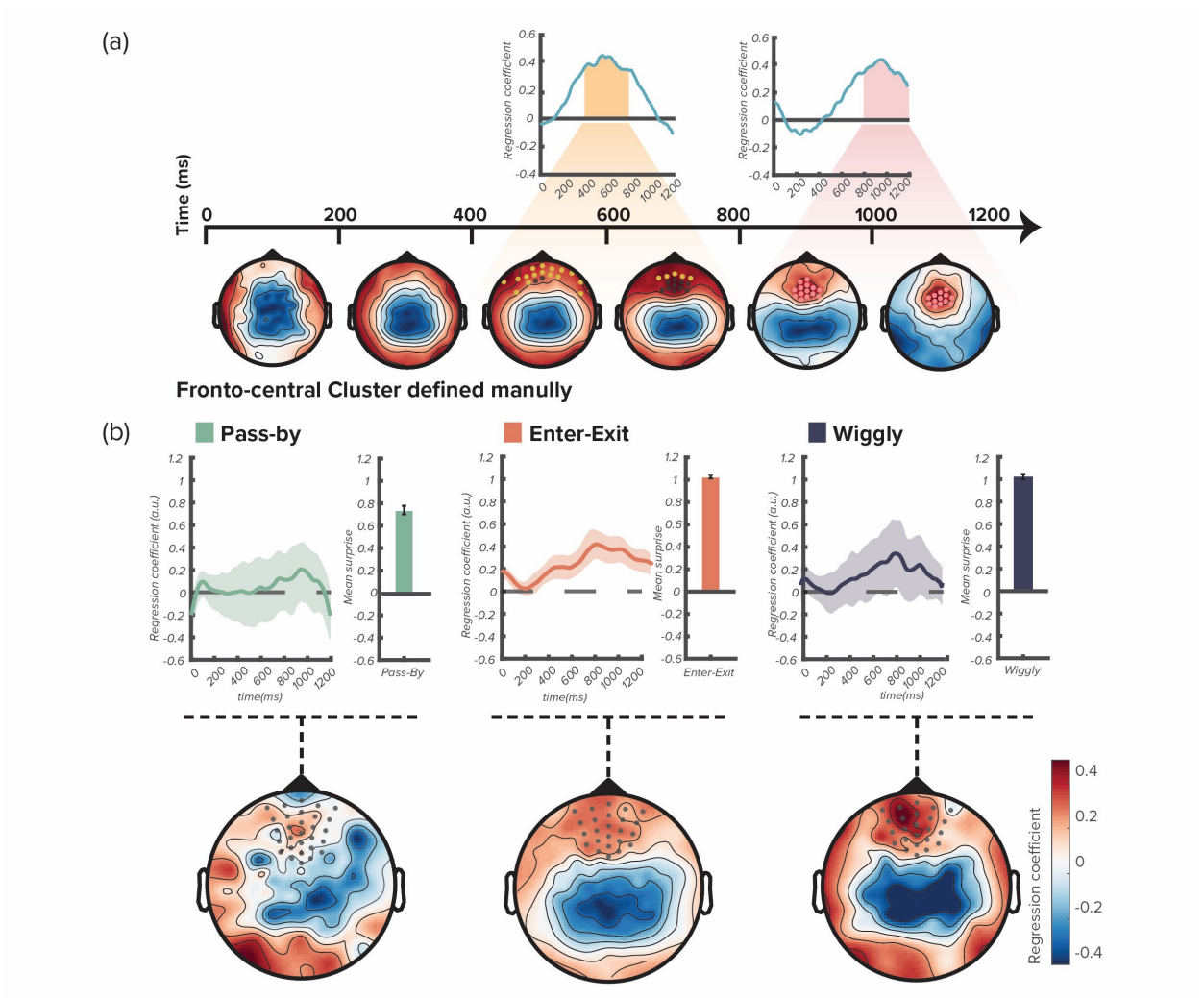


Figure 6 | Model-Based Analysis of EEG Data

(a) Two Distinct Electrode Clusters Sensitive to Surprise: The first cluster showed positive responses in frontal electrodes (represented by yellow dots on the topoplot at 400-800 ms), while the second occurred in frontal-central electrodes (illustrated by pink dots on the topoplot at 800-1200 ms). The time axis is the epoch of one message step during the observation phase of each message (see Figure 1a) which always lasts for 1200ms. **(b) Average Regression Weights for Different Message Types:** Weights were averaged over all-time points for each message type and then across Central-frontal electrodes (gray dots). Bar plots show the average surprise per message type as calculated by the model. Message types such as Enter-Exit and Wiggly, with higher surprise, correlate more strongly with EEG signals from Central-frontal electrodes than less surprising types like Pass-by. Error bars indicate means +/- SEM.

Discussion

In this study, we explored the role of surprise as communicative signals in the absence of a shared language among participants by employing the TCG as an experimental platform. Our innovative Surprise model capitalizes on the information-theoretic surprise as the signal most relevant for selecting the next possible action. A surprise can only unfold based on prior expectations, which the model calculates by integrating dynamic movement priors, rooted in principles of movement kinetics, and static state priors encoding the general goal orientation. A third component — the reward — provides a "push" toward shorter, more efficient messages that also pass through the Receiver's goal. Model simulations (Fig. 2 e-g) as well as model comparison with degenerate versions of the model (i.e., State model and Movement model, Table 1) revealed that each component is necessary to generate messages similar to those created by human participants in two independent data sets. The interplay of these model components suggests that message planning in the TCG is a complex cognitive operation, wherein human players intricately balance different subgoals (represented by the model components) to communicate efficiently.

The novel insight of this study is that when humans are confronted with social situations in which they cannot rely on shared cultural, linguistic, or even gestural cues, they are capable of reverting to an even more fundamental level of physical laws that govern our environment and are universally understood. These form the basis of a new communicative system that is developed and learned through repeated social interactions. In addition, humans then also employ an equally fundamental mechanism – expectancy violations – to communicate important information and guide receivers' attention. This mechanism is seen in careful analysis of the behavioral data and the construction of a computational model that is built around these principles.

We observed supporting evidence for this emergent mechanism of "action selection by surprise" in the analysis of the physiological and neural data. Traditionally, variations in pupil diameter have been indicators of arousal[30–32] and cognitive effort[33]. More recently, these changes have been reinterpreted to provide insights into factors such as surprise, salience, decision biases, and other dynamics that influence information processing[13,34,35]. A significant portion of these physiological responses can be attributed to the activation of norepinephrine (NE)-containing neurons in the locus coeruleus (LC) of the brainstem[36,37]. In examining Receivers' physiological reactions while processing messages, we found a direct correlation between PDRs and model-derived surprise signals. This indicates that Senders, when intentionally communicating with surprise, have a significant impact on the physiological states of Receivers.

An analysis of the neural encoding of surprise in the EEG data unveiled two prominent time-space clusters over frontal and fronto-central electrodes. Positioned above the anterior cingulate cortex (ACC), our model-informed analysis of the electrophysiological activity in the fronto-central cluster suggests that the neural processing within this region is involved in the detection of surprise.

The ACC is often considered as a hub for integrating diverse information and guiding decisions to align with our goals and reward representations[16,17]. Extensive research highlights the ACC's crucial role in linking actions to outcomes, both positive and adverse, and in evaluating the worthiness of specific actions[38,39]. Moreover, recent studies indicate that the ACC is not only essential in evaluating one's own reward representations but also in discerning the intentions of others and adjusting these perceptions in light of new evidence[40,41]. Our findings align with this perspective. The fronto-central cluster, presumably reflecting ACC activity, can be associated with the Receiver's ability to detect surprise as intentionally designed by the Sender. In other words, as Receivers process unexpected messages, such as Wiggly or Enter-exit messages, they might engage the ACC's capabilities to detect the deviation from expected outcomes[23,24].

The second cluster is located above the anterior prefrontal cortex, a brain region often associated with high-level cognitive control. Cognitive control refers to our ability to guide actions in alignment with an internally held goal, a function closely tied to the prefrontal cortex (PFC)[42–44]. In particular, hierarchical cognitive control[45] pertains to the intricate task of directing actions while maintaining more abstract goals. Moreover, the anterior parts of the PFC are thought to play a crucial role in this sophisticated form of cognitive control, potentially being engaged in future-oriented planning[46]. The frontal cluster observed in our study is reminiscent of this form of high-level action programming. Specifically, when the Sender crafts a message aiming for maximum surprise, they engage in advanced action planning. Once both Sender and Receiver settle on a strategy ensuring success, the Receiver might then develop a refined representation of the Sender's intended plan.

Traditionally, surprise and prediction errors arise from a deviation between expectations and real outcomes, serving as a teaching signal, aiming to refine expectations for the future. They play a major role in various theoretical accounts of learning, decision-making and perception, including reinforcement learning[1], predictive coding and Bayesian inference[47], and active inference[48,49]. As learning progresses, these prediction errors gradually decrease to zero as expectations align more closely with reality. Interestingly, all of these frameworks suggest that action selection should choose the policy predicting the environment most accurately[47] or maximizing the long term future reward[1]. This implies that actions with the smallest expected prediction error are selected. In contrast, our study revealed a different aspect of prediction errors. During communication, Senders intentionally employ surprise to disrupt a prior-consistent perception of the environment and thus convey critical information. This results in policies that are diametrically opposed to those from the above-mentioned computational frameworks[1,47,49], i.e., under the Surprise model the action with the largest expected surprise (i.e., the smallest prior) is selected. Yet, in communicative situations without a shared language like the TCG, this action selection produces messages that closely mimic human message design. Thus, it appears that our modeling efforts for the TCG has uncovered a hitherto undescribed policy "action selection by surprise" that contrasts with the "well-established "action selection by value" policy known from reward-based decision-making and that humans can easily shift between these policies based on task demands.

The potential implications of these findings extend across multiple domains, such as human-machine interfaces, artificial intelligence, and language acquisition, emphasizing the importance of harnessing expectancy violations (surprise/prediction errors) to improve communication. Nonetheless, while communication through surprise plays a crucial role in the TCG, it is not the exclusive method of communication in novel environments or of communication in general. Alternative experimental paradigms that leverage distinct cognitive processes may demand additional interpersonal skills, not just the efficient transfer of knowledge without using language. For instance, a therapeutic setting often calls for an understanding of the client's social and emotional situation, which is achieved through empathy and perspective-taking, two essential aspects of Theory of Mind.

Our study establishes a foundation for further exploration and application of surprise across diverse communicative contexts and paradigms, prompting inquiries into whether the "action selection by surprise" policy retains its efficacy across different communicative scenarios, beyond the TCG. Moreover, integrating this approach with AI technologies could be beneficial. Utilizing insights from the Surprise Model, we can investigate how artificial intelligence might effectively use intentionally created surprise signals within a novel communication context, thus enhancing interaction between humans and machines. For example, when generative AIs, such as large language models trained on a vast corpus of text found on the internet, generate a response, they determine the probability of numerous potential subsequent words or phrases based on the context at hand. Typically, they choose the words with the highest probability to ensure clarity. Yet, one could innovate by designing a 'surprise selection algorithm' that occasionally picks less expected alternatives. The application of such algorithms spans various domains, particularly enhancing novelty, creativity and preventing overfitting to typical responses. It holds the potential to enhance AI's comedic or storytelling abilities by introducing surprising plot twists, creating more engaging scenarios. Moreover, it can be applied to generative art and design to break from conventional patterns, leading to more unique creations.

In summary, the Surprise Model demonstrates the effectiveness of incorporating surprise in communication without a shared language, a phenomenon explored through the TCG. Not only did the Surprise Model exhibit a notable alignment with human sender behaviors, but it also validates its impact through both physiological and neural markers, affirming the crucial role of surprise in novel communication.

Methods

Participants

Dataset 1 involved 58 individuals (*Mean age: M=25, SD=4.4*), consisting of 29 pairs of Senders and Receivers, while dataset 2 included 72 participants (*Mean age: M=26, SD=4.1*), comprising

31 pairs of Senders and Receivers. Study participants were recruited through university job boards. The Exclusion criterion was a prior history of neurological or psychiatric diseases. All the participants were naive to the task and have been instructed with standardized written instructions. The ethics committee of the General Medical Council of the City of Hamburg granted approval for this study (PV7114), and all participants provided written informed consent prior to the study and were financially reimbursed for their participation. Each participant received a base payment of 12 euros for every hour of participation in the experiment. In addition to the base payment, participants were eligible for a bonus (Mean bonus: 4 euros) based on their performance.

Experimental Procedure

The TCG game consisted of 120 trials with different goal configurations. The participants were placed in the same room in front of separate screens to play the game after they had been assigned roles and given detailed instructions. They were instructed not to speak with each other during the experiment. In every trial, players had to solve different goal configuration problems. A goal configuration consists of three elements: 1) the starting location in one of 4 middle states of the 4×4 game board (same for both players), 2) the Sender's goal location, and 3) the Receiver's goal location.

At the beginning of every trial, the Sender was presented with a complete goal configuration, while the Receiver could only observe the starting location and the Sender's goal location (Fig. 1a). After seeing the goal configuration and planning the message, the Sender pressed a button to indicate that she was ready for movement. She then had 5 seconds to move her token on the grid. Once the Sender had reached her goal state, the entire message was shown to both players with the same step duration of 1.2 seconds. After the message replay had finished, the Receiver's token appeared at the starting location, and he started moving after pushing the start button. Finally, both players received feedback regarding their accuracy as a team.

Trial Types. To differentiate between different goal configurations, we focused on the location of the receiver's goal relative to the direct path as the key trial characteristic. This characteristic was used to define two distinct types of goal configuration, namely Direct and Indirect, as depicted in Figure 1c. The goal configuration, where the Receiver's goal location can be effectively conveyed by the shortest path from the starting location to the Sender's goal location ("Direct path") is referred to as the Direct configuration. Conversely, the goal configuration in which the Receiver's goal location can only be communicated by diverting from the Direct path is referred to as the Indirect configuration. In the experiment, Dataset 1 comprised 66% of trials with a Direct Goal Configuration and 34% with an Indirect Goal Configuration. In contrast, Dataset 2 exhibited a 76% Direct Goal Configuration to 24% Indirect Goal Configuration trial ratio.

Model-Free Behavioral Indices

We defined different indices to characterize the Sender's behavior in a more fine-grained way, providing insights into the cognitive processes behind creating new messages. These indices cover the message's general category (Message type) and its specific characteristics (Message profile).

Message Types. We've identified three sender-generated message types: Enter-Exit, Wiggly, and Pass-By (as shown in Figure 1b). The Enter-Exit message involves the sender moving towards the goal and then reversing direction, essentially entering, and then exiting the area near the goal. The Wiggly message, while similar to Enter-Exit in its backward movement, is distinguished by its multiple entries and exits, creating a series of back-and-forth movements between two locations. In contrast, the Pass-By message is characterized by a straightforward movement from the starting point to the goal, without any backward or repeated movements.

Message Profile. The message profile is a detailed characteristic of each message and includes information on the number of movements in each direction, specifically, Forward, Backward, Left and Right. An example profile of the Enter-Exit message is shown in Figure 1d, which consists of 4 forward, 2 left and right and 1 backward movements.

Analysis of Pupillary Responses

Data Acquisition and Preprocessing. Pupillary dilation data of the Receiver were collected using the screen-mounted eye-tracker iView Red-m (Sensorimotor Instrument (SMI), Teltow, Germany) with 120 Hz sampling frequency, while the Receiver was observing the Sender's message (Fig. 1a), displayed on the screen with the resolution of 1920 × 1200 pixels. The eye-tracker was positioned 60 cm away from the participant's face. A calibration was done before each game with the iViewX software. To avoid pupil dilation caused by the sunlight, the windows in the testing room were covered with blackout curtains with a light blocking effect. We used the same brightness and screen setting for all the experiments.

In order to identify and remove the noise and invalid data segments from the data, we followed the Guidelines recommended by Kret et al.[50] (https://github.com/ElioS-S/pupil-size) to prepare and filter raw data.

After identifying and interpolating invalid samples, we selected the relevant 1.2 s epochs during the replay phase of the trial for further analysis. Pupillary dilation response (PDR) data were downsampled by a factor of 10, thus rendering each step as a timeseries of 120 data points.

Mixed-Model Analysis for Pupillary Data. We investigated the effect of surprise values generated by the Surprise model on the Receiver's PDR, using a mixed-model analysis implemented in the lme4 package in R version 3.6.1, with the parameters estimated through maximum likelihood estimation[22].

In the model, we included a fixed effect for surprise level (low, medium, and high, split at the 33rd and 66th percentile) and a random intercept for each participant:

$$PDR \sim surprise\ levels + (1 \mid subject) \quad (1)$$

To test the significance of the fixed effects, we used the likelihood ratio test (LRT) implemented in the lme4 package in R[22].

EEG Data Analysis

Data Preprocessing and Cleaning. EEG signals were continuously recorded from both players with a 128-channel Active 2 system from BioSemi (Amsterdam, Netherlands) (0.1-100 Hz band pass; 2048 sampling rate). In the current study, we only analyzed the Receiver's EEG data during the replay section of a trial (Fig. 1a). We divided the continuous EEG time series into segments from 0 to 1200 ms based on the onset of each message step. After visually inspecting the data, we identified and interpolated an average of 6 bad channels out of 128 total channels. Additionally, we rejected an average of 40 bad epochs out of 550 total epochs. Blinks were removed by using Independent Component Analysis (average number of removed components: 5). The electrodes were re-referenced to the average common reference across the channels. These preprocessing steps were carried out using the EEGLAB software version 2022.1.

Model-Informed Regression Analysis. We used a model-informed approach to analyze the EEG data[25–28]. The approach entails a linear regression analysis of the EEG data of all steps of a message onto the Surprise model-derived surprise values for each message step at all-time points in the epoch and across all electrodes.

For this regression analysis, data were bandpass-filtered from 0.5 to 40 Hz, downsampled to 256Hz. Baseline correction was conducted by subtracting the mean activity between 0 and +100 after movement onset because there was no gap between different movement steps (and hence no time for the traditional pre-stimulus baseline) and because inspection of the ERPs did not show any significant deflection in the first 100 ms of each message step. We performed the above-mentioned regression analysis for each participant at each electrode and time point within 1200 ms epoch following movement onset. The predictors were the model-derived surprise values (described below) for each message step and a trial-specific intercept. This procedure creates a "time series" of regression coefficients, which were averaged across trials and then tested against

0 across participants to determine the time point at which the model-derived surprise significantly predicted the epoched EEG data.

The multiple comparison problem was addressed by using a cluster-based permutation test, as implemented in the fieldtrip software package[51]. The cluster formation threshold was set at a significance level of P=0.005 for the t-test.

Computational Models

Surprise Model

The core concept behind the Surprise model for the TCG is to communicate the Receiver's goal location by defying the expectations enforced by probabilistic priors about the next likely step and the general direction of the Sender's goal location. Violating these expectations communicates to the Receiver that he has to pay attention to this part of the message. The model quantifies these expectation violations by calculating the information-theoretic surprise for each available action and then selecting the action with the highest surprise. Thus, a critical part of the model is the definition of the prior probabilities. The model incorporates two types of priors – movement priors and state priors.

Movement Priors. The Movement model assigns probabilities to all available actions, based on the concept of movement kinetics, that an object in motion without an external force is likely to continue moving in a straight line. Deviations from this trajectory are less likely and, therefore, more surprising. This notion implies a dynamic realignment of the movement priors after each step. For instance, if the model takes a turn, the action probabilities will shift towards the new straight direction. Generally, the movement probabilities follow the ranking order of

$$p(f) > p(l) = p(r) > p(b) \tag{2}$$

where $p(f)$ is the probability of going forward, $p(l)$ is the probability of going left, $p(r)$ is the probability of going right, and $p(b)$ is the probability of going backwards. These four probabilities are parameterized using two estimable parameters, namely, λ and γ. They are expressed in the model as follows:

$$\lambda = p(b) \tag{3}$$

$$\gamma = \frac{p(f)}{p(l)+p(r)} \tag{4}$$

λ is restricted to be small (e.g., $0.001 - 0.3$), whereas γ is restricted to a range between $2 - 10$. For instance, if $\gamma = 2$, then $p(f)$ is twice as likely as the combined probability of a turn $(p(l) + p(r))$ and hence, $p(f) = 0.8$ and $p(l) = p(r) = 0.2$.

More generally, the individual probabilities for all four movements can be calculated from these two parameters in the following way:

$$p(b) = \lambda \tag{5}$$

$$p(f) = \frac{\gamma}{\gamma+1}(1-\lambda) \tag{6}$$

$$p(l) = p(r) = 0.5\,(1-\frac{\gamma}{\gamma+1})(1-\lambda) \tag{7}$$

Note that the movement priors remain constant for all steps in all trials, but the orientation (forward, left, right, backward) changes with every step the player takes.

State Priors. State priors play a crucial role in providing a general sense of the Sender's goal or destination. Specifically, the prior state probability is a function that has a value of 1 at the Sender's goal and then decreases according to an inverse power rule as the distance from the goal increases. The function is defined as

$$p(s) = \frac{1}{\alpha^d} \tag{8}$$

Where $p(s)$ is the probability of a state being at distance d from the Sender's goal location. α is the positive estimable slope parameter, determining the rate at which probabilities decrease with increasing distance from the Sender's goal. It remains constant for all trials, but the location of the peak of the state prior changes with every goal configuration.

Combining Movement-State Priors and Surprise. To render the final probabilities for each action *(p(a))*, movement and state priors are multiplicatively combined and normalized as follows.

$$p(a|s) = \frac{p(m)p(s'|s,m)}{\sum_{j=1}^{m=4} \quad p(m)p(s'|s,m)} \tag{9}$$

Where probability *p(m)* is one of the movement probabilities out of four possible movements (*f, l, r, b*), while $p(s' \mid s, m)$ is the state prior probability of state s' after movement m. Note that this is not a softmax filtering, but rather a normalization to insure valid probability values.

The information-theoretic (Shannon) surprise of each action is then calculated as:

$$h(a|s) = -log_2\big(p(a|s)\big) \tag{10}$$

Where $h(a|s)$ is a measure of information-theoretic surprise and shows, how much action a in state s deviates from what the Receiver expected the Sender to do based on the combined movement and state priors. In contrast to other models of learning and decision-making (like Reinforcement Learning) the Surprise model uses the Shannon surprise for the calculation of the final action probabilities.

Rewards. If the actions were only selected based on the Shannon surprise, then the final message trajectory would meander around the game board without efficiently reaching the Sender's goal (see simulations in Figure 2e-g). The rewards (conceptualized as the cost per step in the TCG) provide the Sender with an urge to reach his goal quickly and retain the maximum possible point from the initial endowment. Thus, in the model we defined the rewards $r(s'|a,s)$ at the next state s' that is reached by choosing action a in state s, as the number of points that would remain from the initial 10 points after taking the shortest route thereafter *through the Receiver's goal* to the Sender's goal. These rewards are multiplicatively combined with the Shannon surprise calculated above into an expected value:

$$ev(a|s) = h(a|s) \times r(s'|a,s) \tag{11}$$

Two Phases of Message Planning. Designing a message in the TCG has two components: (1) Creating a trajectory to the Receiver's goal that informs the Receiver of the location. The Surprise model realizes this by using the deviations from prior expectations (Shannon surprise), and (2) afterward moving to the Sender's goal on the shortest route to retain the maximum possible amount from the initial endowment. Thus, the second planning phase no longer uses the Shannon surprise of the combined movement-state prior for action selection, but simply moves along the direction of the state prior toward the Sender's goal. The expected value for this second planning phase is then:

$$ev(a|s) = p(s'|a,s) \times r(s'|a,s) \tag{12}$$

Planning a Message in Steps. In the context of a decision-making problem like a TCG, an agent can use value iteration[52] to plan its next moves up to a certain point in the future (a planning horizon). The planning horizon determines the number of steps the agent plans. For example, if the planning horizon is 2, the agent will plan its actions from the current step until 2 steps into the future. The purpose of this planning horizon is to determine the optimal course of action that maximizes the agent's sum of expected future reward.

In the Surprise model, this is done by generating a tree of possible actions and calculating the expected values of each possible action at each step of the planning horizon, while considering the potential future rewards that might come after. The expected value of getting to the next state in the policy tree from the current state s_i using action a_i is the following:

$$ev_{a_i|s_i} = h(a_i|s_i) \times \epsilon^i r(s'_i|s_i, a_i) \tag{12}$$

where $h(a_i|s_i)$ is the surprise associated with the action probability of getting to the next state in the policy tree from the current state s_i using action a_i and ϵ^i is the discount factor at the planning step i, and $r(s'_i|s_i, a_i)$ is the reward available in the next state s' that is reached from the current state s via action a in the planning step i. Once the model calculates the expected values for every state-action pair on the tree, it then sums up the expected values along each trajectory in the policy tree to render the final expected values for all immediately available actions:

$$EV(a|s) = \sum \; ev_{a_i|s_i} \tag{13}$$

These expected values are then filtered through a softmax function to render the final action probabilities:

$$Pr(a \mid s) = \frac{e^{\tau EV(a|s)}}{\sum_a e^{\tau E(a|s)}} \tag{14}$$

The softmax function is used to calculate the probability of taking a particular action a in a given state s, based on the expected future rewards associated with each action with an estimable temperature parameter τ that models the stochasticity of the choice.

Parameter Estimation. We used the Computational and Behavioral Modeling (CBM) toolbox (version 3, MATLAB R2021)[18], which uses a hierarchical variational Bayes approach to estimate the parameters of the Surprise model for each participant. The model is defined by five key parameters, represented as λ, γ, a, ϵ and τ, which are detailed in equations 3, 4, 8, 12, and 14 respectively. Our approach involved utilizing the Laplace approximation method, implemented in the toolbox, to effectively fit the model to our behavioral data.

Statistical Analysis of Posterior Predictive Checks

One of the objectives of our study was to identify the computational account that most closely replicates the behavior of the human sender in a posterior predictive check. We simulated new synthetic data using the fitted model parameters and calculated the model-free behavioral indices that were defined for the behavioral data (message frequency and message profile). We employed a Bayesian approach[20] to compare the null hypothesis (H_0: no differences between model-generated data and human experimental data) with the alternative hypothesis (H_1: substantial differences between model-generated and human experimental data) for all three model variations. A Bayesian approach allows for a quantification of evidence for both hypotheses.

To evaluate the evidence for H_0 versus H_1 for all models, we computed several Bayes factors (BF_{01}) by comparing the means of the behavioral indices discussed earlier with the means of the corresponding indices generated by the computational models.
A BF_{01} of less than 1 favors the alternative hypothesis (H_1) over the null hypothesis (H_0), indicating that the data provides more evidence supporting the noticeable discrepancy between the model and human sender behavior. Conversely, a BF_{01} greater than 1 favors the null hypothesis (H_0) over the alternative hypothesis (H_1), indicating that the model can closely replicate the behavior of human senders.

Acknowledgements and Author Contributions

We thank Julia Wandschura for help with data collection and Fatemeh Hadaeghi for her valuable comments. This work is supported by a DFG / NSFC grant to J.G. and X.F. (CRC TRR 169 “Cross-modal Learning”, NSFC grant 62061136001) and a DFG grant to J.G. (CRC 1528 “Cognition of Interaction”). Author contributions: research design - J.G. and T.B., experimental task and data

collection - T.B., model development - J.G., model implementation - T.B., data analysis - T.B., paper writing (first draft) - T.B., paper writing (revisions) - T.B., J.G., T.S., K.Z., X.F.

Data Availability Statement

All behavioral and physiological data necessary to run the analysis is available on the authors' GitHub at https://github.com/TatiaBu/CommunicationSurprise_DataCode. EEG data is available upon request from the authors.

Code Availability Statement

The code to run the analysis and reproduce the figures is archived and publicly available on GitHub at https://github.com/TatiaBu/CommunicationSurprise_DataCode. The repository includes a README file with detailed instructions on how to set up the environment, install necessary dependencies, and execute the code.

References

1. Sutton, R. S. & Barto, A. G. *Reinforcement Learning, second edition: An Introduction*. (MIT Press, 2018).
2. Shaw, J. Philosophy of humor. *Philos. Compass* **5**, 112–126 (2010).
3. Reisenzein, R., Horstmann, G. & Schützwohl, A. The Cognitive-Evolutionary Model of Surprise: A Review of the Evidence. *Top. Cogn. Sci.* **11**, 50–74 (2019).
4. Flamson, T. & Clark Barrett, H. THE ENCRYPTION THEORY OF HUMOR: A KNOWLEDGE-BASED MECHANISM OF HONEST SIGNALING. *Journal of Evolutionary Psychology* doi:10.1556/JEP.6.2008.4.2.
5. Goel, V. & Dolan, R. J. The functional anatomy of humor: segregating cognitive and affective components. *Nat. Neurosci.* **4**, 237–238 (2001).
6. Berger, A. A. Anatomy of the joke. *J. Commun.* **26**, 113–115 (1976).
7. Stephens, R. C. & Ward, J. J. Motion In A Circle. in *Applied Mechanics* (eds. Stephens, R. C. & Ward, J. J.) 27–39 (Macmillan Education UK, 1972). doi:10.1007/978-1-349-00870-4_3.
8. Ruiter, J. P. D. *et al.* Exploring the cognitive infrastructure of communication. *Interaction StudiesInteraction Studies. Social Behaviour and Communication in Biological and Artificial Systems*

11, 51–77 (2010).

9. Haggard, P. *et al.* On the origin of intentions. *Sensorimotor Foundations of Higher Cognition* 601–618 (2012) doi:10.1093/acprof:oso/9780199231447.003.0026.
10. Blokpoel, M. *et al.* Recipient design in human communication: simple heuristics or perspective taking? *Front. Hum. Neurosci.* **6**, 253 (2012).
11. Willems, R. M., Benn, Y., Hagoort, P., Toni, I. & Varley, R. Communicating without a functioning language system: Implications for the role of language in mentalizing. *Neuropsychologia* **49**, 3130–3135 (2011).
12. Shannon, C. E. A mathematical theory of communication. *SIGMOBILE Mob. Comput. Commun. Rev.* **5**, 3–55 (2001).
13. Preuschoff, K., 't Hart, B. M. & Einhäuser, W. Pupil Dilation Signals Surprise: Evidence for Noradrenaline's Role in Decision Making. *Front. Neurosci.* **5**, 115 (2011).
14. de Gee, J. W., Correa, C. M. C., Weaver, M., Donner, T. H. & van Gaal, S. Pupil Dilation and the Slow Wave ERP Reflect Surprise about Choice Outcome Resulting from Intrinsic Variability in Decision Confidence. *Cereb. Cortex* **31**, 3565–3578 (2021).
15. Sara, S. J. & Bouret, S. Orienting and reorienting: the locus coeruleus mediates cognition through arousal. *Neuron* **76**, 130–141 (2012).
16. Kolling, N., Behrens, T., Wittmann, M. K. & Rushworth, M. Multiple signals in anterior cingulate cortex. *Curr. Opin. Neurobiol.* **37**, 36–43 (2016).
17. Kennerley, S. W., Walton, M. E., Behrens, T. E. J., Buckley, M. J. & Rushworth, M. F. S. Optimal decision making and the anterior cingulate cortex. *Nat. Neurosci.* **9**, 940–947 (2006).
18. Piray, P., Dezfouli, A., Heskes, T., Frank, M. J. & Daw, N. D. Hierarchical Bayesian inference for concurrent model fitting and comparison for group studies. *PLoS Comput. Biol.* **15**, e1007043 (2019).
19. Stephan, K. E., Penny, W. D., Daunizeau, J., Moran, R. J. & Friston, K. J. Bayesian model selection for group studies. *Neuroimage* **46**, 1004–1017 (2009).
20. Hoijtink, H., Mulder, J., van Lissa, C. & Gu, X. A tutorial on testing hypotheses using the Bayes factor. *Psychol. Methods* **24**, 539–556 (2019).
21. de Weerd, H., Verbrugge, R. & Verheij, B. Higher-order theory of mind in the Tacit Communication

Game. *Biologically Inspired Cognitive Architectures* **11**, 10–21 (2015).

22. Luke, S. G. Evaluating significance in linear mixed-effects models in R. *Behav. Res. Methods* **49**, 1494–1502 (2017).
23. Botvinick, M. M., Braver, T. S. & Barch, D. M. Conflict Monitoring and Cognitive Control. doi:10.1037//0033-295X.I08.3.624.
24. Holroyd, C. B. & Coles, M. G. H. The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychol. Rev.* **109**, 679–709 (2002).
25. Collins, A. G. E. & Frank, M. J. Neural signature of hierarchically structured expectations predicts clustering and transfer of rule sets in reinforcement learning. *Cognition* **152**, 160–169 (2016).
26. Fischer, A. G., Nigbur, R., Klein, T. A., Danielmeier, C. & Ullsperger, M. Cortical beta power reflects decision dynamics and uncovers multiple facets of post-error adaptation. *Nat. Commun.* **9**, 5038 (2018).
27. Cavanagh, J. F. Cortical delta activity reflects reward prediction error and related behavioral adjustments, but at different times. *Neuroimage* **110**, 205–216 (2015).
28. Larsen, T. & O'Doherty, J. P. Uncovering the spatio-temporal dynamics of value-based decision-making in the human brain: a combined fMRI-EEG study. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **369**, (2014).
29. Maris, E. & Oostenveld, R. Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* **164**, 177–190 (2007).
30. Braun, A., Urai, A. E. & Donner, T. H. Pupil-linked arousal is driven by decision uncertainty and alters serial choice bias. *Nat. Commun.* **8**, 1–11 (2017).
31. van Kempen, J. *et al.* Behavioural and neural signatures of perceptual decision-making are modulated by pupil-linked arousal. *Elife* **8**, (2019).
32. Bradley, M. M., Miccoli, L., Escrig, M. A. & Lang, P. J. The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology* **45**, 602–607 (2008).
33. Foroughi, C. K., Sibley, C. & Coyne, J. T. Pupil size as a measure of within-task learning. *Psychophysiology* **54**, 1436–1443 (2017).
34. Hoeks, B. & Levelt, W. J. M. Pupillary dilation as a measure of attention: a quantitative system

analysis. *Behav. Res. Methods Instrum. Comput.* **25**, 16–26 (1993).

35. Yebra, M. *et al.* Action boosts episodic memory encoding in humans via engagement of a noradrenergic system. *bioRxiv* **30**, 71–42 (2017).
36. Joshi, S., Li, Y., Kalwani, R. M. & Gold, J. I. Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex. *Neuron* **89**, 221–234 (2016).
37. Larsen, R. S. & Waters, J. Neuromodulatory Correlates of Pupil Dilation. *Front. Neural Circuits* **12**, 21 (2018).
38. Rushworth, M. F. S., Kolling, N., Sallet, J. & Mars, R. B. Valuation and decision-making in frontal cortex: one or many serial or parallel systems? *Curr. Opin. Neurobiol.* **22**, 946–955 (2012).
39. Rushworth, M. F. S., Walton, M. E., Kennerley, S. W. & Bannerman, D. M. Action sets and decisions in the medial frontal cortex. *Trends Cogn. Sci.* **8**, 410–417 (2004).
40. Apps, M. A. J., Rushworth, M. F. S. & Chang, S. W. C. The Anterior Cingulate Gyrus and Social Cognition: Tracking the Motivation of Others. *Neuron* **90**, 692–707 (2016).
41. Zhang, L. & Gläscher, J. A brain network supporting social influences in human decision-making. *Sci Adv* **6**, eabb4159 (2020).
42. Miller, E. K. & Cohen, J. D. An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* **24**, 167–202 (2001).
43. Duncan, J., Emslie, H., Williams, P., Johnson, R. & Freer, C. Intelligence and the frontal lobe: the organization of goal-directed behavior. *Cogn. Psychol.* **30**, 257–303 (1996).
44. Badre, D. Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. *Trends Cogn. Sci.* **12**, 193–200 (2008).
45. Badre, D. & Desrochers, T. M. Hierarchical cognitive control and the frontal lobes. *Handb. Clin. Neurol.* **163**, 165–177 (2019).
46. Rouault, M. & Koechlin, E. Prefrontal function and cognitive control: from action to language. *Current Opinion in Behavioral Sciences* **21**, 106–111 (2018).
47. Aitchison, L. & Lengyel, M. With or without you: predictive coding and Bayesian inference in the brain. *Curr. Opin. Neurobiol.* **46**, 219–227 (2017).
48. Parr, T. & Friston, K. J. Working memory, attention, and salience in active inference. *Sci. Rep.* **7**,

14678 (2017).

49. Friston, K. *et al.* Active inference and learning. *Neurosci. Biobehav. Rev.* **68**, 862–879 (2016).

50. Kret, M. E. & Sjak-Shie, E. E. Preprocessing pupil size data: Guidelines and code. *Behav. Res. Methods* **51**, 1336–1342 (2019).

51. Oostenveld, R., Fries, P., Maris, E. & Schoffelen, J.-M. FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* **2011**, 156869 (2011).

52. Leslie Pack Kaelbling , Michael L. Littman, Anthony R. Cassandra '. Planning and acting in partially observable stochastic domains. *Artif. Intell.*