

Supplementary Material for The Curse of Shared Knowledge: Recursive Belief Reasoning in a Coordination Game with Imperfect Information

Thomas Bolander,¹ Robin Engelhardt,² and Thomas S. Nicolet²

¹*Department of Applied Mathematics and Computer Science, Technical University of Denmark,
Richard Petersens Plads, building 324, DK-2800 Lyngby, Denmark*

²*Center for Information and Bubble Studies,
Department of Communication, University of Copenhagen,
Karen Blixens Plads 8, DK-2300 Copenhagen S.*

MTurk Walkthrough and Screenshots

After accepting our ‘Human Intelligence Task’ (HIT) and providing informed consent, participants from MTurk were put in a ‘waiting room’ until they were paired up with another participant. After a group of two was formed, participants were directed to an initial instruction page which detailed the rules of the game with a time limit of 240 seconds, see screenshot in Fig. 1. After reading the instructions, participants were directed to round 1 (of 10) where they were given their own arrival time and asked to make a decision between going to the canteen or the office. Each round had a time limit of 61 seconds and rules from the instructions were repeated on the bottom of the page, see example screenshots for round 1 in Fig. 2.

After making their decision (‘Canteen’ or ‘Office’), participants were asked to estimate how certain they were that the other player made the same choice as them, ranging from ‘very uncertain’ over ‘slightly certain’, ‘somewhat certain’ and ‘quite certain’ to ‘very certain’, see Fig. 3.

After both players have made their choices and their certainty estimates, they are prompted to a results page showing them the results of the previous rounds, including arrival times for both players, their choices, their own certainty estimate and resulting payoff, see example screenshot after round 6 in Fig. 4. After 30 seconds, the game would automatically proceed to the next round.

In many instances, groups were not able to play the maximal number of rounds, because one or both of the participants had lost all their bonuses. An example of such a situation is shown in Fig. 5. In such cases, the game would end for both players and they were asked to answer a follow-up question:

1. “The game is over. Do you think it was your fault it is over, your colleagues fault, or do you think it was because of some other reason?” (*Possible answers: ‘My fault’, ‘Other’s fault’, ‘Other reason’*)

This question probes into participants’ ability to rise above their possibly myopic understanding of the game. In addition to this question, all participants were asked three additional post-game questions about their strategy of play and their understanding of the game. The first was:

2. “What strategy did you use while playing this game?” (*open ended*)

The answers to this question provided insight into the reasoning and thoughts of the participants. The next question was used to gauge the depths of recursive reasoning and reads:

3. “Imagine you could have agreed beforehand with your colleague about a point in time where it is safe to go to the canteen. What time would that be?” (*‘I don’t know’, ‘There is no such time’, 8:00, 8:10, 8:20, 8:30, 8:40, 8:50, 9:00, 9:10*)

A final question pertaining all participant’s understanding of the concept of common knowledge was the following:

4. “Imagine you arrive at 8:10 am. Is it common knowledge between you and your colleague that it is safe to go to the canteen, that is, you both arrived before 9:00 am?” (*‘Yes’, ‘No’, ‘Don’t know’*)

The Canteen Dilemma

Time left to complete this page: 3:48

These instructions will also be shown on the following pages.

Instructions for the game:

This game is about trying to do the same as your colleague.

Every morning you arrive at work between 8:00 am and 9:10 am. You and your colleague will arrive by bus 10 minutes apart. Example: You arrive at **8:40 am**. Your colleague may arrive at **8:30 am**, or **8:50 am**.

Both of you like to meet in the canteen for a coffee. If you arrive before 9:00 am, you have time to go to the canteen, but you should only go if your colleague goes to the canteen as well. If you or your colleague arrive at 9:00 am or after, you should go straight to your offices.

At the beginning of each round you will know only your own arrival time. You will have to decide whether to go to the canteen or to the office. As a general rule, you will maximize your payoff by honestly choosing the option you think your colleague will also choose.

Payoff and penalties:

You start the game with \$10.00 and will have to pay various amounts of penalties in each round, depending on how well you both do. Your challenge is to have as much money left as possible when the game ends, after which the remaining amount is paid out to you as a bonus. The game ends after 10 rounds or when you or your colleague has no money left.

- **Both go to canteen**

If you guessed correctly that both of you went to the canteen before 9:00 am, you pay a **small** penalty proportional to how **uncertain** you were, e.g.:

- -\$0.69 if you were very uncertain.
- -\$0.29 if you were somewhat certain.
- -\$0.01 if you were very certain.

- **Both go to office**

If you guessed correctly that both of you went to your offices, no matter what time, your penalty is **doubled** and proportional to how **uncertain** you were, e.g.:

- -\$1.39 if you were very uncertain.
- -\$0.58 if you were somewhat certain.
- -\$0.02 if you were very certain.

- **One goes to the canteen, the other to the office**

If you guessed incorrectly and one of you went to the canteen while the other went to the office - or if any of you went to the canteen at 9:00 am or after, your penalty is **doubled** and proportional to how **certain** you were, e.g.:

- -\$1.39 if you were very uncertain.
- -\$2.77 if you were somewhat certain.
- -\$9.21 if you were very certain.

- In summary, try to do your best doing the same as your colleague. As a general rule you will minimize your losses by giving an honest estimate of the chances of doing the same as your colleague

Next

FIG. 1: Screenshot of instructions page.

Round 1

Time left to complete this page: **0:50**

It is Monday morning and you arrive at your workplace at 9.10 am.

Where will you go?

Canteen

Office

Instructions for the game:

This game is about trying to do the same as your colleague.

Every morning you arrive at work between 8:00 am and 9:10 am. You and your colleague will arrive by bus 10 minutes apart. Example: You arrive at **8:40 am**. Your colleague may arrive at **8:30 am**, or **8:50 am**.

Both of you like to meet in the canteen for a coffee. If you arrive before 9:00 am, you have time to go to the canteen, but you should only go if your colleague goes to the canteen as well. If you or your colleague arrive at 9:00 am or after, you should go straight to your offices.

At the beginning of each round you will know only your own arrival time. You will have to decide whether to go to the canteen or to the office. As a general rule, you will maximize your payoff by honestly choosing the option you think your colleague will also choose.

Payoff and penalties:

FIG. 2: Screenshot of choice page, round 1.

Time left to complete this page: **0:27**

How certain are you that your colleague has made the same choice as you?

- ☐ Very uncertain
- ☐ Slightly certain
- ☐ Somewhat certain
- ☐ Quite certain
- ☐ Very certain

Next

Instructions for the game:

This game is about trying to do the same as your colleague.

Every morning you arrive at work between 8:00 am and 9:10 am. You and your colleague will arrive by bus 10 minutes apart. Example: You arrive at **8:40 am**. Your colleague may arrive at **8:30 am**, or **8:50 am**.

Both of you like to meet in the canteen for a coffee. If you arrive before 9:00 am, you have time to go to the canteen, but you should only go if your colleague goes to the canteen as well. If you or your colleague arrive at 9:00 am or after, you should go straight to your offices.

FIG. 3: Screenshot of page where participants had to estimate the probability that their colleague would make the same choice.

Results (after round 6)

Time left to complete this page: 0:27

Round	You went to the	at	Your colleague went to the	at	Your certainty	Penalty
1	office	9.10	office	9.00	quite certain	-\$0.27
2	office	8.50	canteen	8.40	somewhat certain	-\$2.77
3	canteen	8.10	canteen	8.20	very certain	-\$0.01
4	office	9.00	office	8.50	very certain	-\$0.02
5	canteen	8.10	canteen	8.20	very certain	-\$0.01
6	canteen	8.30	canteen	8.20	quite certain	-\$0.13

You both chose to go to the canteen. Good work!

You lost **-\$0.13** this round. Your total losses are **-\$3.21**. You still have **\$6.79**.

Next

FIG. 4: Screenshot of results page, round 6.

Results

Round	You went to the	at	Your colleague went to the	at	Your certainty	Penalty
1	office	9.10	office	9.00	quite certain	-\$0.27
2	office	8.50	canteen	8.40	somewhat certain	-\$2.77
3	canteen	8.10	canteen	8.20	very certain	-\$0.01
4	office	9.00	office	8.50	very certain	-\$0.02
5	canteen	8.10	canteen	8.20	very certain	-\$0.01
6	canteen	8.30	canteen	8.20	quite certain	-\$0.13
7	canteen	8.20	canteen	8.30	somewhat certain	-\$0.29
8	office	9.00	canteen	8.50	very certain	-\$9.21

You went to the office while your colleague did not.

Unfortunately, you lost all your money, so the game is over.

Your bonus is **\$0.00**.

Thank you for your participation.

Next

FIG. 5: Screenshot after a player has lost all her bonus.

Experiment	Participants	Attrition rate	N	Rounds	AvgPayout (\$)
MTurk	714	0.02	680	10	4.36
DTU1	106	0.13	80	30	(prizes)
DTU2	50	0.08	42	30	-

TABLE I: The main experiment on Amazon Mechanical Turk (MTurk) had 714 participants of which 17 participants (2,4 %) quit prematurely, some of them quite early in the game. These quitters were told (in the consent form) that they would receive no bonus and no participation fee. They are excluded from the data analysis. Their “lucky” colleagues however, got both their bonuses and participation fee, but are likewise excluded from the data analysis. Therefore the final number of subjects, N , is reduced by twice the attrition rate. In the two DTU experiments with students from the Technical University of Denmark (DTU1 and DTU2) attrition rates were slightly higher, mainly due to the higher number of rounds played.

DTU Experiments

The DTU students were asked three additional post-game questions:

5. “Did you ever go to the canteen at an arrival time later than what was safe according to your previous answer? Why or why not?” (*open ended*)
6. “Did you ever choose differently after seeing the same arrival time again at a later point in the game? Why or why not?” (*open ended*)
7. “Imagine you arrived at [8:40/9:00] and you have been secretly informed that your colleague’s arrival time is 8:50. Where do you think your colleague will go?” (*‘Canteen’, ‘Office’*)

In the last question, half of the participants were given 8:40 as their own arrival time while the other half were given 9:00. The question concerns whether player’s own knowledge of the other’s arrival time affect their prediction of the other player’s decision. It relates to the curse of knowledge [1] since participants might attribute their own belief (that it is early enough or too late to go to the canteen) to the other player.

MTurk Settings

Looking at Table I, the average payout to MTurk-workers was \$4.17 (including a general participation fee of \$2) which amounts to an average of more than \$20 per hour, which is considered very generous according to MTurk guidelines and certainly above the recently estimated average of \$6 per hour when excluding un-submitted and rejected work [2]. Students in the DTU experiments (DTU1 and DTU2) did not receive any monetary reward, but were told to try to maximize their payoff, and awarded prizes for doing well.

Participants quitting a study before completing it is prevalent on MTurk, and varies systematically across experimental conditions [3]. In our experiments on Amazon Mechanical Turk attrition rates were 2%, witnessing that we had managed to design the experiment in a way that minimized drop-out rates. A combination of high payouts, a logarithmic scoring rule taking advantage of loss aversion biases, a consent form stipulating the revocation of the participation fee after dropout, and minimization of waiting times may have been the main reasons.

All participants automatically received a ‘qualification’ when accepting a HIT. This qualification ensured that participants could not play the game again. In addition, we required that participants should have had completed at least 500 HITs, have an accepted HIT rate of 98% or above, and should be from the United States or Canada. This ensured that we would get relatively experienced and qualified participants.

MTurk participant attention was expected to be equal to or better than undergraduate participant’s attention [4], while various forms of dishonesty (practical joking or trying to pair up with a friend) was expected to be rare, due to the high turnover rate experienced for our HITs. In addition, during the experiment, participants had easy access to our email for questions and possible bug reports. Apart from a few timeouts, participants had no comments or complaints.

Formal definitions of Private, Shared and Common Knowledge

We can define the notion of common knowledge and related notions a bit more precisely as follows, following the conventions from epistemic logic (see e.g. Herzig and Mauffre [5]). Given a proposition p and an agent i , we use $K_i p$ to denote that agent i knows p . Given a group of agents $G = \{1, \dots, m\}$, we say that p is *private knowledge* in G if at least one of the agents know p , that is, if $K_i p$ is true for some $i \in G$. We use $E_G p$ to denote that everybody in G knows p , that is, for all $i \in G$, it is true that $K_i p$. Whenever it is not necessary to be explicit about the group of agents G , we will just write $E p$ and say “everybody knows p ”. For all n , we then recursively define $E^n p$ to be shorthand for $E E^{n-1} p$, where $E^1 p$ is shorthand for $E p$. So for instance $E^2 p$ expresses that “everybody knows that everybody knows that p ”, and in general $E^n p$ means we have n iterations of “everybody knows that” in front of p . We read $E^n p$ as “everybody knows p to depth/order n ”. We also call this *shared knowledge* (of p) to depth/order n , or *n th-order shared knowledge*. When we say that p is *shared knowledge*, we mean that it is shared knowledge to depth n for some $n \geq 1$. *Common knowledge* of p then means that $E^n p$ is true for all $n \in \mathbb{N}$.

In epistemic logic, the three notions—private, shared and common knowledge—are usually not considered to be mutually exclusive. So if p is common knowledge, it is also automatically both shared and private, since when the conditions for p being common knowledge are satisfied, also the conditions for it being shared and private are satisfied. However, in many cases, as in our paper, we want to make an exclusive distinction between the three types of knowledge. We can define p to be *shared knowledge only* if it is shared knowledge but not common knowledge. Thus, p is shared knowledge only if for some n we have $E^n p$ but not $E^{n+1} p$. Similarly, we can say that p is *private knowledge only* if p is private but not shared knowledge. Thus, p is private knowledge only if $K_i p$ holds for some, but not all, i . In most texts, as in ours, it is left implicit whether private and shared knowledge are interpreted inclusive or exclusive, that is, one doesn’t explicitly distinguish between “shared knowledge” and “shared knowledge only”. Normally it is clear from the context whether one intends the concept to be interpreted exclusively or inclusively.

In our paper, we interpret the concepts exclusively, although we make an exception for private knowledge. When p is known by all agents, we say that p is both private and (first-order) shared knowledge. The exact border between private and shared knowledge vary significantly between different papers. De Freitas et al. [6] consider the case $E p$ to still only be private knowledge, and for p to be considered shared knowledge furthermore requires that there is at least one agent i knowing $E p$ to be true (that is, requires $K_i E p$ to be true for some $i \in G$). The point of De Freitas et al. is that if only $E p$ is true, it is not really shared knowledge, but only private knowledge held by everyone in G . In our paper, we have sought a compromise between the terminology by De Freitas et al. and the standard terminology in epistemic logic, and hence we have the overlap between private and shared knowledge.

Supporting data analysis

When a group experiences a round of miscoordination, we expect some kind of learning to take place. ‘Why did my colleague choose differently than I did?’ should be an obvious question a player asks herself, prompting deeper perspective-taking and possibly an understanding of the lack of common knowledge. We investigate this by partitioning decisions into those in which a participant never before has experienced a miscoordination with her colleague ($m = 0$) and those in which a participant has experienced one or more miscoordinations ($m > 0$). Furthermore, since the number of rounds - and hence miscoordinations - are very different for MTurk participants and DTU students, results from MTurk and DTU are shown separately, with the two DTU experiments combined. Results for the choices in Fig. 6 show significant differences between MTurk-groups having miscoordinated or not, while for DTU students those differences are not significant. Looking into the corresponding certainty estimates, as shown in Fig. 7, we see a similar pattern as before: Much steeper profiles

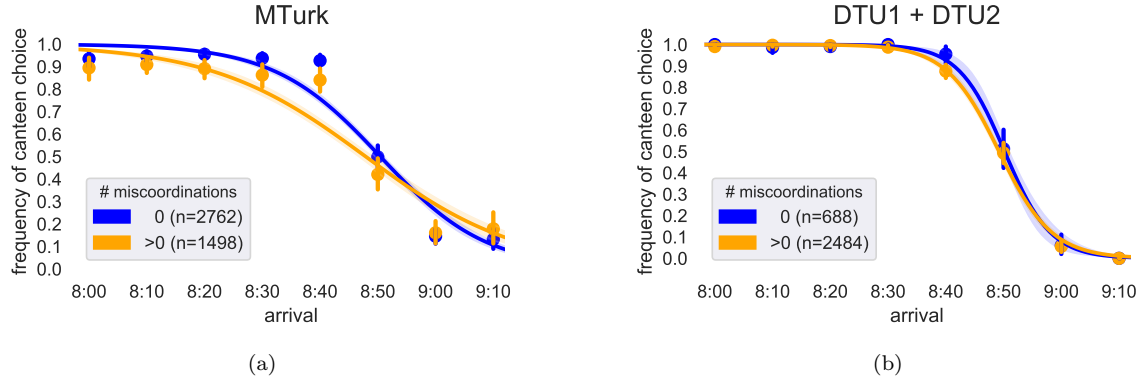


FIG. 6: Participant's decisions of going to the canteen as a function of their arrival time, here partitioned into those groups who previous have experienced zero (blue) or one or more (orange) miscoordinations. MTurk participants are shown one the left, DTU students on the right

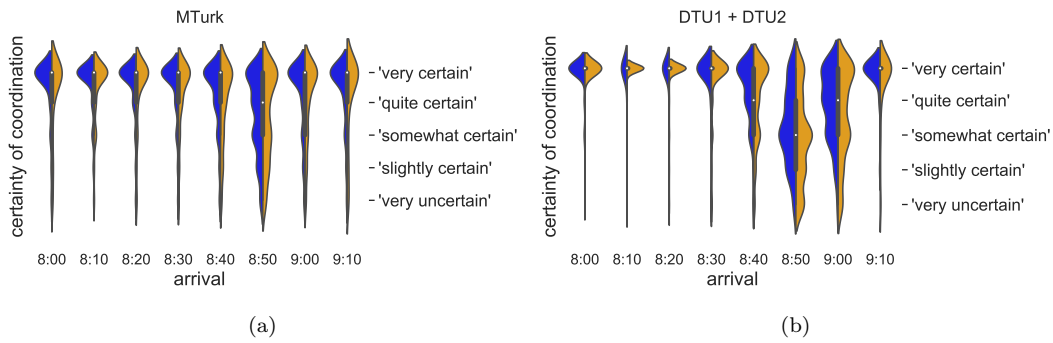


FIG. 7: Participant's certainty estimates of being able to coordinate with their colleague as a function of their arrival time, partitioned into those groups who previous have experienced zero (blue) or one or more (orange) miscoordinations. MTurk participants are shown one the left, DTU students on the right

among DTU students, but also much lower certainty estimates around critical arrival times.

-
- [1] S. A. Birch, P. Bloom, The curse of knowledge in reasoning about false beliefs, *Psychological Science* 18 (2007) 382–386.
 - [2] K. Hara, et al., A data-driven analysis of workers' earnings on amazon mechanical turk, *Proc. of the 2018 Conference on Human Factors in Computing Systems ACM* (2018) 449.
 - [3] H. Zhou, A. Fishbach, The pitfall of experimenting on the web: How unattended selective attrition leads to surprising (yet false) research conclusions, *Journal of Personality and Social Psychology* 111 (2016) 493–504.
 - [4] D. G. Rand, P. Bloom, The promise of Mechanical Turk: how online labor markets can help theorists run behavioral experiments, *Journal of Theoretical Biology* 299 (2012) 172–9.
 - [5] A. Herzig, F. Maffre, How to share knowledge by gossiping, *Springer* (2015) 249–263.
 - [6] J. De Freitas, K. A. Thomas, P. DeScioli, S. Pinker, Common knowledge, coordination, and strategic mentalizing in human social life, *Proceedings of the National Academy of Sciences* 116 (2019) 13751–13758.