A replication of experiment one of the study: “Where the truth lies: how sampling implications drive deception without lying” by Ransom et al. (2019)

Lara Wenzler, Dario Hol Kieslich, Jana Schapka and Emily Seitz

University of Osnabrueck

Abstract

Communication is the simple act of transferring information from one person to another. But what happens when the receiver and sender in communication have differing motives? Do we interpret statements of our partners in communication differently when we take their intent into account and we feel like we can or cannot trust them? In this paper, we aimed to replicate experiment one of the study 'Where the truth lies: how sampling implications drive deception without lying' by Ransom et al. (2019) and found out that people do take the intent of the sender into account when interpreting the offered evidence.

Keywords: Bayesian modelling, deception, meta-inference, sampling assumptions

A replication of experiment one of the study: "Where the truth lies: how sampling implications drive deception without lying" by Ransom et al. (2019)

Everyday communication often requires us to fill in the blanks between what is said and what is meant. This leads to more efficient communication but it also leaves room for misinterpretations. The process behind us filling in these blanks is based on individual knowledge and experiences and it is called inference. In the case of communication, we do not only make use of basic inference but also of meta-inference. This means that as part of encoding a message we reason about how the other person has reasoned. This process is, just like inference, vulnerable to misinterpretations and deception. Deception is the act of the sender leading the receiver to a belief that the sender considers false. The two components lying and detecting lies make up the term 'communication of deception' as Zuckerman et al. (1981) said. Except for lying, people use tactics like *misleading* or being *uninformative* to create deception. The experimental work in this paper excludes outright lying, forcing the senders to either withhold information or to provide information that is factually correct but misleading. In order to get a deeper insight into the issue of deception and inference Ransom et al. (2019) have conducted two related experiments that examine how people reason in cases of deception. In experiment one, participants were in the role of the target of deception and in the second experiment the participants were the perpetrator of deception. The experimenters found that one group assumed the intent based on both the context and content of the message they received. This group was both more likely to mislead and to be misled. They found that another group refrained from such behavior and showed caution in all situations while adopting a more withstanding attitude.

We replicated the first experiment of the study by Ransom et al. (2019) and examined if and how meta-inference is affected when the intent of the sender is taken into account. To be more specific, we first tested whether participants, when presented with *misleading*

evidence, chose the *Lure* significantly more often when they believe to play in the same team as the sender than when they think they play against an opponent. Additionally, we tested the same hypothesis with the difference that the participants were presented with *uninformative* evidence. For the first hypothesis, our research results align with those of the original paper and highlight the interplay between meta-inference and inference when interpreting statements of either a trusted or distrusted partner in communication. We did not find significant evidence to support the second hypothesis.

Method

Participants

The participants were recruited via social media and email. The sample size was roughly guided by the sample size used in the original paper. They recruited 99 participants, we aimed for 80, due to a lack of time and possible language barriers of the mostly German-speaking participants. In the end, 89 participants took part in the online experiment. Of these participants, 24 were excluded because they showed a lack of understanding of the task. The remaining 65 participants were 18-51 years old (median 23).

Materials

Each of the six sets (see rows in Figure 1) consisted of three pieces of evidence (*uninformative*, *misleading* and *helpful*) and four maps (two *Decoy*, *Lure* and *Truth*). The quality of the evidence varied from trial to trial. The three pieces of evidence were designed in the following way:

The *helpful* evidence formed a pattern of locations that bore close similarity to the true map and categorically excluded the two *Decoy* and the *Lure* map. The *uninformative*

evidence was consistent with all of the four maps. The *misleading* evidence was designed to bear a strong similarity to one of the three false maps, namely the *Lure*. Additionally, the informativeness of the *misleading* evidence varied over the sets of stimuli in a way that it ruled out either none, one or two of the *Decoy* maps but never the *Truth* and never the *Lure*. This is indicated with percentages in Figure 1 on the left-hand side. The full set of stimuli can be seen in Figure 1.

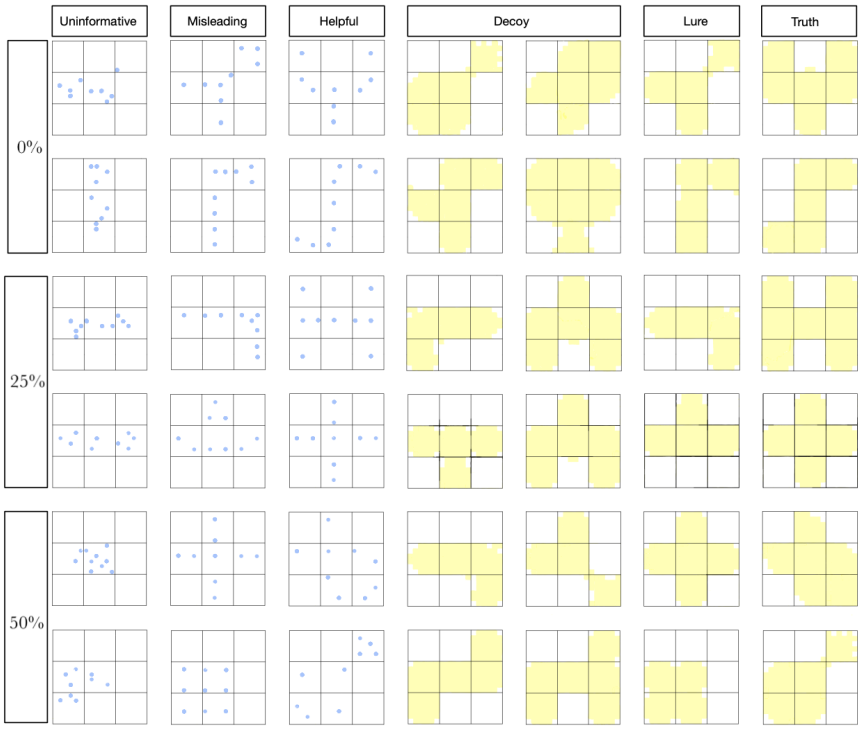

Figure 1: The experimental stimuli

Procedure

The experiment was presented to the participants using _magpie and implemented with JavaScript, CSS and HTML and hosted on Netlify via GitHub. It consisted of five parts. First, there was an introduction view with instructions for the practice phase. Participants were told that they are taking part in an experiment simulating an online game that is based on data provided by past players. They were told that they play the role of an 'explorer' who must decide which of four treasure maps (see bottom row in Figure 2) is the genuine one

based on evidence (see top row in Figure 2) provided by a past player, here called ‘the pirate’. In our terminology regarding communication, the explorer is the receiver and the pirate is the sender. The evidence consisted of points that corresponded to a subset of locations drawn from the genuine map, where each point represents the location of a hidden treasure. Furthermore, the participants were told that the sender can provide *misleading, helpful* or *uninformative* evidence by strategically selecting points. The basis of the within-subjects manipulation is people’s belief about the sender’s intent. There were two conditions. In the teammate condition, participants were told that the goal of the ‘fellow pirates’ is to help the receiver to find the genuine map and that they will provide evidence accordingly. In the opponent condition, the receivers were told that the goal of the ‘opposing pirates’ is to keep the genuine map concealed. Additionally, participants are being informed that ‘pirates’ were not allowed to provide false information, regardless of the condition. Thus, participants knew that they could rule out every map, where the shaded regions in the map and the locations indicated on the evidence map don’t overlap. In Figure 2 there is an example of how one round looked like.

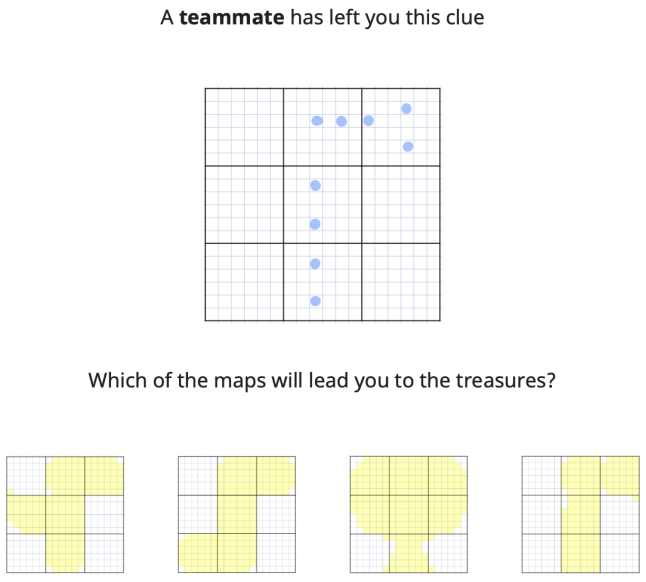

Figure 2: One round in the deception game

For the second part participants were asked to go through some practice trials. There were six in total, three for the teammate condition and three for the opponent condition.

After that, the main phase was introduced with some more instructions leading the participants to go through 30 trials in total for the main phase. Everything in the main phase is randomized, the on-screen order of the maps displayed in each trial, the trial order within each block and the block order itself. On each trial, the participants are asked to consider the four maps, the evidence provided and whether the 'pirate' is a teammate or an opponent when deciding which of the four maps they believe to be the genuine one.

Finally, the participants are asked to participate in the optional post-experiment questionnaire. The post-experiment questionnaire asked the participants for their age, their gender, their level of education and their native languages. Additionally, there was an option to leave a comment.

Data Preparation

For the final analysis, only data from the main trials was used. Furthermore, to prevent participants from clicking through the experiment and accidental double clicking we excluded participants who had a mean reaction time of under three seconds and individual trials which had a trial reaction time of under two seconds. This way, we can ensure that none of the trials we included in the analysis were from someone partaking in the experiment for a second time. Doing so would result in a mean reaction time smaller than three seconds because of practice and knowledge from the first attempt. Moreover, data from participants who had an accuracy score of under 1.0 were also taken out of the analysis as it indicated a lack of understanding of the task. The accuracy score was calculated by looking at how often a participant chose the right option when confronted with a *helpful* cue. In this case, there is only one option that makes sense to choose, because every map where the shaded regions and

the blue dots of the clue map don't overlap can be ruled out. All the outliers are indicated with crosses in Figure 3.

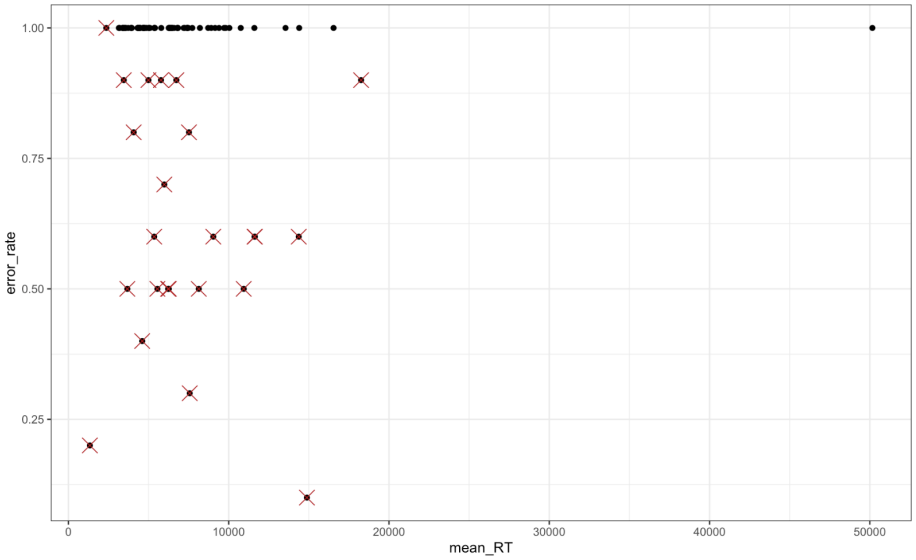

Figure 3: Outliers

Mean values of the metric variables time spent, reaction time and total time spent were computed. The complete analysis was conducted using the statistical programming language R (R Core Team, 2016) where we relied on the 'tidyverse' package for preprocessing, analyzing and visualizing the data and the 'brms' package for analyzing Bayesian regression models.

Results

Our question of interest is whether people take the intention of the sender into account when interpreting the offered evidence. Figures 4 and 5, which plot the responses of participants based on what they were told about the sender, suggest that they do indeed. The bar chart (see Figure 4) shows a clear preference for choosing the *Lure* in the teammate condition when presented with *misleading* evidence, whereas participants in the opponent condition did not show such a clear preference (see Figure 5). This finding is also supported

by the evidence ratio of 'Inf' representing strong evidence in favor of our first hypothesis. Nevertheless, most participants still chose the *Lure* in the opponent condition when presented with *misleading* evidence. This finding may indicate that people tend to trust the evidence provided by a teammate more than they tend to mistrust the evidence provided by an opponent.

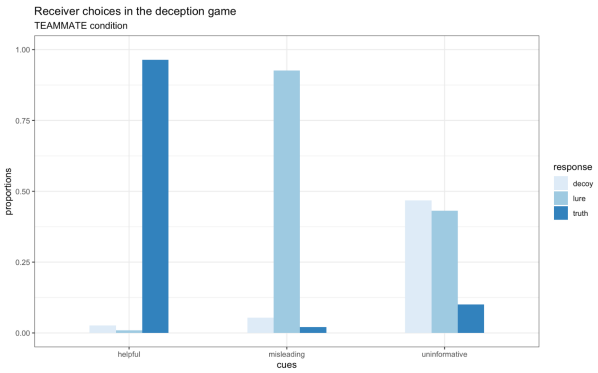

Figure 4: Teammate

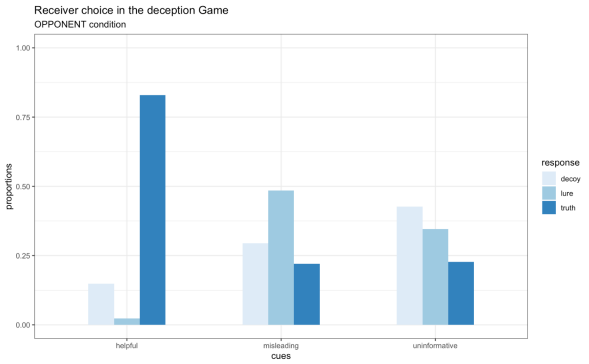

Figure 5: Opponent

The second hypothesis obtained from the original paper, namely that participants in the teammate condition are significantly more likely to choose the *Lure* when presented with *uninformative* evidence than participants in the opponent condition, can not directly be validated by looking at the charts. When we think about it there should not be a logical reason for a preference when presented with *uninformative* evidence, since none of the treasure maps can be ruled out. This is confirmed by the low evidence ratio of 1.15 for the respective hypothesis.

Discussion

The purpose of this study was to gain a better understanding of how people reason in contexts where the goals of sender and receiver are not always aligned. The results of the present study support the hypothesis that participants take the sender's (perceived) intent into account when interpreting the offered evidence. There are two key findings of the present research. First, we found that participants in the teammate condition are significantly more likely to choose the *Lure* when presented with *misleading* evidence than participants in the opponent condition. This confirms the findings by Ransom et al. (2019). Secondly, we did not find significant evidence supporting the hypothesis that participants in the teammate condition are significantly more likely to choose the *Lure* when presented with *uninformative* evidence than participants in the opponent condition. This contrasts with the findings by Ransom et al. (2019).

Although the present results support our first hypothesis, it is appropriate to recognize several potential limitations in our study design. Despite our clear instructions, first in the beginning and immediately before the main trials, that every map where the shaded regions and the blue dots of the clue map don’t overlap can be ruled out, a lot of participants seemed to either overread the instructions or didn’t understand the instructions properly. We have encountered this problem once in our pilot study and have acted on this by clarifying our instructions and by pointing the rule out multiple times instead of only once. In the original study, this problem didn’t occur. We believe this is the case because in the original study the instructions were given in real life from face to face, whereas our experiment took place online. We decided to exclude this data because we highlighted the rules and instructions multiple times, and failing the only rule that was given, shows a clear lack of understanding for the task.

This limitation of the study design could be addressed in future research. We have drafted three possible solutions to this problem. The first one is to conduct the experiment in a lab with experimenters explaining the instructions. The second one is a video version of the instructions that includes a sample trial. The third option would be to include a warning sign that reminds of the rule when a participant breaks it. The second, as well as the third option, would be applicable in an online setting such as ours. A further limitation of the study restricts the assumptions that can be drawn regarding communication. Through the focus on the receivers and the disconnect of senders and receivers, no assumptions about interpersonal relationships and interactions can be drawn. This limitation is echoed by the account of Buller & Burgoon (1996) on communication theory.

Although this study, as well as the studies we have linked to this, support our findings on deception in communication, their most important contribution may be that they raise a variety of intriguing questions for future studies. In future research, it would be useful to extend the current findings by examining the behaviour of receivers of deceptive messages in a less artificial setting, as its current applicability to real-world problems is very limited. Additionally, it would be interesting to examine how receivers behave when the intent of the sender is not as clear, as in this deception game. It would be also interesting to focus future research on the influence of the relationship between the sender and the receiver of deceptive messages.

This study has applied the computational framework provided by Ransom et al. (2019) and put parts of the original study's model of the comprehension of deception without lying in a computational study to the test in our replication. Consistent with the findings from the original study, we found that the participants showed sensitivity towards the (perceived) intent of the sender.

The original study and our replication both showed that receivers in communication do take the intent of the sender into account. This is the fundamental basis of a lot of strategic games like 'Tempel des Schreckens' or 'Poker' that rely on meta-inference in the sense that the players must construct a picture about the other players and their roles in the game. Based on this meta-inference, they make decisions on how trustworthy the statements of the other players are.

Finally, the most compelling contribution of our research is the success in reproducing the experiment by Ransom et al. (2019). The value of reproducibility becomes apparent looking at surveys such as the one conducted by *Nature* Magazine that showed more than 70% of researchers that participated in their online survey, have not been able to reproduce experiments by other scientists and that 50% were unable to reproduce their own experiment (Baker, 2016). Therefore, in the light of the current replication crisis, our success in the reproduction of the original experiment emphasizes the necessity of thorough documentation of experimental procedures in form of preregistrations to reestablish trust in cumulative science.

References

Baker, M. 1,500 scientists lift the lid on reproducibility. Nature 533, 452–454 (2016). https://doi.org/10.1038/533452a

Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. Communication Theory, 6, 203-242.

Dynel, M. (2011). A web of deceit: A neo-gricean view on types of verbal deception. *International Review of Pragmatics*, *3*, 531–538.

Grice, H. P. (1989). Studies in the Way of Words. Cambridge, MA: Harvard University Press.

Levine, T. R., & McCornack, S. A. (2014). Theorizing about deception. Journal of Language and Social Psychology, 33, 431–440.

Ransom, K., Voorspoels, W., Navarro, D., & Perfors, A. (2019, October 2). Where the truth lies: how sampling implications drive deception without lying.

Zuckerman, M., De Paulo, B., Rosenthal, R. (1981). Verbal and Nonverbal Communication of Deception.