Universiteit Leiden
The Netherlands

**Modeling human sequential behavior with deep neural networks in emergent communication**
Kouwenhoven, T.; Verhoef, T.; Raaijmakers, S.A.; Kleijn, R.E. de; Goldwater, M.; Anggoro, F.K.; ... ; Ong, D.C.

**Citation**
Kouwenhoven, T., Verhoef, T., Raaijmakers, S. A., & Kleijn, R. E. de. (2023). Modeling human sequential behavior with deep neural networks in emergent communication. *Proceedings Of The Annual Meeting Of The Cognitive Science Society*, 1747-1754. Retrieved from https://hdl.handle.net/1887/3638763

| | |
|---|---|
| Version: | Publisher's Version |
| License: | Creative Commons CC BY 4.0 license |
| Downloaded from: | https://hdl.handle.net/1887/3638763 |

**Note:** To cite this publication please use the final published version (if applicable).

# Modeling Human Sequential Behavior with Deep Neural Networks in Emergent Communication

**Tom Kouwenhoven (t.kouwenhoven@liacs.leidenuniv.nl)**
Leiden Institute of Advanced Computer Science, Creative Intelligence Lab, Leiden, the Netherlands

**Tessa Verhoef (t.verhoef@liacs.leidenuniv.nl)**
Leiden Institute of Advanced Computer Science, Creative Intelligence Lab , Leiden, the Netherlands

**Stephan Raaijmakers (s.a.raaijmakers@hum.leidenuniv.nl)**
Leiden University Centre for Linguistics, Leiden University, Leiden, the Netherlands

**Roy de Kleijn (kleijnrde@fsw.leidenuniv.nl)**
Cognitive Psychology Unit, Institute of Psychology, Leiden, the Netherlands

## Abstract

In this paper, we study human sequential behavior by integrating cognitive, evolutionary, and computational approaches. Our work centers around the emergence of shared vocabularies in the Embodied Communication Game (ECG). Here, participant pairs solve a shared task without access to conventional means of communication, enforcing the emergence of a new communication system. This problem is solved typically by negotiating a shared set of sequential signals that acquire meaning through interactions. Individual differences in *Personal Need for Structure* (PNS) have been found to influence how this process develops. We trained deep neural networks to mimic the emergence of new communicative systems and used hyperparameter optimization to approximate latent human cognitive variables to explain human behavior. We demonstrate that models based on bidirectional LSTM networks are better at capturing human behavior than unidirectional LSTM networks. This suggests that human sequence processing in the ECG is influenced by expected future states. The approximated variables cannot explain the differences in PNS, but we do provide evidence suggesting that random and uncertainty-directed exploration strategies are combined to develop optimal behavior.

**Keywords:** Computational modeling; Human-machine interaction; Language emergence; Deep neural networks; Sequential behavior modeling

## Introduction

For communication—between humans or between humans and machines—to be successful, the coordinated actions of all interlocutors must adhere to the *grounding criterion*, according to which interlocutors have to agree on the meaning of the current communicative purposes (H. Clark & Brennan, 1991). The fulfillment of this criterion relies extensively on the availability of a (partially) shared vocabulary between interlocutors of a conversation (Pickering & Garrod, 2004). Yet, the exact dynamics of how agents settle on an effective grounded shared vocabulary is still unclear (Tylén, Fusaroli, Bundgaard, & Østergaard, 2013; Mordatch & Abbeel, 2018). Recent work in Computational Linguistics has started modeling emergent communication setups using multi-agent simulations to understand this process better (e.g., Lazaridou, Hermann, Tuyls, & Clark, 2018; Chaabouni, Kharitonov, Dupoux, & Baroni, 2019; Chaabouni, Kharitonov, Bouchacourt, Dupoux, & Baroni, 2020; Chaabouni et al., 2021). But the findings from these simulations often do not match the outcomes of similar experiments with humans Lazaridou, Potapenko, and Tieleman (2020). As such, recent literature proposes to instill human language patterns in machines by including human feedback in the learning loop instead of only learning from large quantities of data (ter Hoeve, Kharitonov, Hupkes, & Dupoux, 2021; Iocchi et al., 2022; Kouwenhoven, Verhoef, de Kleijn, & Raaijmakers, 2022).

The interdisciplinary research presented here attempts to instill such human communicative behavior in machines, using an experimental set-up that allows studying the initial emergence of simple signals where no communication existed before. As such, we explore the grounding problem from an evolutionary perspective, where humans must collaboratively create a novel shared communication system to successfully play the ECG (Scott-Phillips, Kirby, & Ritchie, 2009). This two-player game addresses fundamental steps in the emergence of languages: how does a signal obtain its communicative intent, and how does this signal obtain its meaning? Most human participants can solve this non-trivial task by establishing an initial convention (i.e., settling on a default behavior) and collaboratively bootstrapping new signals onwards (Scott-Phillips et al., 2009; Kouwenhoven, de Kleijn, Raaijmakers, & Verhoef, 2022). These meaningful signals are subsequently used to play the ECG successfully, creating sequences of communicative behavior.

Once a communicative system exists, it must be processed by the brain for comprehension and production. However, it is not entirely clear how this happens for human languages. Traditional views see the human brain as a forward-looking prediction machine (e.g., A. Clark, 2013), but Onnis, Lim, Cheung, and Huettig (2022) found evidence for the importance of backward-looking processing for language comprehension in two self-paced reading and eye-tracking tasks. They showed that context, in the form of preceding words, can be informative for integrating current words, and conclude that both forward *and* backward-looking appear to be important characteristics of language processing. A similar debate exists on processing everyday sequential actions

(De Kleijn, Kachergis, & Hommel, 2014). Early accounts suggested that sequential actions are triggered by the perception of motor execution of the previous action (Washburn, 1916). Yet, there is also evidence that anticipated future states also influence subsequent actions and that planning mechanisms play a role in sequential tasks (e.g., Lashley et al., 1951; Cohen & Rosenbaum, 2004; De Kleijn, Kachergis, & Hommel, 2018), but how exactly this happens is hitherto not well understood.

Context, in the form of preceding behavior or incoming signals, and intended future states also play a role in the ECG. Incoming and produced signals (i.e., context) are informative of future behavior, and anticipated future states can be seen as desired behaviors by the other (i.e., ending on a specific color). The behaviors in the ECG are moreover sequential but less complex than everyday actions and can therefore be studied in a controlled manner. As such, investigating this through computational modeling may reveal how sequential processing possibly played a role in shaping human language, what types of agent architectures are required to facilitate natural communication between humans and machines, and contribute to the debate on sequential action processing in humans.

From a computational view, we use behavior cloning to 1) investigate whether deep learning models can learn the expressed human behaviors during the development of signal–meaning mappings in the ECG; 2) approximate latent human cognitive variables by optimizing model parameters that influence learning and exploration (for an overview of similar work, see Schulz & Gershman, 2019); 3) identify the applicability of networks with different processing directions to model human behavior. We then relate the model parameters with a cognitive measure of Personal Need for Structure (Thompson, Naccarato, & Parker, 1989) and compare the ability to learn human behavior for models with different processing directions. Doing so has the potential to facilitate more natural human-machine interactions through the development of (language) models that possess shared biases, resulting in a more human-like quality. Vice versa, deviations between human and computational biases provide a better understanding of why outcomes might not be as desired. Lastly, a better understanding of the influence of such biases on the emergence of language could steer learning mechanisms in computational simulations of emergent communication and close the gap between evolved human and computational behavior. Ultimately to the benefit of natural interactions with conversational AI.

## Background

The origin of language is extensively studied, but the exact dynamics of language emergence remain unknown. One question concerns the origins of the initial signal–meaning mappings in case no prior communication system exists. If neither form nor meaning is known, a possible way to establish this concerns the cooperative process of agreement on the relations between communicative signals and meanings. This process has been studied extensively through laboratory experiments in which participants invent and negotiate novel signals to solve a cooperative task (Steels, 2006; Scott-Phillips & Kirby, 2010; Tylén et al., 2013). These studies show that humans can establish shared conventions and develop communication systems through social coordination. It is moreover suggested that in addition to language use, human learning and transmission of a language affect the emergence of patterns (Kirby, Tamariz, Cornish, & Smith, 2015; Smith, 2022). A paramount explanation for the highly structured nature of human language is that it emerges due to a human bias for compressible systems through a preference for simplicity (Kemp & Regier, 2012; Kirby et al., 2015; Kirby & Tamariz, 2022).

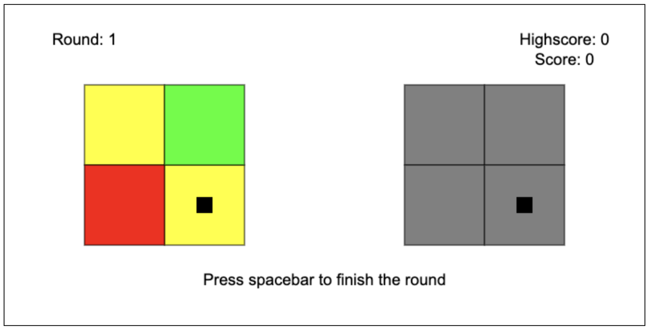
The Personal Need for Structure Scale is a measure of a bias for simplicity (Thompson et al., 1989). This questionnaire quantifies individuals' need for structure (PNS), desire for cognitive simplicity (F1), and the aim of restructuring an environment into a more manageable and simplified form (F2) (Neuberg & Newsom, 1993). Differences in the desire for structure influence how individuals understand and interact with the world (Neuberg & Newsom, 1993) and also affect problem-solving capabilities (Eva, Silvia, & Dáša, 2014; Svecova & Pavlovicova, 2016). Furthermore, PNS affects the task progression of participants playing the ECG, participant pairs who respond differently to a lack of structure are more successful (Kouwenhoven, de Kleijn, et al., 2022).

### Embodied Communication Game

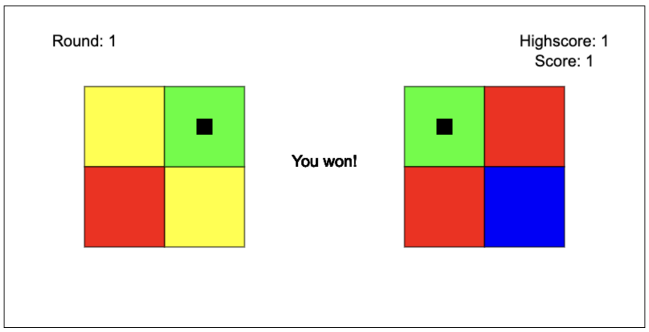
The ECG is a cooperative two-player game that consists of two 2×2 grid worlds. Each quadrant has one of four colors. Both players move between the quadrants, using the arrow keys, and share the goal of ending on identically colored quadrants. When they manage to do so, they score a point. For both grids, the colors and starting positions are determined randomly for each round, with the proviso of one overlapping color, such that it is always possible to score a point. Players see their movements and the movements made by their partner, but only see the colors of their quadrants (Figure 1a). The colors of both worlds are revealed to both players (Figure 1b) when both finish moving. Their goal is to score as many consecutive points as possible, meaning that pairs must find a way to communicate reliably and coordinate behaviors (see Scott-Phillips et al., 2009 for an in-depth explanation).

## Methods

The relationships between computational parameters and cognitive measures are investigated through training deep neural networks on human behaviors in the ECG. Specifically, algorithmic parameters are used as a *proxy* of human preferences, we do not claim the existence of exactly these representations in the human brain, but merely use them as another measuring device of human behavior. Similar work is done by De Kleijn et al. (2018), who used reinforcement




(a) View while playing




(b) View when done

Figure 1: Only the colors from the participants' own grid are visible while players are moving (1a). When both players are done, the colors of all quadrants are revealed to both players (1b). Images were taken from (Kouwenhoven, de Kleijn, et al., 2022).

learning (RL) models to fit human behavior in a serial reaction time (SRT) task and found that good human performance requires a high learning rate and a low discount factor. Suggesting that low-scoring individuals do not update their action-value function or the expected utility of their actions. Curricularized learning for RL agents in the SRT task showed that similar to infants' curiosity-based learning, exploration can promote robust later learning in virtual agents (de Kleijn, Sen, & Kachergis, 2022).

For textual data, Nikolaus and Fourtassi (2021) evaluated the ability of neural networks to acquire meanings of words and sentences through laboratory tasks that involve cross-situational learning used with children. They show that neural networks mirror learning patterns of acquiring semantic knowledge in early childhood and suggest that children might use partial representations of sentence structure to guide semantic interpretation. Additionally, it is shown that language models rely more on word frequency than children, but like children, learn words slower when these are part of longer utterances (Chang & Bergen, 2022). These models notably differed from children in the effects of word length, lexical class, and concreteness on learning, emphasizing the importance of social, cognitive, and sensorimotor experience in child language development.

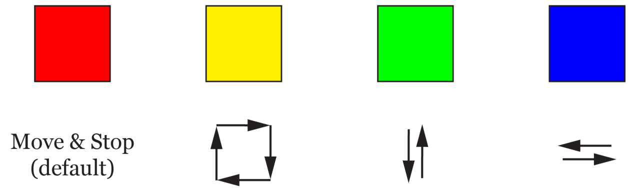



Figure 2: An example communication system.

## Data

The data used in this paper was collected for the study described in (Kouwenhoven, de Kleijn, et al., 2022). Here we collected three additional pairs ($N = 46$: 36 females, 10 males; $M_{\text{age}} = 22.2$, $SD_{\text{age}} = 3.53$). Participants received instructions after which they were separated and placed behind two connected computers. This set-up ensured that conventional communication was impossible and that the problem of emerging signal–meaning mappings had to be solved by the participants. The game was played for 40 minutes, for an average of 256 rounds, after which participants filled out the PNS questionnaire and described the communication systems they attempted to develop. Finally, they were debriefed and allowed to discuss their experience. This study was approved by the Psychology Research Ethics Committee of Leiden University.

Out of 23 pairs, only 14 managed to create (i.e., reported and demonstrated) a robust communicative system. A Bayesian *t*-test showed that these pairs achieved higher scores than pairs that did not establish a system ($\text{BF}_{10} = 26.73$). A typical system contains sequences of movements to indicate different colors (Figure 2). Once established, pairs negotiate which color is available to both by repeating the sequential moves associated with this color. We refer to Scott-Phillips et al. (2009) and Kouwenhoven, de Kleijn, et al. (2022) for a detailed description of the emergence of such communicative behavior.

A sequence of game states, produced by the movements of each participant, is stored for each round. A single state contains the players' position, the position of the other player, the color of the currently occupied quadrant, and the entire color layout of the players' grid. This representation reflects the information that a participant sees during the game. A target label—corresponding to arrow keys and the spacebar—is stored for each game state to create a sequence of actions.

## The model

We trained a deep neural network—implemented with Long Short Term Memory (LSTM, Hochreiter & Schmidhuber, 1997) cells—on the state–action sequences of each participant. The input data, therefore, differs for each model, but its architecture is generic and fixed (Figure 3). The objective of the model is to predict a participant's subsequent move given a particular sequence of states. For unidirectional processing, each state of a sequence is processed chronologically, beginning with the first and ending with the last state[1]. For bidirectional processing, the states are additionally processed in reverse order, thus incorporating (i.e., anticipating) future behavior to predict a subsequent move. The model output layer computes probabilities for subsequent moves using temperature ($\tau$) scaling. Here, high values of $\tau$ cause actions to be approximately equiprobable and therefore lead to exploratory behavior. Low values of $\tau$ cause greater differences between the probabilities, with high probabilities for actions with high expected rewards, and cause deterministic behavior. The learning rate ($lr$) of the model influences how quickly it updates its predictions, where a high learning rate means quick changes. The Adam optimization algorithm (Kingma & Ba, 2014) is used to minimize categorical cross–entropy loss.

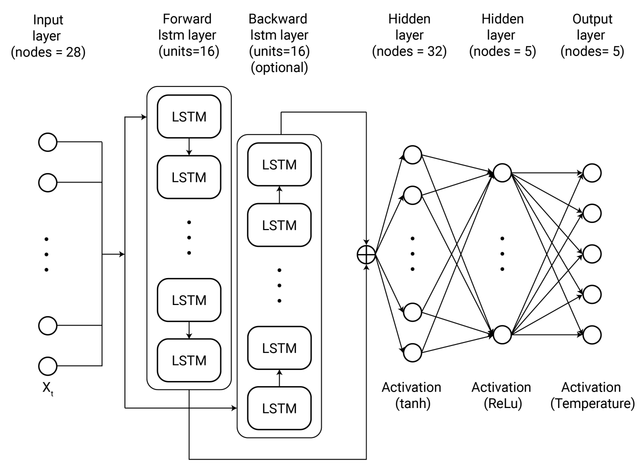



Figure 3: Neural network architecture. The output layer uses temperature scaling as an activation function.

### Measures

Game performance was measured by the number of consecutive successful rounds (*high score*). PNS and its subfactors were collected using a 12-statement questionnaire (see Neuberg & Newsom, 1993), here, high values for *PNS*, *F1*, and *F2* correspond to a high need for structure. To obtain participant-specific $\tau$ and $lr$, we performed hyperparameter optimization on the game data of each participant, resulting in 46 independently trained models. Put differently, grid search was used to optimize model performance using $lr \in \{0.0001, ..., 0.075\}$ and $\tau \in \{0.001, ..., 3.00\}$, with 10 equally spaced steps per parameter, resulting in 100 parameter settings per participant. Each model was trained independently for 5 epochs on each parameter combination. We take the learning rate as a proxy of the extent to which individuals weigh feedback when updating their estimates and use temperature as an approximation of how explorative their behavior was. The ability of the model to predict human sequential behaviors is reflected in model accuracy (*acc*). Lastly, the categorical cross–entropy loss (*cce*, i.e., negative log-likelihood) explains how likely the model and human would perform the same action in a particular game state. For each model, we used three-fold cross-validation to ensure that the model was not learning the data explicitly but captured the underlying structures of that participant. The cross-validation score (i.e., the average over all folds) described model performance. The parameter combination that resulted in the highest cross-validation score was used as a *proxy* for the latent human cognitive variables.

[1]The backward processing layer is not used for unidirectional networks.

## Modeling human sequential behavior

Behavior cloning was used to explain human behavior in the ECG on two accounts. First, by comparing PNS measures with the computational parameters. Neuberg and Newsom (1993) showed that differences in the need for simple structure influence how individuals understand and interact with the world. Arguably, inferred computational parameters such as learning rate and temperature have similar influences. Therefore we sought correspondence between these parameters and the PNS scores of each participant. We hypothesized that learning rate is related to the desire for cognitive simplicity (*F1*) and *high scores* since a desire for structure implies active searching for patterns, which seems crucial to learn signal–meaning mappings in the ECG. Learning these patterns faster (i.e., high $lr$) might result in faster emergence of communicative patterns. Individuals who feel uncomfortable in unstructured environments (i.e., high *F2*) show lower adaptability and flexibility in new environments, preferring to respond with familiar behavioral patterns to counter the uncomfortable feeling (Steinmetz, Loare, & Houssemand, 2011). Since lower values of $\tau$ correspond to less exploratory behavior and a high $lr$ corresponds with high adaptability, it was expected for $lr$ and $\tau$ to correlate negatively with *F2*.

Secondly, we manipulated the sequential processing cells of the models. As argued before, the next move of a signal and the intended finishing color influence immediate action selection and can therefore be formulated as an anticipated future state. As such, optimization as described in the previous section is done for the unidirectional (LSTM) and bidirectional LSTM (biLSTM) model. Whereas unidirectional cells process time steps of sequences in a chronological forward manner, bidirectional cells compute inputs forward *and* backward to make predictions (Schuster & Paliwal, 1997). Note that although the LSTM layer in our model is different for both types, the remaining architecture is identical.

### Results

Statistical analyses were done using R 4.0.5 (R Core Team, 2021) and the BayesFactor 0.9.12-4.3 package (Morey et al., 2018). First, we consider the overall performance of both network types. The mean accuracy (*acc*) over all independently trained models shows that both network types can learn to predict subsequent moves relatively well (Table 1). Comparison between the two network types with a Bayesian $t$-test on

Table 1: The average model performance over the cross-validation scores for each participant. *Uni* and *bi* correspond to the model types LSTM and biLSTM respectively.

| | **acc** | | **cce** | | **lr** | | $\tau$ | |
|---|---|---|---|---|---|---|---|---|
| Type | M | SD | M | SD | M | SD | M | SD |
| Uni | .831 | .112 | .355 | .241 | .019 | .019 | .356 | .745 |
| Bi | .972 | .055 | .084 | .153 | .039 | .020 | 2.28 | .716 |

*acc* and *cce* with network type as a predictor revealed a large performance difference ($\text{BF}_{10}acc = 6.63e+11, d = 1.66$ and $\text{BF}_{10}cce = 1.50e11, d = -1.59$). Indicating that bidirectional sequence processing can better capture the human behavior in the ECG than unidirectional sequence processing (Figure 4). This result is robust when controlled for the number of parameters between the two network architectures. Optimal learning rate and temperature were higher for biLSTM networks when compared to LSTM networks ($\text{BF}_{10}lr = 5.85e3, d = .790$ and $\text{BF}_{10}\tau = 3.46e+14, d = 2.00$). Since the learning rate was taken as a proxy for the extent to which individuals update their estimates, a higher learning rate implies flexible behavior. Therefore this result suggests that bidirectional processing requires more flexibility toward updating behavior policies. Additionally, it implies that explorative behavior might complement updating these policies. We can assume that a higher learning rate translates to better learning in humans since learning is required to play the ECG successfully and learning rates were significantly higher for pairs that managed to establish a communicative system compared to those that did not ($\text{M}_{successful} = .047, \text{M}_{unsuccessful} = .025, \text{BF}_{10} = 556, d = 1.39$).

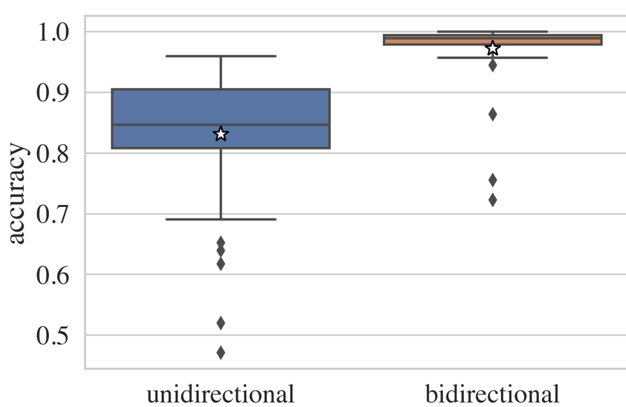
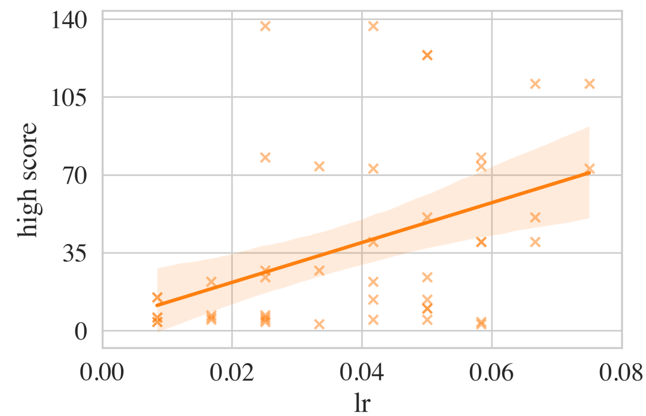
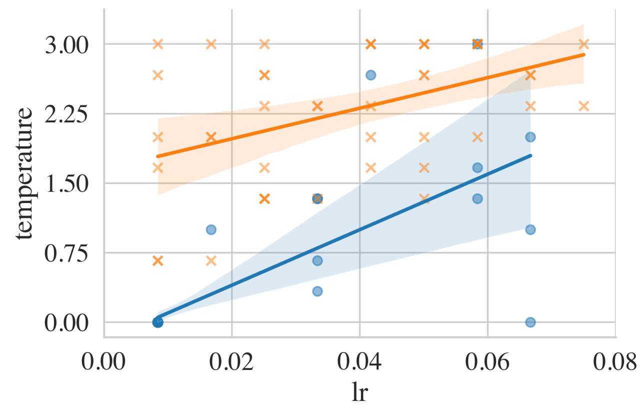



Figure 4: BiLSTM models show greater accuracy than LSTM models. Stars indicate mean accuracy.

We now consider the relationships between model parameters, cognitive measures, and high scores as described earlier. Successful participants (i.e., participants with a high score) performed complex and structured sequences in order to communicate. Nevertheless, we find that for LSTM networks, but not for biLSTM networks, *high score* negatively influences *acc* ($\text{BF}_{10} = 3.07, r = -.346, r^2 = .120$). This suggests that unidirectional processing is able to learn simpler human behavior relatively well but has difficulties capturing more elaborate behaviors. Importantly, this finding explains the difference observed in Figure 4.




(a) Relationship between learning rate and temperature.




(b) Relationship between learning rate and high score.

Figure 5: Relationships between learning rate, high score, and temperature. Each point corresponds to one participant. Note: darker marks are overlapping data points and the shaded area is the 95% confidence interval. Blue is used for *unidirectional* networks and orange is used for *bidirectional networks.*

Bayesian regression showed that for biLSTM networks, there is a positive linear relationship between learning rate and high score (Figure 5a $\text{BF}_{10} = 12.8, r^2 = .183$), confirming our hypothesis and suggesting that participants who adopt new behaviors faster are more successful in creating new signal–meaning mappings in the ECG. We moreover find that regardless of processing directionality, temperature, and learning rate are related (Figure 5b, $\text{BF}_{10}biLSTM = 28.1, r = .452, r^2 = .204$ and $\text{BF}_{10}LSTM = 1.40e7, r = .772, r^2 = .597$), suggesting that participants who explored more also adapted new behaviors faster. Surprisingly, we did not find a relation between exploration and *high score*. This was to be expected since explorative behavior may lead to new conventions in the ECG. Lastly, in the current data learning rate or temperature cannot explain *PNS, F1*, or *F2* for LSTM and biLSTM networks. Thereby rejecting the remaining hypotheses.

## Discussion

Here we modeled human sequential behavior in the Embodied Communication Game with deep neural networks and investigated possible relationships between human cognitive preferences and computational parameters. Specifically, we looked at relationships between participants' personal need for structure, learning rate, and temperature parameters. Although we showed that the current deep neural networks are able to learn the behavior associated with creating signal–meaning mappings, we did not find any correspondences between cognitive and computational variables. As such, PNS, used here as a human bias for structure (Kirby & Tamariz, 2022), cannot be captured with this setup. Further research should investigate how parameters of various network architectures may have correspondences with cognitive measures or look at different games that investigate emergent communication (e.g., Galantucci, 2005; Steels & Loetzsch, 2012; Mordatch & Abbeel, 2018). Being able to capture human biases, such as the human bias for compressible and simple systems (Kemp & Regier, 2012; Kirby et al., 2015), in computational systems is insightful for simulations of emergent communication as they are then closer to human experiments. Furthermore, playing these collaborative games between humans and machines might also result in shared grounded vocabularies that are adapted to the biases of humans *and* computers, ultimately resulting in better conversational AI (Kouwenhoven, Verhoef, et al., 2022).

Manipulation of the processing directionality of action sequences showed that participants' behavior was explained better by biLSTM models than by LSTM models. This hereby provides additional arguments for bidirectional processing of sequential actions in humans (Lashley et al., 1951; Cohen & Rosenbaum, 2004; Onnis et al., 2022). For communicative purposes in the ECG, integrating current actions is dependent on the preceding shared context (i.e., the negotiations of signals and intended final colors), and must be taken into account when deciding what moves to take next. The difficulties for LSTM networks to learn more complex behaviors performed by more successful participants also indicate that unidirectional processing is insufficient to capture more elaborate human behavior. Although more research is needed to support this, these findings suggest that the effect of a backward-looking mechanism found by Onnis et al. (2022) in a self-paced reading task might originate in the very early stage of forming signaling conventions. To verify this, simulations of emergent communication with deep learning agents should look at the effect of processing directionality of network architectures on the structure of emergent communicative protocols. Integrating bidirectional networks may close the current gap between human experiments and simulations.

We showed that for biLSTM networks learning rate positively influences high scores and was positively related to temperature (Figure 5b). This seems to support the recent view which suggests that humans combine random and uncertainty-directed exploration strategies to develop optimal behavior (Jepma et al., 2016; Schulz & Gershman, 2019). An explanation for this could be that explorative behavior in the ECG led to the emergence of new signals, which need to be learned quickly (i.e., require a high learning rate) to be useful. In other words, the correlation between learning rate and temperature likely reflects the fact that participants who are more explorative benefit from higher learning rates (i.e., there is no benefit to explorative behavior if you do not use the explored options to update expected values). However, in-depth analysis is required to strengthen this link further. For optimal behavior, learning rate and explorative behavior would be expected to decrease over time as strategies are learned and exploration becomes less necessary, instead exploiting the knowledge gathered thus far. However, literature on how learning rate and temperature parameters develop with age and experience has yielded conflicting results (Nussenbaum & Hartley, 2019). Games like the ECG could be extended over time to investigate the dynamic nature of the temperature and learning rate parameters.

Lastly, we acknowledge that the ECG is a highly simplified setup, thereby limiting the generalizability to real-world processing (Nastase, Goldstein, & Hasson, 2020). It also goes without saying that these models are mere approximations of the human brain and do not capture its breadth, but we simply use them as a proxy to mimic human processes. These findings must therefore be replicated in more ecological settings.

## Conclusion

In this work, we modeled sequential human behavior captured in the Embodied Communication Game with deep neural networks. Here participants have to establish a communication system from scratch to solve a collaborative task. We demonstrate that neural networks can learn the human behaviors associated with the creation of a new communication system. Manipulation of network types shows that bidirectional processing of sequential actions better explains human behavior than unidirectional processing, hereby providing additional arguments for the existence of a planning mechanism for sequential signal production in humans. No relationship between Personal Need for Structure and participant–specific computational parameters was found, but our results suggest that humans combine random and uncertainty-directed exploration strategies to develop optimal behavior in the ECG. Future research should attempt to extrapolate our results to communicative settings with complex linguistic signal exchange (e.g., between chatbots and humans). Additionally, experiments on the emergence of a more complex human–AI language will deepen the understanding of the relationship between natural and artificial biases for communication.

## References


Chaabouni, R., Kharitonov, E., Bouchacourt, D., Dupoux, E., & Baroni, M. (2020, July). Compositionality and generalization in emergent languages. In *Proceedings of the 58th annual meeting of the association for computational linguistics* (pp. 4427–4442). Online: Association for Computational Linguistics. Retrieved from `https://aclanthology.org/2020.acl-main.407` doi: 10.18653/v1/2020.acl-main.407

Chaabouni, R., Kharitonov, E., Dupoux, E., & Baroni, M. (2019). Anti-efficient encoding in emergent communication. *Advances in Neural Information Processing Systems*, *32*.

Chaabouni, R., Strub, F., Altché, F., Tarassov, E., Tallec, C., Davoodi, E., . . . Piot, B. (2021). Emergent communication at scale. In *International conference on learning representations.*

Chang, T. A., & Bergen, B. K. (2022, 01). Word Acquisition in Neural Language Models. *Transactions of the Association for Computational Linguistics*, *10*, 1-16. Retrieved from `https://doi.org/10.1162/tacl_a_00444` doi: 10.1162/tacl_a_00444

Clark, A. (2013). Whatever next? predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204. doi: 10.1017/S0140525X12000477

Clark, H., & Brennan, S. (1991, 01). Grounding in communication. In (Vol. 13, p. 127-149). doi: 10.1037/10096-006

Cohen, R. G., & Rosenbaum, D. A. (2004). Where grasps are made reveals how grasps are planned: generation and recall of motor plans. *Experimental Brain Research*, *157*(4), 486–495.

De Kleijn, R., Kachergis, G., & Hommel, B. (2018). Predictive movements and human reinforcement learning of sequential action. *Cognitive science*, *42*, 783–808.

de Kleijn, R., Sen, D., & Kachergis, G. (2022). A critical period for robust curriculum-based deep reinforcement learning of sequential action in a robot arm. *Topics in Cognitive Science*.

De Kleijn, R., Kachergis, G., & Hommel, B. (2014). Everyday robotic action: lessons from human action control. *Frontiers in Neurorobotics*, *8*, 13.

Eva, S., Silvia, H., & Dáša, M. (2014). Personal need for structure in relation to language variables. *Procedia-Social and Behavioral Sciences*, *159*, 665–670.

Galantucci, B. (2005). An experimental study of the emergence of human communication systems. *Cognitive science*, *29*(5), 737–767.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, *9*(8), 1735–1780.

Iocchi, L., et al. (2022). Emergent communication in human-machine games. In *Emergent communication workshop at iclr 2022.*

Jepma, M., Murphy, P. R., Nassar, M. R., Rangel-Gomez, M., Meeter, M., & Nieuwenhuis, S. (2016). Catecholaminergic regulation of learning rate in a dynamic environment. *PLoS computational biology*, *12*(10), e1005171.

Kemp, C., & Regier, T. (2012). Kinship categories across languages reflect general communicative principles. *Science*, *336*(6084), 1049–1054.

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Kirby, S., & Tamariz, M. (2022). Cumulative cultural evolution, population structure and the origin of combinatoriality in human language. *Philosophical Transactions of the Royal Society B*, *377*(1843), 20200319. doi: https://doi.org/10.1098/rstb.2020.0319

Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, *141*, 87–102.

Kouwenhoven, T., de Kleijn, R., Raaijmakers, S., & Verhoef, T. (2022). Need for structure and the emergence of communication. In *44th annual meeting of the cognitive science society (cogsci 2022).*

Kouwenhoven, T., Verhoef, T., de Kleijn, R., & Raaijmakers, S. (2022). Emerging grounded shared vocabularies between human and machine, inspired by human language evolution. *Frontiers in Artificial Intelligence*, *5*.

Lashley, K. S., et al. (1951). *The problem of serial order in behavior* (Vol. 21). Bobbs-Merrill Oxford.

Lazaridou, A., Hermann, K. M., Tuyls, K., & Clark, S. (2018). Emergence of linguistic communication from referential games with symbolic and pixel input. *arXiv preprint arXiv:1804.03984*.

Lazaridou, A., Potapenko, A., & Tieleman, O. (2020, July). Multi-agent communication meets natural language: Synergies between functional and structural language learning. In *Proceedings of the 58th annual meeting of the association for computational linguistics* (pp. 7663–7674). Online: Association for Computational Linguistics. Retrieved from `https://aclanthology.org/2020.acl-main.685` doi: 10.18653/v1/2020.acl-main.685

Mordatch, I., & Abbeel, P. (2018). Emergence of grounded compositional language in multi-agent populations. In *Proceedings of the aaai conference on artificial intelligence* (Vol. 32).

Morey, R., Rouder, J., Jamil, T., Urbanek, S., Forner, K., & Ly, A. (2018). Bayesfactor: Computation of bayes factors for common designs (r package version 0.9. 12-4.3)[computer software]. *Retrieved form https://CRAN. R-project. org/package= BayesFactor*.

Nastase, S. A., Goldstein, A., & Hasson, U. (2020). Keep it real: rethinking the primacy of experimental control in cognitive neuroscience. *NeuroImage*, *222*, 117254. Retrieved from `https://www.sciencedirect.com/science/article/pii/S1053811920307400` doi: https://doi.org/10.1016/j.neuroimage.2020.117254

Neuberg, S. L., & Newsom, J. T. (1993). Personal need for structure: Individual differences in the desire for simpler structure. *Journal of Personality and Social Psychology*, *65*(1), 113.

Nikolaus, M., & Fourtassi, A. (2021, June). Evaluating the acquisition of semantic knowledge from cross-situational learning in artificial neural networks. In *Proceedings of the workshop on cognitive modeling and computational linguistics* (pp. 200–210). Online: Association for Computational Linguistics. Retrieved from https://aclanthology.org/2021.cmcl-1.24 doi: 10.18653/v1/2021.cmcl-1.24

Nussenbaum, K., & Hartley, C. A. (2019). Reinforcement learning across development: What insights can we draw from a decade of research? *Developmental Cognitive Neuroscience*, *40*, 100733. Retrieved from https://www.sciencedirect.com/science/article/pii/S1878929319303202 doi: https://doi.org/10.1016/j.dcn.2019.100733

Onnis, L., Lim, A., Cheung, S., & Huettig, F. (2022). Is the mind inherently predicting? exploring forward and backward looking in language processing. *Cognitive Science*, *46*(10), e13201. doi: https://doi.org/10.1111/cogs.13201

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and brain sciences*, *27*(2), 169–190.

R Core Team. (2021). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from https://www.R-project.org/

Schulz, E., & Gershman, S. J. (2019). The algorithmic architecture of exploration in the human brain. *Current Opinion in Neurobiology*, *55*, 7-14. Retrieved from https://www.sciencedirect.com/science/article/pii/S0959438818300904 (Machine Learning, Big Data, and Neuroscience) doi: https://doi.org/10.1016/j.conb.2018.11.003

Schuster, M., & Paliwal, K. K. (1997). Bidirectional recurrent neural networks. *IEEE transactions on Signal Processing*, *45*(11), 2673–2681.

Scott-Phillips, T. C., & Kirby, S. (2010). Language evolution in the laboratory. *Trends in cognitive sciences*, *14*(9), 411–417.

Scott-Phillips, T. C., Kirby, S., & Ritchie, G. R. (2009). Signalling signalhood and the emergence of communication. *Cognition*, *113*(2), 226–233.

Smith, K. (2022). How language learning and language use create linguistic structure. *Current Directions in Psychological Science*, *31*(2), 177–186.

Steels, L. (2006). Experiments on the emergence of human communication. *Trends in Cognitive Sciences*, *10*(8), 347–349.

Steels, L., & Loetzsch, M. (2012). The grounded naming game. *Experiments in cultural language evolution*, *3*, 41–59.

Steinmetz, J.-P., Loare, E., & Houssemand, C. (2011). Rigidity of attitudes and behaviors: A study on the validity of the concept. *Individual Differences Research*, *9*(2).

Svecova, V., & Pavlovicova, G. (2016). Screening the personal need for the structure and solving word problems with fractions. *SpringerPlus*, *5*(1), 1–9.

ter Hoeve, M., Kharitonov, E., Hupkes, D., & Dupoux, E. (2021). Towards interactive language modeling. *arXiv preprint arXiv:2112.11911*.

Thompson, M. M., Naccarato, M. E., & Parker, K. E. (1989). Assessing cognitive need: The development of the personal need for structure and personal fear of invalidity scales. In *annual meeting of the canadian psychological association, halifax, nova scotia, canada.*

Tylén, K., Fusaroli, R., Bundgaard, P. F., & Østergaard, S. (2013). Making sense together: A dynamical account of linguistic meaning-making. *Semiotica*, *2013*(194), 39–62.

Washburn, M. F. (1916). *Movement and mental imagery: Outlines of a motor theory of the complexer mental processes*. Houghton Mifflin.