

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.*

Common knowledge is a necessary condition for safe group coordination. When common knowledge can not be obtained, humans routinely use their ability to attribute beliefs and intentions in order to infer what is known. But such shared knowledge attributions are limited in depth and therefore prone to coordination failures, because any finite-order knowledge attribution allows for an even higher order attribution that may change what is known by whom. In three separate experiments we investigate to which degree human participants (N=802) are able to recognize the difference between common knowledge and *n*th-order shared knowledge. We use a new two-person coordination game with imperfect information that is able to cast the recursive game structure and higher-order uncertainties into a simple, everyday-like setting. Our results show that participants have a very hard time accepting the fact that common knowledge is not reducible to shared knowledge. Instead, participants try to coordinate even at the shallowest depths of shared knowledge and in spite of huge payoff penalties.

Introduction

Successful group coordination requires complementary choices among group members, which, in turn, requires communication of beliefs and intentions in such a way that they become common knowledge [1]. A fact is said to be *common knowledge* if everyone knows it, and everyone knows everyone knows it, and everyone knows everyone knows everyone knows it, and so on, ad infinitum [2–5]. If the premise of “everyone knows” is not infinitely nested, but only nested to finite depth, we instead have *shared knowledge*. If there is no nested knowledge about knowledge at all and not everyone necessarily knows the fact, we have *private knowledge*.

Let us illustrate the difference between these notions with an example: Two friends, Agnes and Bertram, are taking different trekking routes to the top of a mountain. In the morning they agree that if the weather gets bad, they will go back down and sleep in the mountain hut at the base. Otherwise, they strongly prefer to stay overnight at the top. The equipment essential for an overnight stay at the top has been divided between their backpacks. It is therefore crucial that if one of them decides to go to the top, the other one does the same. On the way to the top, they both observe a thunderstorm approaching, but are uncertain about whether the other person has seen it. At this point they both know the fact that “a thunderstorm approaches”, but don’t know whether the other knows. In this situation, we would say that Agnes and Bertram both have *private knowledge* that a thunderstorm approaches, and since they both know it, it is also *shared knowledge* between them. More generally, a fact is *private knowledge* in a group of agents if some non-empty subset of the agents know the fact. For the special case where everybody in the group knows it, we say that there is *shared knowledge to depth one* (or *first-order shared knowledge*) of the fact [3].

Since Bertram doesn’t know whether Agnes knows about the thunderstorm, he would like to warn her. So he sends a text message: “Thunderstorm approaching. Let’s meet at base.” However, due to the unstable mobile network signals, he is not certain that the message will go through. Therefore he asks Agnes to confirm that she has received the message. A few minutes later, he receives her confirmation. At this point it has become *shared knowledge to depth two* (or *second-order shared knowledge*) that a thunderstorm is approaching: She knows that he knows, since she received his message, and he knows that she knows, because he received her confirmation. To have *shared knowledge to depth three* (*third-order shared knowledge*), it would additionally be required that A) he knows that she knows that he knows, and B) that she knows that he knows that she knows. In fact, A already holds, since she confirmed receiving his message. However, B doesn’t hold, since she will be uncertain about whether her confirmation was received in good order. Thus we have an asymmetry in the level of knowledge of the two agents. If she also asks him to confirm her message, and she receives such a confirmation, then of course she will get to know that he knows that she knows. Then there will be shared knowledge to depth three. However, there will still be a (higher-order) knowledge asymmetry, since Bertram can’t be certain that the last message was received.

How many messages back and forth does it take for Agnes and Bertram to coordinate going back to the base in the evening? At first it might seem that it is sufficient for both of them to know that at least one of them plans to go to the base. However, that is not so. After the first message has been received, both know that Bertram plans to go to the base, but he is still uncertain whether she knows. And if she doesn’t, he might risk leaving her alone at the top. So shared knowledge to depth one is clearly not sufficient. Even shared knowl-

edge to depth two is not sufficient. After she has confirmed receiving his original message, from her perspective it is still entirely possible that he doesn't know that she knows, and hence, he might decide to go to the top to not leave her alone. And in this case, she also has to go to the top in order not to leave him alone. This argument can be generalised to prove that even n th-order shared knowledge for any arbitrary large n is insufficient for safe coordination.¹ The consequence is that no finite number of messages successfully delivered will guarantee that Agnes and Bertram manage to meet at the base. In order to guarantee meeting at the base, they would need to have shared n th-order knowledge for *all* $n \in \mathbb{N}$. Shared n th-order knowledge of some fact for all $n \in \mathbb{N}$ is called *common knowledge* of the fact. For a more formal definition of private, shared and common knowledge, see the Supplementary Information.

In practice, shouldn't it be possible to coordinate meeting at the base after one or two messages being sent back and forth? Isn't it a purely mathematical problem with no practical implications for humans trying to coordinate their actions? One of the main goals of this paper is to argue that the answer to both questions is no. In order to make that argument, we have designed a coordination game that has a similar underlying mathematical structure as the example just given, but cast in a simpler, more everyday-like setting, where the higher-order uncertainty is established already when the game is initialised and not, as above, through a series of message passings. The game was originally developed to illustrate how the difference between shared and common knowledge can have a real impact on human behaviour. Humans generally find the concept of common knowledge hard to grasp, and even harder to grasp the practical relevance of, due to the unbounded nesting of knowledge involved. The game we developed intends to make it clear that human intuitions about common knowledge can be misleading

¹ To derive a contradiction, suppose n th-order shared knowledge for some n is sufficient to make it safe to go to the base, in the sense of guaranteeing that the other person will also go there. This implies that the successful delivery of the n th message is sufficient to guarantee meeting at the base. There must then exist a smallest number n_0 such that the successful delivery of the n_0 th message is sufficient to guarantee meeting at the base. Since n_0 is the smallest such number, the successful delivery of the $(n_0 - 1)$ st message is not sufficient to guarantee meeting at the base. Now note that at the moment when the n_0 th message has been successfully delivered, the sender of this message is still uncertain about whether it was actually received, and hence the sender is only certain that the first $n_0 - 1$ messages has been successfully delivered. In other words, the sender of the n_0 th message considers it possible—even after the successful delivery of the message—that only the first $n_0 - 1$ messages were successfully delivered, in which case it is *not* safe to go to the base. Hence that person will after having sent the n_0 th message still consider it unsafe to go to the base, and will choose to go to the top. This is a contradiction, completing the proof. This proof is stated in rather informal terms, but can be turned into a formal, mathematical proof [1].

and may have costly consequences.

The curse of shared knowledge

Reasoning about the knowledge of others, their reasoning about you, and your reasoning about their reasoning, and so on, is famous in cognitive science for its presumed computational intractability [6]. Because of this, coordinating species typically use heuristic shortcuts in order to work with nested knowledge states like common knowledge, such as joint perceptual cues and broadcasted signals [7–9]. Humans may obtain common knowledge via mutually accessible first-order sensory experiences [10–13], eye contact [14], public rituals and conventions [2], or salient focal points [15]. What most prominently is believed to distinguish human coordination from other animals, however, is the enormous flexibility by which humans can imagine and articulate the mental states of their peers [16, 17]. The abilities to blush, to tell jokes, and to write novels testify that humans readily attribute *higher-order* beliefs, intentions, and reasoning capabilities to other people, such as thinking explicitly about the mental states of others who think about the thoughts and beliefs of others and so on, while at the same time appreciating that those thoughts and beliefs can differ from each other and from reality. Such higher-order cognition has seen substantial scientific attention, and has brought about various technical terms such as “theory of mind” (ToM) [18], “mentalizing” [19], “mind reading” [20, 21], “mental models” [22], “mind perception” [23], “perspective taking”, and “social intelligence” [24], which often are used interchangeably for studying the cognitive mechanisms of shared knowledge, but sometimes focus on slightly different ideas and associated meanings [25], depending on the field of investigation.

Looking at the ToM-literature, conclusions about human belief reasoning abilities are rather heterogeneous [26, 27]: Human reasoning about the reasoning processes of other humans is limited [18, 28–33], contextual, and possibly domain specific [34–36]. Three-year-old children tend to fail in the well-known first-order false belief tasks by falsely assuming that their private information is shared by others, while second-order false belief tasks are mastered around ages 5–7 [37, 38]. Adults may reliably master up to four orders [39], but still have difficulties ignoring the private information they possess when assessing the beliefs of others, resulting in a *curse of knowledge* bias which can compromise their ability to make predictions about other people's beliefs and actions [1, 40]. When the nested mental states represent a succession of different people, such as “Alice thinks that Bob thinks that Carol is contemplating the idea that David is thinking about Evelyn”, we have less problems following along² than when the nested mental states

² Especially when those successions of mental states are qualified by psychological attribute words such as ‘Alice thinks that Bob is

are successions of the same people over and over again, and thus are truly recursive, such as “I think that you think that I contemplate the idea that you are thinking about me”. We get confused more easily by the latter formulation, since we need to keep track of several representations of ourselves and of the other, each representation differing in its perspective and in the number of mental states it presupposes [6]. When humans compete or try to detect cheaters, higher-order belief reasoning seems to perform better than lower-order belief reasoning [44]. In negotiations and other mixed motive situations, where innuendo, threads, bribes and other kinds of indirect propositions are common, humans are very good at the *strategic* use of higher-order belief reasoning, for instance as a means to *prevent* common knowledge in certain groups of agents, or as a means to form specific knowledge alliances [45–47]. ToM proficiency may also be facilitated by providing games with stepwise increase in ToM [48].

In pure coordination problems, such as pedestrians choosing sides, or people agreeing on new words or on new technical standards, common knowledge is the preferred informational state for all members of the group, because it ensures that all sides find an optimal common equilibrium. If there are no or limited means by which to communicate, however, people face an equilibrium-selection problem for which neither game theory nor the ToM literature has any clear solution. Although some experimental evidence [49] suggests that higher-order ToM reasoning may improve coordination efforts, other work seems to suggest that coordination favours lower orders of ToM sophistication [50, 51]. The challenge of tacit coordination is particularly relevant for artificial intelligence research and for social cognitive robotics, where the implementation of ToM-like processes into artificial social agents is believed to be an important step towards reliable human-robot interaction [52–55].

Recently, researchers have investigated whether humans have adapted specifically to recognizing common knowledge as a separate cognitive category, distinct from both private and shared knowledge [6]. Controlled pure coordination experiments in social settings on market collaboration [56], the bystander effect [57], indirect speech [58], self-conscious emotions [59], and charity [60], consistently find that people indeed make strategically different choices under common knowledge conditions (typically presented in the form of public announcements), compared to situations in which there is only private knowledge (in the form of private messages) or shared knowledge (private messages that elaborate on the depth of knowledge of other participants). Apart from seeing a clear benefit of common knowledge, some of these studies also showed that people have a hard time discriminating

between various orders of shared knowledge, and that coordination efforts do not correlate with payoff conditions [56], which is in contrast to the assumptions of standard rational choice theory in which payoffs are expected to be maximized [61].

So if humans indeed have adapted to recognize common knowledge *in the wild*, the question remains if they are also able to recognize the difference between common knowledge and n th-order shared knowledge for some (potentially large) n . In other words, while humans are able to reliably detect proper common knowledge in a wide range of situations, how good are they at refraining from inferring common knowledge in situations with only n th-order shared knowledge? We do not know, in part because many existing experimental designs stop after describing 2-3 orders of belief reasoning to the participants, as higher orders require quite convoluted sentences that tend to become incomprehensible and increase experimental error. Or, as in the mountain trekking example, they require reasoning about the consequences of a high number of (message passing) actions that each change the mental state of the involved agents.

The latter has been explored theoretically in the ‘electronic mail game’ by Rubinstein [62], a game version of the mountain trekking example presented above (and of the structurally equivalent ‘coordinated attack problem’ [1]). The Rubinstein paper shows that ‘almost common knowledge’ in the sense of n th-order shared knowledge for some large n , leads to a very different expected player behaviour than ‘absolute common knowledge’. Essentially his conclusion, translated into the context of the mountain trekking example, is that common knowledge will make the two mountain hikers both go to the base, whereas if there is only n th-order shared knowledge for some n , then both will meet at the top, independent of n and despite the bad weather condition (resulting in non-maximal payoffs). Rubinstein does a pure game-theoretic analysis of the game with no experiments, and only speculates what people playing the game might do. We find it interesting to dig deeper into how humans would play and reason about such games in practice. What would their intuition recommend them to do? Which depth of shared knowledge (if any) would be enough to attempt risky coordination that would lead to maximal payoff if successful? How would they be certain that the person they try to coordinate with thinks that the same depth is sufficient?

The electronic mail game and the mountain trekking example are complicated in terms of the dynamics of iterated message passing. In this paper, we devise a novel game in which the higher orders of shared knowledge are not achieved dynamically via actions, but are already present at the beginning of the game, using uncertainty about arrival times. This, we believe, makes the game easier to understand. Letting humans play our game, we have been able to address the previous questions in more detail. Our results show that people indeed have a very hard time accepting the fact that common knowl-

mistakenly worrying that Carol is offended by misunderstanding something Dave had said to Evelyn [42].

edge is not reducible to shared knowledge of finite depth. On the contrary, participants try to coordinate even at the shallowest depths of shared knowledge and in spite of huge payoff risks. The reason, we believe, is that the sole presence of shared knowledge is enough to make participants try to coordinate, and that moderate depths of shared knowledge become effectively indistinguishable from common knowledge due to the recursive nature of the game. We call this effect “the curse of shared knowledge” because even small depths of shared knowledge raises the participant’s expectation of being able to coordinate in spite of repeated payoff penalties for having miscoordinated before.

Experimental Design

The experiment is designed as a two-player coordination game with imperfect information. The game is inspired by the structure of the consecutive number riddle, also called the Conway paradox, see e.g. van Emde Boas et al. [63], van Ditmarsch and Kooi [64]. Our game is framed as an everyday situation, where two colleagues arrive at their workplace in the morning, and have to decide whether to meet in the canteen for a morning coffee or go straight to their offices and start working immediately. We call the game the ‘Canteen Dilemma’. The purpose of framing it in an everyday situation is to attempt to make some of the recursive reasoning easier to comprehend [65, 66]. The introductory story of the game goes as follows:

“Every morning you arrive at work between 8:10 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 cup of 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.”

The game has 10 rounds on MTurk and participants are told that at the beginning of each round they will know only their own arrival time, and based on this will have to decide whether to go to the canteen or the office. After choosing an option, participants are asked to estimate their certainty that their colleague will choose the same option (on a five-point Likert scale). We call the value chosen the *certainty estimate* of the participant. A fixed participation fee of \$2 is given to all players who finish the game. Additional bonuses are calculated with a logarithmic scoring rule whereby each participant is given an initial bonus of \$10, which is then reduced by a variable penalty in each round, depending on the players’ decisions and certainty estimates.

The game has three possible outcomes: 1) both choose the canteen which we refer to as *coordination into the canteen*; 2) both choose their respective offices which we

refer to as *coordination into the offices*; 3) one chooses the canteen and the other chooses the office which we refer to as *miscoordination*. Penalties are tiered in such a way that a small penalty is deducted for successful coordination into the canteen (achieving the highest payoff), which is doubled for coordination into the offices (achieving the second-highest payoff), while the penalty for miscoordination or forbidden choices, i.e. going to the canteen at 9 am or after, is much larger (up to 921 times larger, meaning a significantly lower payoff than the previous two). See Materials and Methods for details about the payoff structure. In the instructions shown to the participants beforehand, we also include payoff examples of both successful and failed coordinations. Screenshots and full descriptions of the experimental setup can be found in the Supplementary Information.

For the main experiment, we recruited a total of 680 participants from Amazon Mechanical Turk (MTurk) to play for a maximum of 10 rounds, making a total of $n = 4260$ choices. In addition, we conducted two supplementary classroom experiments with 80 students (DTU1 - $n = 2160$) from the Technical University of Denmark (DTU) taking a course on Artificial Intelligence and Multi-Agent Systems, and 42 additional student (DTU2 - $n = 1012$) taking an introductory course in Artificial Intelligence. The two classroom experiments differed slightly from the MTurk experiment in that the students got an initial bonus (endowment) of \$30 instead of \$10, and played 30 rounds instead of 10. Also, in the classroom experiments, all students were told that they would not receive any monetary rewards for playing the game, but they should still try to do their best. The students also had to answer a few additional post-game questions, see the Supplementary Information for a full list of those questions.

Game Strategies

What are the relevant strategies for this game? First note that going to the canteen at 9:00 or after results in the worst possible payoff. So both players should always go to the office if they arrive at 9:00 or after. How about if both arrive strictly before 9:00? If both choose canteen, they get a better payoff than if both choose office. Now consider a case where you are one of the players, and you arrive at 8:50. Then your colleague will be arriving at either 9:00 or 8:40. If your colleague arrives at 9:00, she has to choose office according to the previous argument, and then you would have to choose office as well to avoid the large penalty of miscoordination. However, if your colleague arrives at 8:40, you may both choose the canteen, and this will lead to the highest payoff. In other words, depending on the arrival time of your colleague, a piece of information that you don’t have access to, the best choice is either office or canteen. So which one to choose?

Since the penalty of miscoordination is very high, it would seem best to choose office. What if you then instead arrive at 8:40? In this case, your colleague either

arrives at 8:30 or 8:50. In both cases, you have time to meet for a cup of coffee in the canteen, and doing so will give you the highest payoff. At first, it might seem like an easy choice. However, we just concluded that the best strategy at 8:50 would be to go to the office. So, if you arrive at 8:40 and contemplate that your colleague might arrive at 8:50—and if you believe your colleague would reason as yourself and go to the office at 8:50—you also ought to go to the office at 8:40. This argument can of course be iterated, because if the optimal choice at 8:40 is to go to the office, then the optimal choice at 8:30 must also be to go to the office. In other words, the optimal strategy seems to be to always go to the office, independent of arrival time! And, indeed, so it is. If both players go to the office in all rounds and declare the highest possible certainty in their decision, they will both leave the experiment with \$9.80, excluding the \$2 participation fee. This is the highest possible payoff that can be guaranteed by any strategy in the game, and very close to the \$10 that the players start out with. As we will see later, the payoffs that people actually get when playing the game are significantly lower than this.

The *all-office* strategy described above, where you always decide to go to the office independent of arrival time, is a safe strategy if both players follow it. By safe is meant that there is never any risk of miscoordination, and hence no risk of getting the highest penalty (the penalty for miscoordination is up to \$9.21 in a single round). It is actually the *only* safe strategy. The reason is that if at least one of the players, say *a*, has the strategy of going to the canteen at some time *t* before 9:00, then since they both have to go to the office at 9:00 or later, there must exist at least one pair of arrival times for which the two players are miscoordinated.³

The fact that the all-office strategy is the only safe one is counter-intuitive to most people before being presented with the proof, and for some people even after. The issue is that, intuitively, it would seem to be safe to go to the canteen at, say, 8:30. Why would you ever go to the office that early? You know that your colleague will then be arriving at 8:40, which is still plenty of time to get a cup of coffee before 9:00. The issue is of course that if you take the perspective of your colleague, then your colleague arriving at 8:40 will consider it possible that you arrived at 8:50. And if you had indeed arrived at 8:50, you would consider it possible that your colleague had arrived at 9:00. In that case you would be forced to

choose the office. A major point of our experiments is to test whether this kind of recursive perspective-taking is utilized by human players of the game.

The argument of the all-office strategy being safe of course relies on the other player following the same strategy. Since we don't allow players to agree on a strategy with their co-player beforehand, the all-office strategy doesn't necessarily in practice lead to the highest payoff for a particular player. Another issue is that one might decide to play risky instead of safe. Consider the *canteen-before-9* strategy of always going to the canteen before 9:00 and going to the office at later times, all with the highest certainty estimate. If both players choose this strategy and are fortunate to play 10 rounds without any of them arriving at 9:00 or later, they will get the highest possible payoff of \$9.90—slightly higher than the guaranteed payoff \$9.80 of the all-office strategy. However, if all pairs of arrival times are equally likely, the probability of miscoordination is then 1/6 (there are 12 pairs of arrival times in total, and 2 of those have one player arriving at 8:50 and the other at 9:00). Miscoordination with the highest possible certainty estimate gives a penalty of \$9.21, so in practice this strategy is of course still significantly worse than the all-office strategy, even if only playing one round (the expected payoff for a single-round game will be $\$10.00 - \$9.21 \cdot 1/6 - \$0.01 \cdot 5/6 = \8.46).

In the Methods section we make the reasoning about game strategies formally precise. We show that independently of the particular payoff structure (only using the order of the payoffs, not their exact values), there will only be two candidates for the optimal strategy, the all-office strategy and the canteen-before-9 strategy. Which one is then optimal depends on the particular payoff structure and the number of arrival times before and after 9 am. In our specific experiments with our specific arrival times and payoff structure, the all-office strategy has a significantly higher expected utility than the canteen-before-9 strategy, as already argued. As we will see, the human players in our experiments very rarely play any of these strategies, but seem to believe that it is safe to go to the canteen if arriving sufficiently ahead of 9 am, e.g. before 8:50, but unsafe when arriving later.

A strategy to always go to the canteen if arriving before some cut-off time t_c and always go to the office if arriving after t_c is called a *cut-off strategy* (with cut-off t_c). The canteen-before-9 strategy is a cut-off strategy with cut-off 8:55 (see Materials and Methods for more details).

Results

The maximal theoretical payoff described in the previous section was never observed in the experiments—actually quite far from it, despite doing the experiment with more than 800 people. Recall that the payoff of the all-office strategy is \$9.80 independent of arrival times. The average bonus paid to our MTurk participants was a mere \$2.36. Due to the penalty-based payoff structure, only 46 out of 340 MTurk groups (14%) were able to play 10 rounds and still have any bonus left, while the average

³ Since *a* chooses to go to the canteen at time *t*, player *b* also has to go to the canteen at time *t* + 10, since otherwise they would be miscoordinated when *a* arrives at *t* and *b* at *t* + 10. But if *b* goes to the canteen at time *t* + 10, *a* also has to go to the canteen at time *t* + 20, since otherwise they would be miscoordinated when *b* arrives at *t* + 10 and *a* at *t* + 20. This can be generalized to conclude that *a* would have to go to the canteen at any time *t* + 20*x* for *x* ≥ 0 and *b* would have to go to the canteen at any time *t* + 10 + 20*y* for *y* ≥ 0. Clearly this implies going to the canteen after 9:00.

number of rounds played was 6.3, see Table I. As soon as one of the players had no money left, the game would terminate.

Exp.	N	R	\bar{r}	Ruin (%)	Payoff (%)	\bar{s} (\$)
MTurk	680	10	6.3	52.8	23.6	-1.59
DTU1	80	30	27.0	17.5	27.0	-0.83
DTU2	42	30	24.1	31.0	24.1	-0.98

TABLE I: Exp. = experiment; N = number of subjects; R = maximum number of rounds; \bar{r} = average number of rounds played; Ruin = percentage of participants losing all their bonus before (or in) round R; Payoff = average earnings (given as the retained percentage of the initial endowment); \bar{s} = average penalty per player per round.

Comparing the MTurk experiment with the DTU experiments in Table I, shows that the DTU participants were slightly better on average. While more than half of the MTurk participants had lost their initial bonus and had to end the game before the last round, only 18% and 31% of the DTU participants, respectively, had done so. Especially the students from the Artificial Intelligence and Multi-Agent Systems course (DTU1) managed well by retaining 27% of the initial endowment and losing only \$0.83 per round on average. This may come as no surprise, because these students were later into their studies, and had already been taught about social cognition.

Looking at Figure 1, we see the frequency of participants' canteen choices as a function of their arrival time together with a fitted binary logistic regression line. While DTU students (orange and green) show similar steep profiles, MTurk participants have a slightly more gradual decline in canteen choices for increasing arrival times. However, the point at which there is a 50% probability of choosing the canteen or the office (see Materials and Methods) is close to 8:50 in all three experiments. In the Discussion we therefore combine all three experiments in Figure 5 in order to understand the experimental results in terms of degrees of shared knowledge.

Figure 2 shows the distribution of certainty estimates for each arrival time. It clearly shows that it is exceedingly rare for any of the participants to consider it even problematic to go to the canteen when arriving early. Arriving at 8:40 or earlier is deemed sufficiently early to visit the canteen with high confidence, arriving at 9:00 or later is deemed office time with high confidence, and arriving at 8:50 is deemed either office or canteen with at least being "somewhat certain". The difference in certainty estimates between MTurk participants and DTU students show that the latter tend to be more certain that their co-players follow a similar strategy (higher certainty estimates for the early and late arrival times), and also that they are more aware of the danger of miscoordination (lower certainty estimates around the cut-off). This is in particular the case for the DTU1 experiment that has the steepest profile. Being more certain that your co-players follow a similar strategy probably indicates that

you believe such a strategy to be optimal. So, interestingly, the DTU1 participants are both the ones that appear to be most aware of the danger of miscoordination, and at the same time those who most firmly believe a cut-off strategy is optimal, i.e., believing that the risk of miscoordination is unavoidable. The differences between the three experiments are however still relatively minor, and in the following we will combine data from all three experiments.

Group dynamics

Starting with the group dynamics in Figure 12, we see the number of successful group coordinations into the canteen/office (green/purple) together with the number of miscoordinations (red) as a function of all possible arrival time combinations. The figure shows clearly that players are able to coordinate into the canteen more than 80% of the time if both of them arrive before 8:50. As soon as a group has a player who arrives at 8:50, however, the result changes drastically. Suddenly almost half of such groups miscoordinate. As players experience harsh penalties for miscoordinating, one could perhaps expect to see a tendency of choosing office more often when arriving at 8:40 or 8:50 in subsequent rounds. That is, we might expect that players learn and converge to the all-office strategy in order to avoid miscoordination altogether. But this is not what we see.

Figure 4 shows the mean frequency of canteen choices as a function of rounds played for all three experiments. Each color corresponds to a certain arrival time. Clearly, the only arrival times that do not converge towards either the canteen or the office are the arrival times of 8:40 and 8:50, with the former fluctuating around 90% canteen choices and the latter fluctuating around 50% canteen

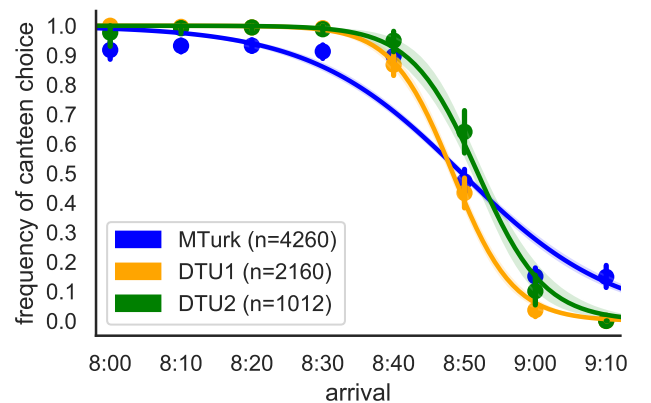


FIG. 1: Frequency of canteen choices as a function of arrival times. Circles indicate the mean frequency of participants choosing the canteen at a certain arrival time with error bars. Colored lines are logistic regression lines with bootstrapped 95% confidence intervals (10,000 resamples) shown as translucent bands. Fitted parameters show significant differences for all three experiments.

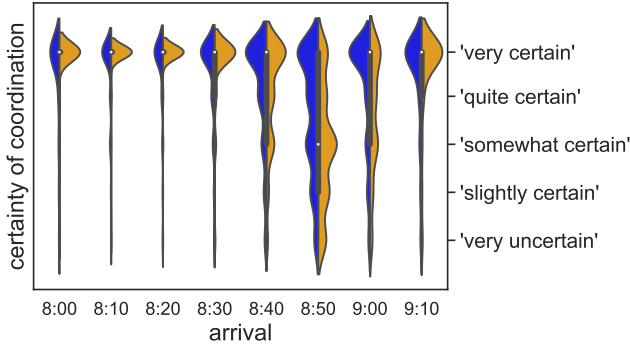


FIG. 2: Violin plots of certainty estimates. In each round, participants were asked how certain they were of successful coordination with their colleague. Blue areas show the results from MTurk ($n = 4260$) and orange areas show the results from DTU1 and DTU2 combined ($n = 3172$). We predefined a five point likert scale of certainty estimates as: 'very uncertain', 'slightly certain', 'somewhat certain', 'quite certain', and 'very certain', and translated them into the numerical values of probability estimates used in the payoff calculations (see Materials and Methods).

choices. This indicates that participants arriving at 8:40 or 8:50 do not feel incentivized to change their behavior significantly in subsequent rounds, even though there is a high risk of miscoordination. This is not to say that participants do not learn that canteen choices at 8:40 or 8:50 are dangerous. Partitioning the data from Figure 1 into two bins, corresponding to groups having had no miscoordination and groups having had one or more miscoordination (see the supplementary data analysis in the

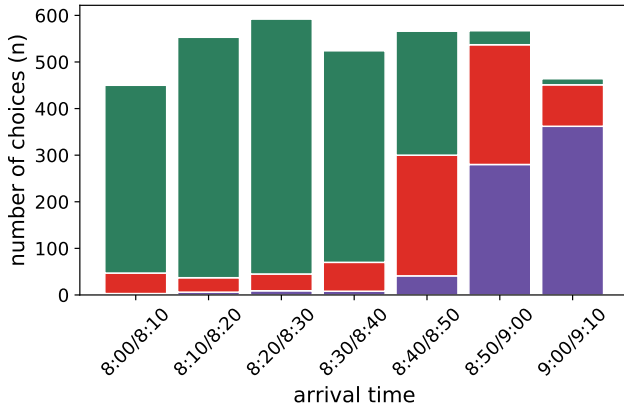


FIG. 3: Number of coordinations and miscoordination as a function of arrival times. Green means coordinating into the canteen, purple means coordinating into the office, and red means miscoordination. We use the notation 8:00/8:10 to denote the union of the arrival pairs (8:00, 8:10) and (8:10, 8:00), i.e., the arrival time combinations where one of the players arrive at 8:00 and the other at 8:10. Miscoordinations approach 50% at 8:40/8:50 and 8:50/9:00.

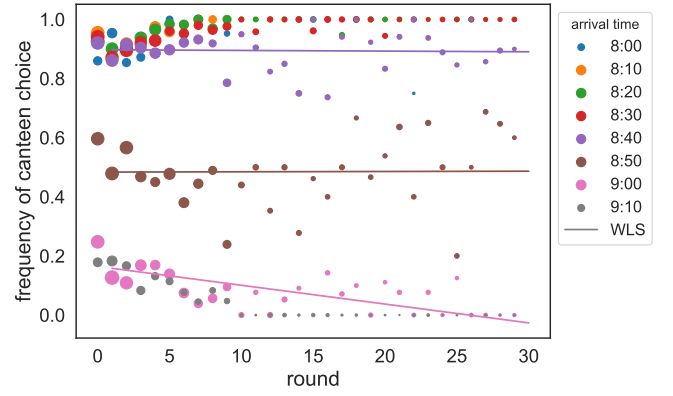


FIG. 4: Mean frequencies of canteen choices for all possible arrival times as a function of the number of rounds played. The fitted straight lines are weighted linear squares (WLS) with the weights chosen to be the square root of the number of data points constituting the mean frequencies for each round, also shown by dot size.

Supplementary Information), shows somewhat decreasing certainty estimates around the critical arrival times, especially for DTU student. However, this does not affect their actual choices. MTurk participants do choose the canteen a little less often after a miscoordination (see Figure S6 in the Supplementary Information), but this does not translate into better payoffs as later miscoordination just move to earlier arrival times. So even though participants learn that their choices are risky, they don't see any way to improve their strategy. Specifically, they never converge to the optimal all-office strategy, and also not to the alternative canteen-before-9 strategy (cf. Theorem 4 in Materials and Methods). This apparent lack of behavioral change in higher-order social reasoning games is also shown in Verbrugge & Mol [67].

Discussion

Let us try to analyse the experimental results in terms of the depth of knowledge of the participants. The highest payoff is achieved when coordinating into the canteen before 9 am. With the aim of achieving the highest possible payoff, each participant can be expected to consider her own arrival time and try to assess whether there is still time to meet in the canteen. When a participant arrives strictly before 9 am, i.e. at 8:50 or earlier, she has private knowledge that she arrives sufficiently early to go to the canteen. If participants only make choices based on their private knowledge, we should then expect participants to always go to the canteen at 8:50. This is not what we see, cf. Figure 1. Thus, other considerations in addition to the player's private knowledge must play a role in their decision-making.

When participants know that they both have arrived before 9:00, they have shared knowledge of having arrived in time for going to the canteen. This happens for any arrival pair (t_1, t_2) with $t_i \leq 8:50$, $i = 1, 2$ (where,

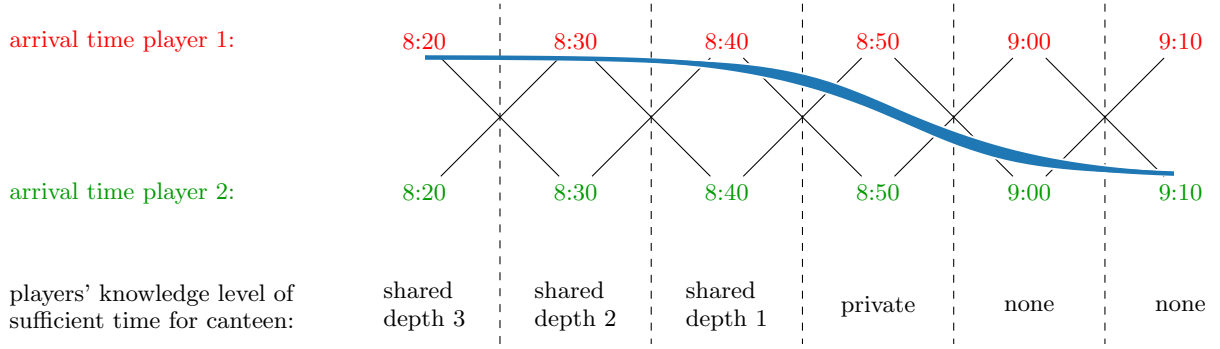


FIG. 5: The solid black lines express indistinguishability for the players, e.g. the arrival time 8:40 for player 1 has a line to both of the arrival times 8:30 and 8:50 for player 2, since these are the two arrival times for player 2 that player 1 will consider possible when herself arriving at 8:40. Below each possible arrival time, we have marked the highest level of knowledge concerning whether there is sufficient time to go to the canteen, e.g. when arriving at 8:40 there is shared knowledge to depth 1 of this fact, but not shared knowledge to depth 2. In blue, a binary logistic regression model was used to predict the probability of a participant going to the canteen (upper limit) or to the office (lower limit) at the shown arrival times. The width of the regression line indicates the 95% confidence interval using 10.000 bootstrapped resamples of all choices in all three experiments ($N = 7432$).

again, an arrival pair (t_1, t_2) denotes that player 1 arrives at time t_1 and player 2 at time t_2). Note that for an arrival pair $(8:50, 8:40)$, there is shared knowledge of there being sufficient time to go to the canteen, but only player 2 knows this fact: Player 2 knows that also player 1 must have arrived before 9:00, but player 1 doesn't know this about player 2. In other words, when a player arrives at 8:50, that player considers it possible that there is shared knowledge of being sufficient time for a cup of coffee in the canteen, but only if arriving at 8:40 or before will that player *know* there to be shared knowledge (to depth 1). When arriving at 8:30 or before, the player additionally knows there to be shared knowledge to depth 2. We illustrate this in Figure 5. Note that in general, if a player arrives at time $8:50 - 10n$, $n > 0$, then that player knows that there is n th-order shared knowledge, but the player doesn't know there to be $(n + 1)$ st-order shared knowledge. This follows a similar pattern as the mountain trekking example, except here the depth of shared knowledge is determined by how early ahead of 9 am the agents arrive, rather than how many messages have successfully been delivered. No number of messages was sufficient to achieve common knowledge in the mountain trekking example. We similarly get that no arrival time is sufficiently early to establish common knowledge about having time to meet in the canteen.

The participants seem to clearly be able to distinguish between private and shared knowledge, which is supported by their significantly different choices at 8:50 and 8:40 (see again Figure 5). However, it is less clear whether they are able to robustly distinguish different levels of shared knowledge, and whether they are able to distinguish that from common knowledge. Indeed, most participants relatively robustly choose the canteen at 8:40 and any time before that, despite the difference in depth

of shared knowledge in those possible arrival times. The certainty estimates are however slowly decreasing from 8:10 to 8:50 in all three experiments (see Figure 2), showing that the participants are not completely ignorant to the differences. This could suggest that many participants are aware that it is less safe to go to the canteen based on n th-order shared knowledge than $(n + 1)$ st-order shared knowledge. However, very few seem to draw the conclusion that it is never safe to go to the canteen. Our game theoretic analysis showed that they ought to only choose the canteen when there is common knowledge that it is safe, which in this case actually means never.

Why do participants not regard earlier office choices as viable options? Why do participants not continue their train of thought and deduce that when 8:50 turns out to be unsafe, 8:40 will become unsafe as well, which means that 8:30 will be unsafe also, etc.? One reason may be that the benefits of an all-office strategy are cognitively unavailable for the participants in the sense that participants have a limited ability to take the perspective of each other recursively. Another reason may be that the benefits of an all-office strategy are (vaguely) understood, but participants do not believe that their colleague will reason the same way as they do themselves, and instead try to guess what their colleague will choose. One candidate of such a (mixed) strategy may be the following: 1) always go to the canteen before 8:50, 2) always going to the office after 8:50, and 3) do some guesswork at 8:50. The arrival time combinations 8:40/8:50 and 8:50/9:00 will then coordinate 50% of the time, matching well with what we observe in Figure 12. If this is the strategy followed, players should also be aware of the 50% probability of miscoordination at 8:50. In Figure 2, we indeed see a much lower certainty estimate at those arrival times.

Probing these questions further, we asked participants the following post-game question:

“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)

The results in Figure 6A show that most answers range from 8:30 to 8:50 (approximately 75 % of all answers), giving support to the verdict that participants are not able to continue taking the perspective of each other recursively, or at least that they believe that shared knowledge of some modest finite depth is sufficient for the canteen choice to be safe. Rather, they stop after one or two, possibly three, iterations, thus believing that as long as they arrive sufficiently early, they can be sure to coordinate safely in the canteen. Notice that the correct answer “there is no such time” is chosen by less than 4% of all participants, close to the margin of random error.

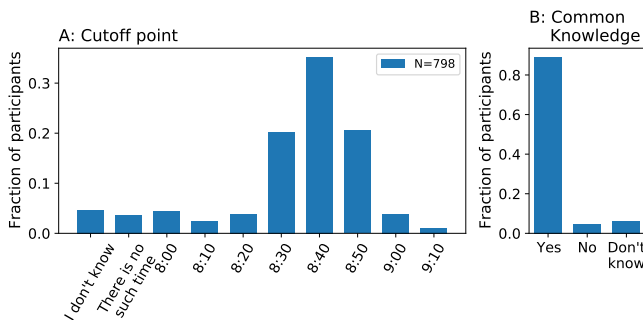


FIG. 6: A) Frequencies of answers to the question: “Imagine you could have agreed beforehand with your colleague about a point in time where it is safe to go to the canteen. What time would that be?” Due to the pragmatics of language, we assume that an answer like 8:30 entails the belief that all earlier arrival times would also be deemed safe. B) Frequencies of answers to the question: “Imagine you arrive at 8:10. Is it common knowledge between you and your colleague that it is safe to go to the canteen, that is, that you both arrived before 9:00?”

This indicates that participants indeed do believe that there exists a strategy that includes canteen choices without the risk of miscoordination, supporting the conjecture stated in the introduction that moderate depths of shared knowledge become effectively indistinguishable from common knowledge. Participants might of course not necessarily have a precise idea of the technical notion of common knowledge, but as discussed in the introduction, there is actually quite a number of studies demonstrating that humans have adapted to recognize common knowledge and making distinct strategic choices depending on whether there is common, shared or private knowledge. In our experiments, we see the player behavior stabilizing already at relatively modest depths

of shared knowledge, both in terms of action choices and certainty estimates. And that player behavior matches what we would expect to see if they indeed wrongly infer common knowledge from shared knowledge to some finite depth.

To specifically address the issue of whether they wrongly infer common knowledge, we asked a final post-game question:

“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’*)

This question inquires about participants’ understanding of the term ‘common knowledge’, and how it applies to the given situation. In Figure 6B, the results show that 89% of all participants responded that it was common knowledge that both players arrived before 9:00, when they themselves had arrived at 8:10. The answers may signify that indeed they believe there to be common knowledge in the strict technical (logical) sense. But of course the answers could also pertain to the everyday linguistic usage of the term ‘common knowledge’, which is less strict.

Conclusion

We have devised a new coordination game, the Canteen Dilemma, to investigate human higher-order social reasoning. Our experimental results show that high levels of recursive perspective-taking are cognitively unavailable to the vast majority of players of the game. We see a significant amount of miscoordination, which seems to occur due to a “curse of shared knowledge”: the guise of common knowledge existing in situations where there is only shared knowledge to some limited depth.

Our experience from playing the Canteen Dilemma with many people and explaining to them its unintuitive result, is that many players simply do not accept the argument that they cannot at any time coordinate safely into the canteen. On top of this, the certainty of participants that they will coordinate into the canteen at early arrival times indicates that when participants fail at n th-order reasoning, they do not default to agnosticism, but the opposite. That is, when there is a sufficiently large order of shared knowledge about a fact, it is possible that such a fact is mistaken for proper common knowledge. An interesting avenue for future research would be to investigate if there may be any social and psychological benefits of having an *illusion of common knowledge*, such as a higher willingness to cooperate. Thus, what we have called the ‘curse’ of shared knowledge in the Canteen Dilemma, may turn out to be a blessing in other settings.

An obvious question is how often this illusion of common knowledge occurs in real life. For instance, in the real-world version of the Canteen Dilemma scenario, the two colleagues would be likely to simply coordinate their

actions via cell phone (“I’ll arrive at 8:50 today. Are you up for a cup of coffee in the canteen?”). This suggests that the advent of modern technology could have made the information asymmetry inherent in shared knowledge situations less widespread. However, modern technology has also given us the Internet and social media, where the flow of information is much more complex, creating more intricate cases of information asymmetry than ever before.

That humans tend to confuse shared and common knowledge could possibly be due to a limited evolutionary importance of being able to make the distinction. It could also be due to the distinction requiring too many cognitive resources. Or it could be that the confusion actually leads to evolutionary benefits in terms of higher degrees of cooperation in most practically occurring settings. What exactly has lead to the confusion, and to what degree it has any practical importance today, we leave as open problems.

Methods

Experimental design and data collection

Experiments on Amazon Mechanical Turk had a total of 714 participants (including dropouts, see Supplementary Information), while the two classroom experiments at the Technical University of Denmark (DTU1 and DTU2) had a total of 106 and 50 participants, respectively. The average payout to MTurk workers was \$4.17 (including a general participation fee of \$2). After accepting our task and providing informed consent, participants from MTurk were put in a ‘waiting room’ until they were paired up with another participant. After an instructions page, detailing the rules of the game, participants were given an arrival time $t \in \{8:00, 8:10, 8:20, 8:30, 8:40, 8:50, 9:00, 9:10\}$ and asked to make a decision between going to the canteen or to the office. Next, participants were asked to estimate how certain they were that their ‘colleague’ made the same choice as them, ranging from ‘very uncertain’ over ‘slightly certain’, ‘somewhat certain’ and ‘quite certain’ to ‘very certain’, which were translated into numerical values, e_i , used in the payoff calculations (see below). A results page was shown between each round, showing the results of the previous rounds, including arrival times for both players, their choices, their own certainty estimates and resulting payoffs. After 30 seconds, the game would automatically proceed to the next round. After the last round, we asked all participants a few final questions about their strategy and their understanding of the game. The experiments were implemented using oTree 2.1.35 [68].

The two classroom experiments DTU1 and DTU2 differed from the MTurk experiment in a few aspects: 1) the maximum number of rounds played was increased from 10 to 30; 2) the initial bonus given each participant was increased accordingly from \$10 to \$30; 3) three additional questions were asked in order to elicit more explicitly some of the implicit assumptions and explicit behaviours by the students; 4) participants were told that they would not receive any monetary rewards, but that they should try to do their best. DTU1 received prizes. Screenshots, additional questions, experimental settings, and a detailed walk-through can be found in the Supplementary Information.

Payoffs and penalties

All MTurk players finishing the game were paid a participation fee of \$2. In addition, a bonus could be earned if players did well. Before the game started, the bonus was set to \$10 for all participants. After each round, the bonus was reduced by a personal penalty, depending on the two players’ choices. Penalties are calculated using a logarithmic scoring rule and by ordering them to be minimized by successful coordinations into the canteen. Penalties are maximized by any type of miscoordination or forbidden choice (i.e. going to the canteen at 9 am or later). Office coordinations are designed to have larger penalties than canteen coordinations, but smaller penalties than miscoordinations in order to make sure that coordination remains the main objective of the game. Penalties are defined as negative utility values in the following way. First, we define the chosen action a_i by player i , $i = 1, 2$, to take binary values encoding the canteen option ($a_i = 0$) and the office option ($a_i = 1$), and define their respective certainty estimates $e_i \in \{0.5, 0.625, 0.75, 0.875, 0.99\}$. We can then express the utility u received by player 1 as $u(e_1, a_1, a_2) = (1 - |a_1 - a_2| + a_1 a_2) \ln(e_1) + 2|a_1 - a_2| \ln(1 - e_1)$, and symmetric for player 2. If any of the players choose the canteen at 9 am or after, the utility becomes $u(e_1, a_1, a_2) = 2 \ln(1 - e_1)$ for player 1 (and symmetric for player 2), corresponding to a miscoordination. As an example, imagine player 1 arrives at 8:40 and chooses the canteen, $a_1 = 0$. She estimates the probability that her colleague also will go to the canteen to “somewhat certain”, $e_1 = 0.75$. If her colleague indeed chooses the canteen, $a_2 = 0$, her utility will be $u(e_1, a_1, a_2) = \ln(e_1) = -0.29$, but if her prediction proves false and her colleague chooses the office instead, her utility will be $u(e_1, a_1, a_2) = 2 \ln(1 - e_1) = -2.77$. If she goes to the office just like her colleague, her utility is $u(e_1, a_1, a_2) = 2 \ln(e_1) = -0.58$. It should be noted that the logarithmic scoring rule used here is not strictly proper since office and canteen coordinations are penalized differently. Nevertheless, we find a good match between estimates and actual choices at arrival times different from those that are prone to miscoordinations, as seen in Figure 2, indicating that loss minimization remained a central concern and that participants made their choices and estimates as honestly as possible [69, 70].

Formal Analysis

The game can be represented as a game with three players, *nature*, *player 1* and *player 2*. Nature is the player that initially decides the arrival times of player 1 and 2. Then player 1 and 2 are each informed of their own arrival time, and each have to choose among two actions: o for going to the office and c for going to the canteen. Based on the choice of actions by all three agents, player 1 and 2 receive a payoff, and they always receive the same payoff (we are disregarding the certainty estimates for now). The action choice of nature can be represented as an *arrival pair* $\mathbf{t} = (t_1, t_2)$ consisting of the arrival time t_1 for player 1 and t_2 for player 2. Any arrival pair \mathbf{t} has to satisfy that $|t_1 - t_2| = 10$ minutes (we will suppress mentioning the unit, minutes, in the following). In our specific version of the game, we additionally have the restriction that $8:10 \leq t_i \leq 9:10$ for $i = 1, 2$. The analysis of optimal strategies however doesn’t depend on the exact arrival times available, so we will make things a bit more general and only assume that there is an earliest arrival time t_{\min} and a latest arrival time t_{\max} , and that $t_{\min} \leq 8:50$ and $t_{\max} \geq 9:00$.

Given t_{\min} and t_{\max} , the set of possible arrival pairs is defined as $T = \{(t, t+10) \mid t_{\min} \leq t \leq t_{\max} - 10\} \cup \{(t, t-10) \mid t_{\min} + 10 \leq t \leq t_{\max}\}$.

The game starts by nature choosing an element $\mathbf{t} \in T$. Nature is not a strategic player, so we assume that \mathbf{t} is chosen uniformly at random, which is exactly how \mathbf{t} is chosen in our experiments. The participants do not know that the arrival times are chosen uniformly at random, as this is left implicit in the description of the game. The following analysis of optimal strategies in the game could potentially change if arrival times were chosen according to a highly skewed probability distribution.

When nature has chosen its action $\mathbf{t} \in T$ and player 1 and 2 have chosen their actions a_1 and a_2 , player 1 and 2 receive their payoff, which we denote $u_{\mathbf{t}}(a_1, a_2)$ (the utility resulting from player 1 choosing a_1 and player 2 choosing a_2 given that nature played \mathbf{t}). We don't need to make any assumptions regarding the exact utility values (payoff values), except that successful coordination into the canteen is always better than successful coordination into the offices, which again is always better than being miscoordinated. Hence, we put the following constraints on the utility function, for all $\mathbf{t} \in T$,

$$(U1) \text{ If } t_1, t_2 < 9:00 \text{ then } u_{\mathbf{t}}(c, c) > u_{\mathbf{t}}(o, o) > u_{\mathbf{t}}(c, o) = u_{\mathbf{t}}(o, c).$$

$$(U2) \text{ If } t_i \geq 9:00 \text{ for some } i, \text{ then } u_{\mathbf{t}}(o, o) > u_{\mathbf{t}}(c, o) = u_{\mathbf{t}}(o, c) = u_{\mathbf{t}}(c, c).$$

A *strategy* for player i , $i = 1, 2$, is a mapping from arrival pairs to actions, that is, a mapping $s_i : T \rightarrow \{c, o\}$. A strategy for i simply determines which action i will choose given the arrival pair. Each agent only observes her own arrival time, that is, any two arrival pairs \mathbf{t} and \mathbf{t}' with $t_i = t'_i$ will be indistinguishable to player i , $i = 1, 2$. This immediately leads to the following formal definition of the indistinguishability relation \sim_i for player i : $\mathbf{t} \sim_i \mathbf{t}'$ iff $t_i = t'_i$. We need to require the strategy of each player to be *uniform*, that is, any two arrival pairs that are indistinguishable by that player should be mapped to the same action: if $\mathbf{t} \sim_i \mathbf{t}'$ then $s_i(\mathbf{t}) = s_i(\mathbf{t}')$. Due to the uniformity condition, we can allow ourselves to overload the meaning of the symbol s_i and write $s_i(t_i)$ as an abbreviation of $s_i(t_1, t_2)$ for $i = 1, 2$.

Note that the defined strategies are memoryless (Markov strategies), that is, a player's choice only depends on the observed arrival time in the current round, not the history of arrival times and chosen actions in earlier rounds. Human players playing the game should not be expected to necessarily play memoryless strategies, as they might seek to adopt to the observed strategy of the other player. However, since it is a repeated game (every round is a new instance of the same game), perfectly rational players playing the game for a sufficient number of rounds should converge to an optimal memoryless strategy. We will leave further discussion of history-dependent strategies and focus on the optimal memoryless strategies in the following.

Given strategies s_1 and s_2 , the pair $\mathbf{s} = (s_1, s_2)$ is called a *strategy profile*. Given an arrival pair $\mathbf{t} = (t_1, t_2)$, we use $\mathbf{s}(\mathbf{t})$ as a shorthand for $(s_1(t_1), s_2(t_2))$. Hence $\mathbf{s}(\mathbf{t})$ denotes the choices made by players 1 and 2 when their strategies are given by \mathbf{s} and their arrival times are given by \mathbf{t} . The payoff of those choices is then $u_{\mathbf{t}}(\mathbf{s}(\mathbf{t}))$. We will use $u_{\mathbf{t}}(\mathbf{s})$ as an abbreviation of $u_{\mathbf{t}}(\mathbf{s}(\mathbf{t}))$, i.e., $u_{\mathbf{t}}(\mathbf{s})$ is the utility received by player 1 and 2 when they play by strategy profile \mathbf{s} in the game with arrival pair \mathbf{t} . The *expected utility* $EU(\mathbf{s})$ of a

strategy profile \mathbf{s} is the average of the payoffs [71]:

$$EU(\mathbf{s}) = \frac{1}{|T|} \sum_{\mathbf{t} \in T} u_{\mathbf{t}}(\mathbf{s}).$$

Note again that player 1 and 2 get the same payoff (common-payoff game), so there is only one expected utility value to be computed. A strategy profile \mathbf{s}' *Pareto dominates* another strategy profile \mathbf{s} if $EU(\mathbf{s}') > EU(\mathbf{s})$ [71]. A strategy profile is *Pareto optimal* if there does not exist another strategy profile dominating it. A strategy profile \mathbf{s}' *weakly Pareto dominates* another strategy profile \mathbf{s} if $EU(\mathbf{s}') \geq EU(\mathbf{s})$. The game is cooperative (between player 1 and 2), so both players should seek to play a Pareto optimal strategy profile. Also, since it is a common-payoff game, all Pareto optimal strategy profiles have the same expected utility [71].

Define subsets of arrival pairs $T_1, T_2 \subseteq T$ by:

$$T_1 = \{(8:10 + 20x, 8:20 + 20y) \in T \mid y \leq x \leq y + 1\}$$

$$T_2 = \{(8:20 + 20y, 8:10 + 20x) \in T \mid y \leq x \leq y + 1\}$$

Note that T_1 and T_2 are disjoint and that $T = T_1 \cup T_2$. We hence get, for any strategy profile \mathbf{s} ,

$$EU(\mathbf{s}) = \frac{1}{|T|} \sum_{\mathbf{t} \in T} u_{\mathbf{t}}(\mathbf{s}) = \frac{1}{|T|} \sum_{i \in \{1, 2\}} \sum_{\mathbf{t} \in T_i} u_{\mathbf{t}}(\mathbf{s}) \quad (1)$$

Given a strategy profile \mathbf{s} , we let $\mathbf{s} \upharpoonright T_i$ be the restriction of \mathbf{s} to T_i , that is, $\mathbf{s} \upharpoonright T_i$ is as \mathbf{s} except it is only defined on the arrival pairs in T_i . So $\mathbf{s} \upharpoonright T_i$ is the strategy profile for the subgame in which only the arrival pairs in T_i can be chosen. We can now rewrite formula (1) as

$$EU(\mathbf{s}) = \frac{1}{|T|} \sum_{\mathbf{t} \in T} u_{\mathbf{t}}(\mathbf{s}) = \frac{1}{|T|} \sum_{i \in \{1, 2\}} \sum_{\mathbf{t} \in T_i} u_{\mathbf{t}}(\mathbf{s} \upharpoonright T_i) \quad (2)$$

Note that there exists no $\mathbf{t} \in T_1$, $\mathbf{t}' \in T_2$ and $i \in \{1, 2\}$ such that $\mathbf{t} \sim_i \mathbf{t}'$. Hence the strategy profile $\mathbf{s} \upharpoonright T_1$ can be chosen completely independently of the strategy profile $\mathbf{s} \upharpoonright T_2$. Using formula (2) it then follows that a strategy profile \mathbf{s} is Pareto optimal in the full game if and only if each of the strategy profiles $\mathbf{s} \upharpoonright T_1$ and $\mathbf{s} \upharpoonright T_2$ are Pareto optimal on the subgames with arrival pairs only in T_1 and T_2 , respectively. When looking for Pareto optimal strategy profiles in the game, we can hence look for Pareto optimal strategy profiles on each of the two subgames independently (essentially the game consists of two disjoint subgames). Note also that the two games are completely symmetric, since $(t_1, t_2) \in T_1$ if and only if $(t_2, t_1) \in T_2$. Hence the two subgames necessarily have the exactly the same strategy profiles up to symmetry (swapping the roles of player 1 and 2). It is hence sufficient to only investigate Pareto optimal strategies for one of these subgames, say the subgame with arrival pairs in T_1 .

We will now try to determine the possible candidates for being Pareto optimal strategy profiles for the subgame with arrival pairs in T_1 . We do this by iteratively removing strategy profiles that are not Pareto optimal.

Lemma 1. *Going to the canteen at 9:00 or after can never be part of a Pareto optimal strategy profile. More precisely: No strategy profile \mathbf{s} with $s_i(t_i) = c$ for some $i \in \{1, 2\}$ and $t_i \geq 9:00$ can be Pareto optimal.*

Proof. Consider a strategy profile \mathbf{s} with $s_i(t_i) = c$ for some $i \in \{1, 2\}$ and some $\mathbf{t} \in T_1$ with $t_i \geq 9:00$. We only consider

the case of $i = 1$, the case of $i = 2$ being proved similarly. Then we have $s_1(t_1) = c$ and $t_1 = 9:10 + 20x$ for some x (recall that we have restricted attention to the arrival pairs in T_1). Now define a strategy profile \mathbf{s}' which is identical to \mathbf{s} except $s'_j(t'_j) = o$ for all $j \in \{1, 2\}$ and $t'_j \geq 9:00$. We want to show that \mathbf{s}' Pareto dominates \mathbf{s} . First note that:

$$\begin{aligned} u_{(9:10+20x, 9:00+20x)}(\mathbf{s}') &\neq u_{(9:10+20x, 9:00+20x)}(o, o) \text{ def. of } \mathbf{s}' \\ &> u_{(9:10+20x, 9:00+20x)}(c, \cdot) \text{ by U2} \\ &= u_{(9:10+20x, 9:00+20x)}(\mathbf{s}), \end{aligned}$$

where the last equality follows from $s_1(9:10+20x) = s_1(t_1) = c$. This proves the existence of an arrival pair for which \mathbf{s}' has a strictly higher utility than \mathbf{s} . To prove $EU(\mathbf{s}') > EU(\mathbf{s})$, we hence only need to prove that $u_{\mathbf{t}'}(\mathbf{s}') \geq u_{\mathbf{t}'}(\mathbf{s})$ for all $\mathbf{t}' \in T_1$. When $t'_1, t'_2 < 9:00$ this is trivial, as we then have $\mathbf{s}'(\mathbf{t}') = \mathbf{s}(\mathbf{t}')$ by definition of \mathbf{s}' . When $t'_j \geq 9:00$ for some j , $\mathbf{s}'(\mathbf{t}')$ will by definition of \mathbf{s}' necessarily have an o in each position where $\mathbf{s}(\mathbf{t}')$ also has one. It follows by constraint U2 that $u_{\mathbf{t}'}(\mathbf{s}') \geq u_{\mathbf{t}'}(\mathbf{s})$, as required. \square

Lemma 2. Assume $s_i(t) = o$ and $s_{3-i}(t+10) = c$ for some strategy profile \mathbf{s} , some arrival time t , and some $i \in \{1, 2\}$. Then \mathbf{s} is not Pareto optimal.

Proof. Let \mathbf{s} , t and i be as stated above. We need to find a strategy profile \mathbf{s}' Pareto dominating \mathbf{s} . We only consider the case $i = 1$, the case of $i = 2$ being symmetric. Then we have $s_1(t) = o$ and $s_2(t+10) = c$. If $s_j(t'_j) = c$ for some $j \in \{1, 2\}$ and $t'_j \geq 9:00$, the existence of a strategy dominating \mathbf{s} follows immediately from Lemma 1. We can hence in the following assume that $s_j(t'_j) = o$ for all $j \in \{1, 2\}$ and $t'_j \geq 9:00$. Since $s_2(t+10) = c$, we can thus also conclude that $t+10 < 9:00$.

Now define \mathbf{s}' to be identical to \mathbf{s} except that we let

$$\begin{aligned} s'_1(t-20x) &= c & \text{for } x \geq 0 \\ s'_2(t-10-20y) &= c & \text{for } y \geq 0, \end{aligned}$$

recalling that we are only considering arrival pairs in T_1 . We want to show that $EU(\mathbf{s}') > EU(\mathbf{s})$. First note that

$$\begin{aligned} u_{(t, t+10)}(\mathbf{s}') &\neq u_{(t, t+10)}(c, s_2(t+10)) \text{ by def. of } \mathbf{s}' \\ &= u_{(t, t+10)}(c, c) & \text{by def. of } \mathbf{s} \\ &> u_{(t, t+10)}(o, c) & \text{by U1, as } t+10 < 9:00 \\ &= u_{(t, t+10)}(\mathbf{s}) & \text{by def. of } \mathbf{s} \end{aligned}$$

To prove $EU(\mathbf{s}') > EU(\mathbf{s})$, we hence only need to prove that $u_{\mathbf{t}'}(\mathbf{s}') \geq u_{\mathbf{t}'}(\mathbf{s})$ for all $\mathbf{t}' \in T$. The only non-trivial cases are when either $t'_1 = t - 20x$ for some $x \geq 0$ or $t'_2 = t - 10 - 20y$ for some $y \geq 0$ (in all other cases, $\mathbf{s}'(\mathbf{t}') = \mathbf{s}(\mathbf{t}')$). Consider first a \mathbf{t}' with $t'_2 = t - 10 - 20y$ for some $y \geq 0$. Then $t'_1 = t - 20x$ for some $x \geq 0$, and hence $u_{\mathbf{t}'}(\mathbf{s}') = u_{\mathbf{t}'}(c, c)$. Constraint U1 now immediately gives $u_{\mathbf{t}'}(c, c) \geq u_{\mathbf{t}'}(\mathbf{s})$, and hence $u_{\mathbf{t}'}(\mathbf{s}') \geq u_{\mathbf{t}'}(\mathbf{s})$, as required. Consider instead \mathbf{t}' with $t'_1 = t - 20x$ for some $x \geq 0$. Then either $t'_2 = t - 10 - 20y$ for some $y \geq 0$ or we have $t'_1 = t$ and $t'_2 = t + 10$. Both cases have already previously been covered. This completes the proof. \square

Definition 3. A cut-off strategy with cut-off t' is a strategy s with $s(t) = c$ for all $t < t'$ and $s(t) = o$ for all $t > t'$. A cut-off strategy profile with cut-off t' is a pair (s_1, s_2) where both s_1 and s_2 are cut-off strategies with cut-off t' . A cut-off strategy (profile) with cut-off before 8:10 is called an all-office strategy (profile).

Note that the strategy we in the informal discussions above referred to as the “canteen-before-9” strategy is the cut-off strategy with cut-off 8:55. We now get the result on optimal strategies claimed in the informal discussion.

Theorem 4. Any Pareto optimal strategy profile is either the all-office strategy profile or the cut-off strategy profile with cut-off 8:55.

Proof. Let \mathbf{s} be a Pareto optimal strategy profile. Assume that t_{\min} is of the form $8:50 - 20x$ for some $x \geq 0$ and t_{\max} is of the form $9:00 + 20y$ for some $y \geq 0$, the other cases being treated symmetrically. Let σ be the following string over the alphabet $\{o, c\}$, where we alternate between the strategy choices of player 1 and 2 from t_{\min} to t_{\max} :

$$s_1(t_{\min})s_2(t_{\min} + 10)s_1(t_{\min} + 20)s_2(t_{\min} + 30) \cdots s_2(t_{\max})$$

Note that when $(t_1, t_2) \in T_1$, then $s_1(t_1)$ and $s_2(t_2)$ both occur in the string σ . From Lemma 2 it follows that σ cannot contain the substring oc . Suppose the first letter of σ is o . Then since σ does not contain the substring oc , we have $\sigma = o^{|\sigma|}$ (a string of only o s). Hence $s_1(\mathbf{t}) = s_2(\mathbf{t}) = o$ for all $\mathbf{t} \in T_1$. This means that \mathbf{s} is the all-office strategy profile. Suppose alternatively that the first letter of σ is c . The last letter of σ is necessarily o since $t_{\max} \geq 9:00$ and \mathbf{s} is Pareto optimal, cf. Lemma 1. Since σ is then a string that starts with c and ends with o , but doesn't contain oc , it must have the form $c^n o^m$ for some $n, m \geq 1$ with $m + n = |\sigma|$. Hence there exists a t' such that $s_i(t) = c$ for all $t < t'$ and $s_i(t) = o$ for all $t > t'$ (and all $i \in \{1, 2\}$). In other words, \mathbf{s} is a cut-off strategy profile with cut-off t' . What is left to prove is then only that \mathbf{s} has cut-off 8:55. First note that we must necessarily have $t' < 9:00$, since otherwise \mathbf{s} would not be Pareto optimal according to Lemma 1. From this it follows that we for all arrival pairs $(t_1, t_2) \in T_1$ must have

1. If $t_1, t_2 < t'$, the two players coordinate into the canteen, receiving the highest possible payoff.
2. If $t_1 < t' < t_2$ or $t_2 < t' < t_1$, the two players are miscoordinated (one chooses canteen, the other office), receiving the lowest possible payoff.
3. If $t_1, t_2 > t'$, the two players coordinate into their offices, receiving a payoff strictly between the lowest and highest.

Note that there will always be exactly one arrival pair in T_1 of type 2, independent of t' . Since arrival pairs of type 1 have a higher payoff than arrival pairs of type 3, and \mathbf{s} is Pareto optimal, \mathbf{s} must have the maximal number of arrival pairs of type 1, that is, it is the cut-off strategy with the latest possible cut-off. That is exactly the cut-off 8:55 (or, more precisely, any cut-off strictly between 8:50 and 9:00). \square

The theorem proves what was argued in the main text: There are only two candidates for an optimal strategy, the all-office strategy or the canteen-before-9 strategy. This does not in any way imply that we should expect human players to adopt any of these two strategies, but if two perfectly rational players were to play the game, and if they knew they could expect the other player to play perfectly rational as well, of course the optimal strategy would be played. And, as earlier mentioned, in our particular version of the game, the optimal strategy is the all-office strategy.

Logistic regression

The experimental results were analyzed using a logistic regression model with the arrival time t as predictor. The model was specified as $\mu_i = \alpha_i + \beta_i t$, with μ_i being the log-odds $\mu_i = \log(p_i/(1 - p_i))$ and $i = 1, 2, 3$ denoting the MTurk, DTU1 and DTU2 experiments, respectively, allowing for differing slopes and intercepts. Results are shown in Figure 1. The steepest slope occurs at $p(t) = 1/2 = -\alpha/\beta$, which for the MTurk and DTU1 experiments is $t = 8:48$ and for DTU2 is $t = 8:52$. The regression line in Figure 5 is obtained similarly by combining observations from all three experiments. The high number of observations imply small confidence bands. Hence, conclusions from the models can be viewed as robust.

Acknowledgments

Server infrastructure and devops was handled by Mikkel Birkegaard Andersen. The authors wish to thank Vincent

F. Hendricks for enabling the project. This research was approved by the Institutional Review Board at the University of Copenhagen and included informed consent by all participants in the study. The authors gratefully acknowledge the support provided by The Carlsberg Foundation under grant number CF 15-0212.

Author contributions

Thomas Bolander developed the game, made the theoretical analysis, and contributed as lead author; Robin Engelhardt designed the experiment, analyzed the data, and contributed as lead author; Thomas S. Nicolet designed the experiment, analyzed the data, and contributed as lead author. The authors declare no conflict of interest.

-
- [1] R. Fagin, J. Y. Halpern, Y. Moses, M. Y. Vardi, Reasoning About Knowledge, MIT Press, 1995.
 - [2] D. Lewis, Convention: A Philosophical Study, 1969.
 - [3] H. H. Clark, C. R. Marshall, Definite knowledge and mutual knowledge (1981).
 - [4] T. C. Schelling, The strategy of conflict, Harvard university press, 1980.
 - [5] R. J. Aumann, Agreeing to disagree, The annals of statistics (1976) 1236–1239.
 - [6] I. van de Pol, I. van Rooij, J. Szymanik, Parameterized complexity of theory of mind reasoning in dynamic epistemic logic, Journal of Logic, Language and Information 27 (2018) 255–294.
 - [7] P. Milgrom, An axiomatic characterization of common knowledge, Econometrica: Journal of the Econometric Society (1981) 219–222.
 - [8] H. H. Clark, Using language, Cambridge university press, 1996.
 - [9] J. W. Bradbury, S. L. Vehrencamp, et al., Principles of animal communication (1998).
 - [10] M. Tomasello, Joint attention as social cognition, Joint attention: Its origins and role in development (1995) 103–130.
 - [11] E. Lorini, L. Tumminoli, A. Herzig, Establishing mutual beliefs by joint attention: towards a formal model of public events, in: Proc. of CogSci, pp. 1325–1330.
 - [12] T. Bolander, H. van Ditmarsch, A. Herzig, E. Lorini, P. Pardo, F. Schwarzenruber, Announcements to attentive agents, Journal of Logic, Language and Information (2015) 1–35.
 - [13] H. Gintis, Rationality and common knowledge, Rationality and Society 22 (2010) 259–282.
 - [14] M. F. Friedell, On the structure of shared awareness, Behavioral Science 14 (1969) 28–39.
 - [15] T. C. Schelling, Bargaining, communication, and limited war, Conflict Resolution 1 (1957) 19–36.
 - [16] J. Tooby, L. Cosmides, Groups in mind: The coalitional roots of war and morality, Human morality and sociality: Evolutionary and comparative perspectives (2010) 91–234.
 - [17] Y. N. Harari, Sapiens: A brief history of humankind, Random House, 2014.
 - [18] D. Premack, G. Woodruff, Does the chimpanzee have a theory of mind?, Behavioral and brain sciences 1 (1978) 515–526.
 - [19] U. Frith, C. D. Frith, Development and neurophysiology of mentalizing, Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences 358 (2003) 459–473.
 - [20] K. Vogeley, P. Bussfeld, A. Newen, S. Herrmann, F. Happé, P. Falkai, W. Maier, N. J. Shah, G. R. Fink, K. Zilles, Mind reading: neural mechanisms of theory of mind and self-perspective, Neuroimage 14 (2001) 170–181.
 - [21] I. Apperly, Mindreaders: the cognitive basis of "theory of mind", Psychology Press, 2010.
 - [22] P. N. Johnson-Laird, Mental models: Towards a cognitive science of language, inference, and consciousness, 6, Harvard University Press, 1983.
 - [23] K. Gray, A. C. Jenkins, A. S. Heberlein, D. M. Wegner, Distortions of mind perception in psychopathology, Proceedings of the National Academy of Sciences 108 (2011) 477–479.
 - [24] S. Baron-Cohen, H. A. Ring, S. Wheelwright, E. T. Bullmore, M. J. Brammer, A. Simmons, S. C. Williams, Social intelligence in the normal and autistic brain: an fmri study, European journal of neuroscience 11 (1999) 1891–1898.
 - [25] S. M. Schaafsma, D. W. Pfaff, R. P. Spunt, R. Adolphs, Deconstructing and reconstructing theory of mind, Trends in cognitive sciences 19 (2015) 65–72.
 - [26] I. A. Apperly, S. A. Butterfill, Do humans have two systems to track beliefs and belief-like states?, Psychological review 116 (2009) 953.
 - [27] R. Saxe, L. Young, Theory of mind: How brains think about thoughts, The Oxford handbook of cognitive neuroscience 2 (2013) 204–213.
 - [28] A. Gopnik, J. W. Astington, Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction, Child development (1988) 26–37.
 - [29] D. O. Stahl, P. W. Wilson, On players' models of other players: Theory and experimental evidence, Games and Economic Behavior 10 (1995) 218–254.
 - [30] R. Nagel, Unraveling in guessing games: An experimental study, The American Economic Review 85 (1995)

- 1313–1326.
- [31] T. Hedden, J. Zhang, What do you think i think you think?: Strategic reasoning in matrix games, *Cognition* 85 (2002) 1–36.
 - [32] B. Keysar, S. Lin, D. J. Barr, Limits on theory of mind use in adults, *Cognition* 89 (2003) 25–41.
 - [33] S. Pinker, *The language instinct: How the mind creates language*, Penguin UK, 2003.
 - [34] A. M. Leslie, L. Thaiss, Domain specificity in conceptual development: Neuropsychological evidence from autism, *Cognition* 43 (1992) 225–251.
 - [35] R. Saxe, Theory of mind (neural basis), *Encyclopedia of consciousness* 2 (2009) 401–410.
 - [36] C. Heyes, Submentalizing: I am not really reading your mind, *Perspectives on Psychological Science* 9 (2014) 131–143.
 - [37] J. Perner, H. Wimmer, “john thinks that mary thinks that...” attribution of second-order beliefs by 5-to 10-year-old children, *Journal of experimental child psychology* 39 (1985) 437–471.
 - [38] K. Sullivan, D. Zaitchik, H. Tager-Flusberg, Preschoolers can attribute second-order beliefs., *Developmental Psychology* 30 (1994) 395.
 - [39] P. Kinderman, R. Dunbar, R. P. Bentall, Theory-of-mind deficits and causal attributions, *British Journal of Psychology* 89 (1998) 191–204.
 - [40] C. Camerer, G. Loewenstein, M. Weber, The curse of knowledge in economic settings: An experimental analysis, *Journal of political Economy* 97 (1989) 1232–1254.
 - [1] S. A. Birch, P. Bloom, The curse of knowledge in reasoning about false beliefs, *Psychological Science* 18 (2007) 382–386.
 - [42] Academian, Unrolling social metacognition: Three levels of meta are not enough., *Lesswrong*, 2018.
 - [6] J. De Freitas, K. 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.
 - [44] A. S. Goodie, P. Doshi, D. L. Young, Levels of theory-of-mind reasoning in competitive games, *Journal of Behavioral Decision Making* 25 (2012) 95–108.
 - [45] S. Pinker, *The stuff of thought: Language as a window into human nature*, Penguin, 2007.
 - [46] S. Pinker, M. A. Nowak, J. J. Lee, The logic of indirect speech, *Proceedings of the National Academy of sciences* 105 (2008) 833–838.
 - [47] H. de Weerd, R. Verbrugge, B. Verheij, Negotiating with other minds: the role of recursive theory of mind in negotiation with incomplete information, *Autonomous Agents and Multi-Agent Systems* 31 (2017) 250–287.
 - [48] R. Verbrugge, B. Meijering, S. Wierda, H. van Rijn, N. Taatgen, et al., Stepwise training supports strategic second-order theory of mind in turn-taking games, *Judgment and decision making* 13 (2018) 79–98.
 - [49] O. Curry, M. J. Chesters, ‘putting ourselves in the other fellow’s shoes’: The role of ‘theory of mind’ in solving coordination problems, *Journal of Cognition and Culture* 12 (2012) 147–159.
 - [50] M. Devaine, G. Hollard, J. Daunizeau, Theory of mind: did evolution fool us?, *PLoS One* 9 (2014) e87619.
 - [51] H. De Weerd, R. Verbrugge, B. Verheij, Higher-order theory of mind in the tacit communication game, *Biologically Inspired Cognitive Architectures* 11 (2015) 10–21.
 - [52] B. Erb, *Artificial intelligence & theory of mind* (2016).
 - [53] T. Bolander, Seeing is believing: Formalising false-belief tasks in dynamic epistemic logic, in: *Jaakko Hintikka on Knowledge and Game-Theoretical Semantics*, Springer, 2018, pp. 207–236.
 - [54] N. Bard, J. N. Foerster, S. Chandar, N. Burch, M. Lancot, H. F. Song, E. Parisotto, V. Dumoulin, S. Moitra, E. Hughes, et al., The hanabi challenge: A new frontier for ai research, *Artificial Intelligence* 280 (2020) 103216.
 - [55] L. Dissing, T. Bolander, Implementing theory of mind on a robot using dynamic epistemic logic, in: *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*.
 - [56] K. A. Thomas, P. DeScioli, O. S. Haque, S. Pinker, The psychology of coordination and common knowledge., *Journal of personality and social psychology* 107 (2014) 657.
 - [57] K. A. Thomas, J. De Freitas, P. DeScioli, S. Pinker, Recursive mentalizing and common knowledge in the bystander effect., *Journal of Experimental Psychology: General* 145 (2016) 621.
 - [58] J. J. Lee, S. Pinker, Rationales for indirect speech: The theory of the strategic speaker., *Psychological review* 117 (2010) 785.
 - [59] K. A. Thomas, P. DeScioli, S. Pinker, Common knowledge, coordination, and the logic of self-conscious emotions, *Evol Hum Behav* 39 (2018) 179–190.
 - [60] J. De Freitas, P. DeScioli, K. A. Thomas, S. Pinker, Maimonides’ ladder: States of mutual knowledge and the perception of charitability., *Journal of Experimental Psychology: General* 148 (2019) 158.
 - [61] G. S. Becker, *The economic approach to human behavior*, University of Chicago press, 1976.
 - [62] A. Rubinstein, The electronic mail game: Strategic behavior under “almost common knowledge”, *The American Economic Review* (1989) 385–391.
 - [63] P. van Emde Boas, J. Groenendijk, M. Stokhof, The conway paradox: Its solution in an epistemic framework, in: *Proceedings of the third Amsterdam Montague Symposium*, pp. 159–182.
 - [64] H. van Ditmarsch, B. Kooi, One hundred prisoners and a light bulb, in: *One Hundred Prisoners and a Light Bulb*, Springer, 2015, pp. 83–94.
 - [65] B. Meijering, L. van Maanen, H. van Rijn, R. Verbrugge, The facilitative effect of context on second-order social reasoning, in: *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 32.
 - [66] P. C. Wason, D. Shapiro, Natural and contrived experience in a reasoning problem, *Quarterly Journal of Experimental Psychology* 23 (1971) 63–71.
 - [67] R. Verbrugge, L. Mol, Learning to apply theory of mind, *Journal of Logic, Language and Information* 17 (2008) 489–511.
 - [68] D. Chen, M. Schonger, C. Wickens, otree—an open-source platform for laboratory, online, and field experiments, *Journal of Behavioral and Experimental Finance* 9 (2016) 88–97.
 - [69] T. Seidenfeld, Calibration, coherence, and scoring rules, *Philosophy of Science* 52 (1985) 274–294.
 - [70] T. R. Palfrey, S. W. Wang, On eliciting beliefs in strategic games, *Journal of Economic Behavior & Organization* 71 (2009) 98–109.
 - [71] Y. Shoham, K. Leyton-Brown, *Multiagent systems: Algorithmic, game-theoretic, and logical foundations*, Cambridge University Press, 2008.

Supplementary Information

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. 7.

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 (Fig. 8).

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. 9.

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 (Fig. 10). 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. 11. 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’*)

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 II, 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.

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 II: 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.

Participants quitting a study before completing it is prevalent on MTurk, and varies systemically 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 [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 Ep to be true (that is, requires K_iEp to be true for some $i \in G$). The point of De Freitas et al. is that if only Ep 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 Figure 12 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 Figure 13, we see a similar pattern as before: Much steeper profiles among DTU students, but also much lower certainty estimates around critical arrival times.

Datasets

MTurk_anonymous.xlsx, DTU1_anonymous.xlsx, and DTU2_anonymous.xlsx: Anonymized data set of all Mechanical Turk experiments. Parameters: session = name of experiment; code = anonymized participant id; group = group number in session; id_in_group = player id in group; round = round number; arrival = arrival time; choice = choice made by participant; certainty = certainty estimate by participant; bonus = penalty in dollars; strategy = free text question after game has ended; simple = answers to question 4, cutoff = answers to question 3; fault = answers to question 1; payoff = money left after game has finished.

-
- [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.

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. 7: 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. 8: 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. 9: 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. 10: 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. 11: Screenshot after a player has lost all her bonus.

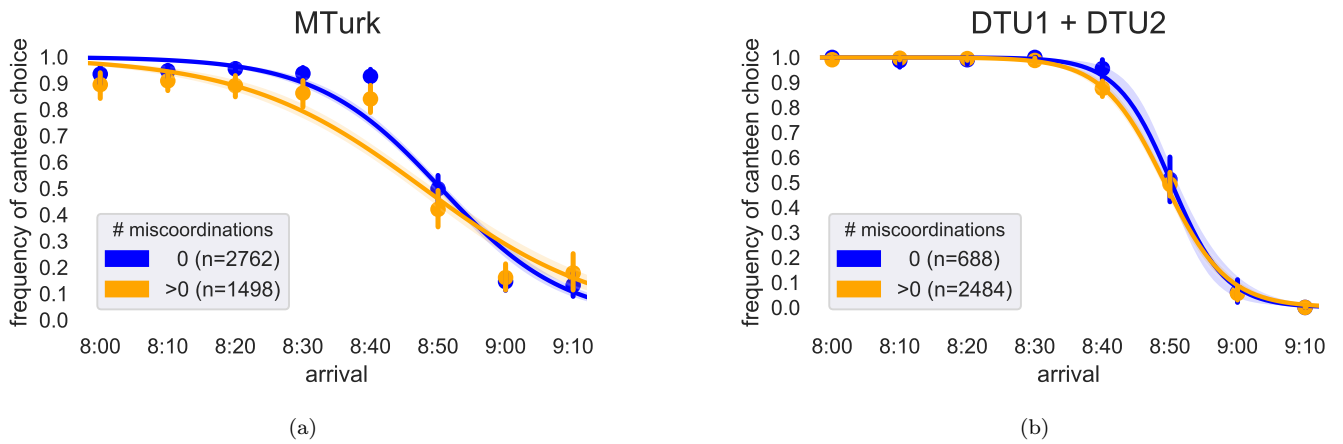


FIG. 12: Participant's decisions of going to the canteen as a function of their arrival time, here partitioned into those groups who previous have experienced zero (blue) or one or more (orange) miscoordinations. MTurk participants are shown one the left, DTU students on the right

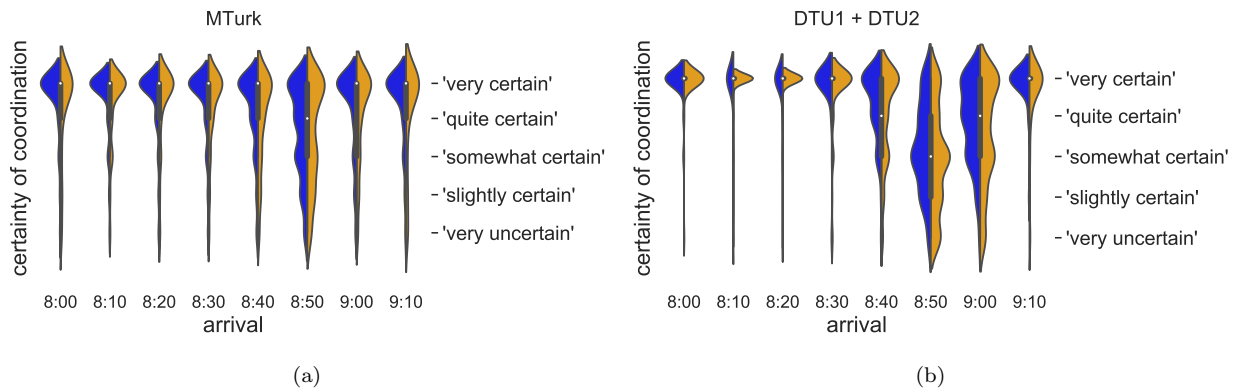


FIG. 13: Participant's certainty estimates of being able to coordinate with their colleague as a function of their arrival time, partitioned into those groups who previous have experienced zero (blue) or one or more (orange) miscoordinations. MTurk participants are shown one the left, DTU students on the right