

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

Thomas Bolander¹^{*}, Robin Engelhardt², Thomas S. Nicolet³

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

2 Affiliation Dept/Program/Center, Institution Name, City, State, Country

3 Affiliation Dept/Program/Center, Institution Name, City, State, Country

 All authors contributed equally to this work.

* tobo@dtu.dk

Abstract

Common knowledge is crucial for safe group coordination. In its absence, humans must rely on shared knowledge, which is inherently 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 involving 802 participants, we investigate the extent to which humans can differentiate between common knowledge and n th-order shared knowledge. We designed a two-person coordination game with imperfect information to simplify the recursive game structure and higher-order uncertainties into a relatable everyday scenario. In this game, coordination for the highest payoff requires a specific fact to be common knowledge between players. However, this fact cannot become common knowledge in the game. The fact can at most be n th-order shared knowledge for some n . Our findings reveal that even at quite shallow depths of shared knowledge (low values of n), players behave as though they possess common knowledge, and claim similar levels of certainty in their actions, despite incurring significant penalties when falsely assuming guaranteed coordination. We term this phenomenon ‘the curse of shared knowledge’. It arises either from the players’ inability to distinguish between higher-order shared knowledge and common knowledge, or from their implicit assumption that their co-player cannot make this distinction.

1 Introduction

Anybody who has been caught up in a sidewalk dance knows that human coordination without common knowledge about each other’s intentions is prone to failure. Thus, cooperating humans have evolved techniques by which to create common knowledge in practice, such as eye contact, conventions, and broadcasted messages. When common knowledge cannot be obtained, humans routinely act upon their private knowledge, or upon their shared reasoning about other minds. While we know much about human limits to such higher-order belief reasoning, little is known about at which point, and under which circumstances, shared knowledge becomes effectively indistinguishable from common knowledge. We investigate this question with a novel recursive coordination game with imperfect information, and show that people consistently act as if they have common knowledge about a fact, even in situations where they share only the shallowest levels of their state of mind.

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 on the same mountain, and have agreed to meet at the top. The equipment essential for an overnight stay at the top has been divided between their backpacks, so it’s crucial for their safety that they meet there before it gets dark. On the way to the top, they both observe bad weather approaching, which makes it safer to go back to the base. However, they are both uncertain about whether the other person has seen the bad weather, so they cannot expect the other person to necessarily go to the base. The most dangerous situation is if one of them goes to the base and the other goes to the top, since then one of them is left at the top without the necessary equipment to stay there for the night.

At this point they both know the fact that “bad weather approaches”, but don’t know whether the other knows. In this situation, we would say that Agnes and Bertram both have *private knowledge* that bad weather is approaching, 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]. The exact border between what is considered private and shared knowledge vary between authors. Some consider the case where everybody knows a fact to still only be private knowledge, and that for it to be shared knowledge, at least one agent also needs to know that they all know. The notion of shared knowledge we use in this paper, where shared knowledge only requires that everybody knows, is the one originally used by [3], and also the one often seen in (epistemic) logic [6]. Formal definitions of private, shared and common knowledge as well as a more detailed discussion of our choice of conventions can be found in S1 Supplementary Material.

Since Agnes doesn’t know whether Bertram knows about the bad weather, she would like to warn him and suggest to meet back at the base. So she sends a text message: “Bad weather approaching. Let’s meet at base.” However, due to the unstable mobile network signals, she is not certain that the message will go through. Therefore she asks Bertram to confirm that he received the message. A few minutes later, she receives his confirmation. At this point it has become *shared knowledge to depth two* (or

second-order shared knowledge) that bad weather is approaching: He knows that she knows, since he received her message, and she knows that he knows, because she received his confirmation. To have *shared knowledge to depth three* (*third-order shared knowledge*), it would additionally be required that A) she knows that he knows that she knows, and B) that he knows that she knows that he knows. In fact, A already holds, since she received (and read) his confirmation. However, B doesn't hold, since he will be uncertain about whether his confirmation was successfully received. Thus we have asymmetry in the level of knowledge of the two agents. If he also asks her to confirm his message, and he receives such a confirmation, then of course he will get to know that she knows that he knows. Then there will be shared knowledge to depth three. However, there will still be a (higher-order) knowledge asymmetry, since Agnes 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? 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 Agnes plans to go to the base, but she is still uncertain whether he knows. And if he doesn't, she might risk leaving him alone at the top. So shared knowledge to depth one is clearly not sufficient. Even shared knowledge to depth two is not sufficient. After he has confirmed receiving her original message, from his perspective it is still entirely possible that she doesn't know that he knows, and hence, she might decide to go to the top to not leave him alone up there. And in this case, he also has to go to the top in order not to leave *her* alone up there. 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 . Shared n th-order knowledge of some fact for all n is called *common knowledge* of the fact (see again S1 Supplementary Material for a more formal definition). The essence of the argument is that any plan for Agnes and Bertram to coordinate without common knowledge would leave one stranded at the top in the case of a faulty signal at a critical point in their communication.

In practice, shouldn't it be possible to coordinate meeting at the base after one or two messages being sent back and forth? It is likely that this is what Agnes and Bertram would attempt, at least. The question is whether they are fully aware of the risk of miscoordination in this case? After Agnes has sent the message and Bertram has successfully confirmed receiving it, are they aware that the content of the message is still not common knowledge, and that they still risk miscoordination? Or do they now believe it completely safe to go to the base? The questions here are whether A) they

¹To 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]. The example here is in fact modelled after the coordinated attack problem studied extensively in that book, where it is also used to illustrate that common knowledge is a necessary condition for coordination. We are making a similar point here about the necessity of common knowledge for coordination, just using another example.

can distinguish between the case of n th-order shared knowledge and common knowledge, and whether B) they would act differently in the case of common knowledge than in the case of n th-order shared knowledge for some suitable large n .

We have designed a coordination game that has a similar underlying mathematical structure as this trekking example, but cast in a simpler setting, where the higher-order uncertainty is established already when the game is initialised and not, as above, through a series of message passings. In this game, our experiments show that human players indeed try to coordinate on the preferred outcome even when only having low orders of shared knowledge (corresponding to trying to meet at the base after only a few message passings in the trekking example). However, this often results in severe cases of miscoordination with high penalty, corresponding to the situation in the trekking example where one of the trekkers would leave the other at the top without the necessary equipment. Our game suggests that the answer to B above is sometimes a “no”: Most players definitely act as if they have common knowledge, even though they don’t, i.e., for sufficiently high levels of shared knowledge, they act as in the common knowledge situation. For sufficiently high levels of shared knowledge, they play strategies that would have given them the highest payoff if they had had common knowledge, but since they don’t, they end up with payoffs that are significantly lower than what they could have received if they had explicitly distinguished the common knowledge situation from the n th-order shared knowledge situation. One might however still speculate that this is a conscious choice of strategy, and that the players are nevertheless aware that they don’t have common knowledge. However, through follow-up questions and interviews, it seems that the answer to A is also “no”: Players claim they have common knowledge and that coordination is guaranteed, even when they only have low-order shared knowledge.

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. We originally developed our game to illustrate that the difference between n th-order shared knowledge and common knowledge can have actual strategic implications, hence helping to understand the notion of common knowledge better (and why n th-order shared knowledge for sufficiently high values of n is not sufficient). The game we developed makes it clear that human intuitions about common knowledge can be misleading and may have costly consequences.

2 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 [7, 8]. Because of this, coordinating species typically use heuristic shortcuts in order to reduce complex shared knowledge states, such as joint perceptual cues and broadcasted signals [9–11]. Humans may obtain common knowledge via mutually accessible first-order sensory experiences [12–15], eye contact [16], public rituals and conventions [2], or salient focal points [17]. 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 [18, 19]. 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) [20], “mentalizing” [21], “mind

reading” [22,23], “mind perception” [24], “perspective taking”, and “social intelligence” [25], which often are used interchangeably for studying the cognitive mechanisms of shared knowledge, but sometimes focus on slightly different ideas and associated meanings [26], depending on the field of investigation.

Looking at the ToM-literature, conclusions about human belief reasoning abilities are rather heterogeneous [27,28]: Human reasoning about the reasoning processes of other humans is limited [20,29–34], contextual, and possibly domain specific [35–37]. 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 [38,39]. Adults may reliably master up to four orders [40], 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 [41,42]. 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 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 [44]. When humans compete or try to detect cheaters, higher-order belief reasoning seems to perform better than lower-order belief reasoning [45]. In negotiations and other mixed motive situations, where innuendo, threats, 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 [46–48]. ToM proficiency may also be facilitated by providing games with stepwise increase in ToM [49].

In pure coordination problems, such as pedestrians choosing sides, or people agreeing on new words or on new technical standards, common knowledge is the normatively preferred informational state for all members of the group, because it allows to coordinate on 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 [50] suggests that higher-order ToM reasoning may improve coordination efforts, other work seems to suggest that coordination favours lower orders of ToM sophistication [51,52]. 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 [53–56].

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 [44]. Controlled pure coordination experiments in social settings on market collaboration [57], the bystander effect [58], indirect speech [59], self-conscious emotions [60], and charity [61], consistently find that people indeed make strategically different choices under common knowledge conditions (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

²Especially when those successions of mental states are qualified by psychological attribute words such as ‘Alice thinks that Bob is mistakenly worrying that Carol is offended by misunderstanding something Dave had said to Evelyn’. [43].

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 [57], which is in contrast to the assumptions of standard rational choice theory in which payoffs are expected to be maximized [62].

Most of these studies are designed one way or the other to determine how far people typically unpack shared knowledge and whether they can recognize common knowledge when presented with it. They show that humans are indeed able to discriminate common knowledge from shared knowledge, and adapt their actions accordingly, at least in cases where it is simple to discriminate between the two states of knowledge. The purpose of our paper is somewhat different. The purpose is not to determine whether players can for instance distinguish n th order and $(n + 1)$ th order shared knowledge, but whether they can in general distinguish shared knowledge (of some order) from common knowledge. In the above-mentioned experiments, common knowledge was achieved via public announcements, and shared knowledge via private announcements in which the depth of knowledge was explicated. In other words, these were experiments in which the distinction between shared and common knowledge was made clear and explicit. The question then remains how good humans are at recognizing the difference between shared and common knowledge when the difference is not made explicit, but has to be deduced? The fact that humans are able to adapt their actions to whether they are in a shared or common knowledge situation might have significantly less practical relevance if humans in general have a hard time to distinguish these in the wild. The results of this paper suggest that humans are indeed not very good at recognizing the difference, in particular that they are prone to mistake n th-order shared knowledge for common knowledge—even for relatively low values of n . Or, at least, they behave as if they had common knowledge even when they only have n th-order shared knowledge for very moderate values of n . This issue has not been thoroughly investigated in previous research, 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 [63], a game version of the mountain trekking example 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 is expected to make the two mountain hikers both go to the base, whereas if there is only n th-order shared knowledge for some n , they are expected to both 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 optimal strategic behaviour in the game with no experiments, and only speculates what people playing the game might do. This paper aims at elucidating what actual human intuitions command in situations like this. Which depth of shared knowledge (if any) would be enough to attempt risky coordination that would lead to maximal payoff if successful?

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 indicate that the lack of common knowledge does not defer people from being confident in their ability to coordinate, except at the very lowest depths of shared knowledge. So people behave as if having common knowledge, despite only having a low level of shared knowledge, and despite the significant payoff penalties incurred. Our conjecture 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. Our follow-up interviews seem to confirm this conjecture. 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.

3 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. [64, 65]. 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 [66, 67]. 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.”

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 three supplementary classroom experiments, with respectively 80 students (DTU1: $n = 2160$) from the Technical University of Denmark (DTU) taking a course on Artificial Intelligence and Multi-Agent Systems, 42 students (DTU2: $n = 1012$) taking an introductory course in Artificial Intelligence, and 12 students (DIS: $n = 180$) taking a course on Artificial Intelligence at DIS—Study Abroad in Scandinavia (DIS). The two DTU classroom experiments differed slightly from the MTurk experiment in that the students got an initial endowment of \$30 instead of \$10 given to the MTurk participants, and played 30 rounds instead of 10 (see below). Also, in the DTU 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 S1 Supplementary Material for a full list of those questions). The DIS classroom experiment also differed in the formulation of the game itself, and the way it was presented to the students, see details further below.

The game has three possible outcomes in each round: 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*. At the beginning of each round, participants are told only their own arrival time, and based on this they will have to decide whether to go to the canteen or to the office. After choosing one of those, participants are asked to estimate their certainty that their colleague will make the same choice (on a five-point Likert scale).

A fixed participation fee of \$2 is given to all MTurk players. Additional bonuses are calculated as penalties subtracted from the initial endowment of \$10 (MTurk) or \$30 (DTU) in each round. Payoffs 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). This means that if a pair of players miscoordinate often, they risk losing all their endowment, in which case their game terminates prematurely. Penalties are calculated by a logarithmic scoring rule, which depends on the decision and the certainty estimate by each player. Using penalties instead of bonuses may at first seem an unintuitive way to reward participants, but logarithmic scoring rules have been shown to work well due to their ability to ensure that loss minimization remains central, and that the players' actions represent their actual beliefs [68–71]. As noted in Appendix B, we find a good match between certainty estimates and choices at arrival times different from those that are prone to miscoordination. This corroborates that players tried to minimize their losses and that they made their choices and certainty estimates as honestly as possible.

We included payoff examples of both successful and failed coordinations in the instructions shown to the participants before the game started, as well as during the game in each round. Screenshots and full descriptions of the experimental setup can be found in S1 Supplementary Material and details of the payoff structure are provided in Appendix B, Payoffs and Penalties.

The DIS classroom version of the game experiments were different, since we wanted to test the robustness of our findings against variations in the specific formulations of the game and its setup. Additionally, we wanted to address two specific potential criticisms of the original game setup. First of all, one might speculate that the formulation of the game might create a certain behavioural bias, for instance a bias towards going to the canteen due to a personal preference for coffee over office. Second, when doing experiments on MTurk, the players in principle cannot be guaranteed that their co-players have seen the same rules, or that they are even human. Hence the rules and the game setup might not be common knowledge, as is otherwise expected. In the DIS classroom experiments, the following rules were publicly broadcast by being shown on the classroom projector and being read up loud (see S1 Supplementary Material for the full version):

The 2 players have to try to coordinate their actions, and if they are successful, they will both achieve the same positive reward [...]. The positive reward can either be \$1 or \$2. If you're miscoordinated, you lose all the money you earned so far in the game. [...]

In each round, each player is dealt a face down card. On the face of the card there is a number in the interval 2-10 [...]. The numbers on the two dealt cards are always exactly 1 apart, so if for instance one of the players gets a 3, then the other will either be getting a 2 or a 4. [...]

When the cards are dealt, each player looks at the number on their own card without showing it to the other. The two players are not allowed to communicate or exchange any kind of information during the game. After

each player has inspected their own card, they should hide either a white or a black marble stone in their hand and put the hand on the table. [...]

Both players then receive rewards depending on the colors of their stones and the numbers on their cards. The white marble stone is worth \$2, and the black marble stone is worth \$1, but you only get to keep your stones (money) if you choose the same stone as the opponent, and additionally, if you both choose the white stones, you only get to keep them if both card numbers are strictly below 9. In more details:

- a) If one player chose a white stone and the other a black stone, then they both lose all the stones (money) they have received in the game so far.
- b) If both players chose a black stone, they both receive \$1 (i.e., a black stone).
- c) If both players chose a white stone, then they both receive \$2 (i.e., a white stone), but only if both numbers are strictly below 9, meaning they are both in the interval 2-8. If not, the players lose all the money (stones) they have received in the game so far.

This new version of the game is structurally equivalent to the original, but now formulated in a completely abstract way. This potentially makes it harder to grasp the rules, but avoids the potential problem of players making choices affected by everyday personal preferences. By reading up the rules and showing them on the projector in class solves the issue regarding whether the game rules are common knowledge. The reward structure is different in this version of the game. Instead of logarithmic scoring, we chose a simpler scoring with an even higher penalty (lose everything) for miscoordination (different stones) and “forbidden” choices (choosing white when having 9 or 10). This might help making it more clear to the players that choosing white is never worth the risk of miscoordination.

The overall point of both versions of the game is that the only ‘safe’ strategy is to always coordinate on the choice that gives the lower payoff (going to the offices or choosing black stones). Even if it seems safe to try to coordinate for the higher payoff when arriving sufficiently ahead of 9:00 am or getting a card with a number sufficiently below 9, it is never common knowledge that both arrived before 9:00 am or both have a card below 9, which makes it unsafe to attempt coordinating for the higher payoff.

4 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 decisions, they will both leave the experiment with \$9.80, excluding the \$2 participation fee (see again Appendix B for details of the payoff structure). 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, see Appendix B). 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 a person/player, say Agnes, ever go to the office that early? She knows that her co-player/colleague, say Bertram, will then be arriving at latest at 8:40, which still leaves plenty of time to get a cup of coffee together before 9:00. The issue is of course that if Agnes takes the perspective of Bertram, then assuming he arrives at 8:40, he will consider it possible that she arrives at 8:50. And if she had indeed arrived at 8:50, she would consider it possible that he arrives at 9:00. In that case she 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

³ We now prove this claim. Suppose *a* has the strategy of going to the canteen at some time *t* before 9:00, and that both *a* and *b* will go to the office if they arrive at 9:00 or later. To derive a contradiction, suppose *a* and *b* are always coordinated, i.e., always make the same choice for any pair of arrival times. Since *a* chooses to go to the canteen at time *t*, player *b* also has to go to the canteen at time *t* + 0:10, since otherwise they would be miscoordinated when *a* arrives at *t* and *b* at *t* + 0:10. But if *b* goes to the canteen at time *t* + 0:10, *a* also has to go to the canteen at time *t* + 0:20, since otherwise they would be miscoordinated when *b* arrives at *t* + 0:10 and *a* at *t* + 0:20. This can be generalized to conclude that *a* would have to go to the canteen at any time *t* + 0:20*x* for *x* ≥ 0 and *b* would have to go to the canteen at any time *t* + 0:10 + 0:20*y* for *y* ≥ 0. Clearly this implies going to the canteen after 9:00, which is a contradiction.

Table 1

Treatment	N	MaxRounds	avg. rounds	n	Ruin	payoff	Avg penalty
AMT	680	10	6.3	4260	52.79	23.59	-1.59
DTU1	80	30	27	2160	17.5	27.03	-0.83
DTU2	42	30	24.1	1012	30.95	24.11	-0.98

Aggregate statistics of all three experiments. *N = number of subjects; pairs = number of pairs; R = maximum number of rounds; \bar{r} = average number of rounds played; n = number of choices; Ruin = percentage of participants loosing 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.

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 of the canteen-before-9 strategy in a single-round game will be $\$10.00 - \$9.21 \cdot 1/6 - \$0.01 \cdot 5/6 = \8.46 , whereas the all-office strategy guarantees a payoff of $\$10.00 - \$0.02 = \$9.98$, cf. Appendix B).

In Appendix C, Formal Analysis, we make the reasoning about game strategies formally precise. We show that independently of the particular payoff structure (only using the relative 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 when arriving sufficiently ahead of 9 am (e.g. before 8:50), but unsafe when arriving later (e.g. at 8:50 or 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 at or 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 9:00 (see Appendix C for more details).

5 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 pairs (14%) were able to play 10 rounds and still have any bonus left, while the average number of rounds played was 6.3, see Table 1. As soon as one of the players had no money left, the game would terminate.

Comparing the MTurk experiment with the DTU experiments in Table 1, 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 loosing only \$0.83 per round on average.

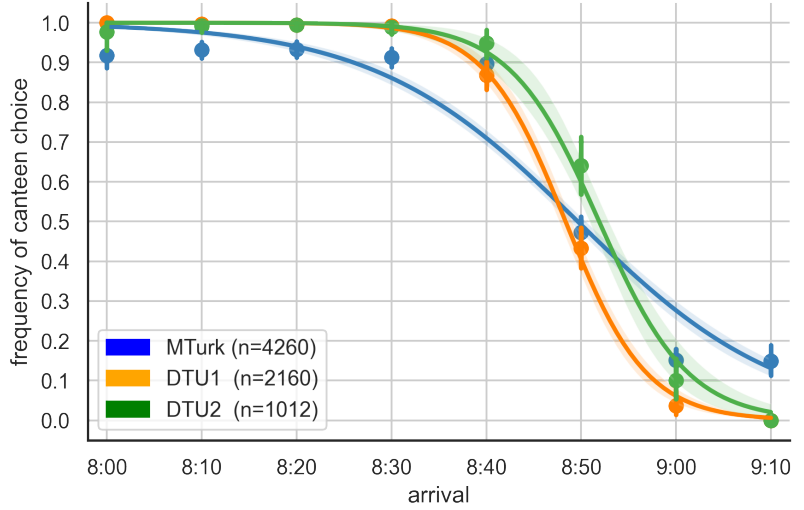


Fig 1. Percentage of canteen choices as a function of arrival times. The colored lines are logistic regression lines with 95% confidence intervals shown as translucent bands. Fitted parameters show significant differences in the slope and intercept between MTurk and DTU experiments ($p < .0001$).

To model the relationship between arrival time and the probability of choosing the canteen (choice), a logistic regression is performed in Fig. 1. The model computes a binomial model with bootstrapped confidence intervals (CI = 95%) for the estimated parameters. Bootstrapping is performed with 10,000 iterations to ensure stable estimation of confidence intervals. MTurk participants in blue have a slightly more gradual decline of canteen choices for increasing arrival times. However, the point at which there is a 50% probability of choosing the canteen or the office is close to 8:50 in all three experiments (see further specifications in an additional mixed-effects regression analysis in Appendix A.1). Later, in Fig. 6, we make a combined regression analysis of all three experiments in order to understand the overall experimental results in terms of degrees of shared knowledge.

The violin plots in Fig. 2 show the distribution of certainty estimates for each arrival time. The DTU experiments are merged and compared with the MTurk experiment, as the differences between the two DTU experiments generally are smaller than their respective difference to the MTurk experiment. This is quantified in Table 1, as the calculated average penalties per round are a direct consequence of the certainty estimates given by the participants. In general, however, Fig. 2 shows that it is exceedingly rare for any of the participants to consider it problematic to go to the canteen when arriving early. Arriving at 8:30 or earlier is deemed sufficiently early to visit the canteen with very high confidence. The subtle differences in certainty estimates between MTurk participants and DTU students show that the latter tend to be slightly 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 at 8:50 and 9:00). 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

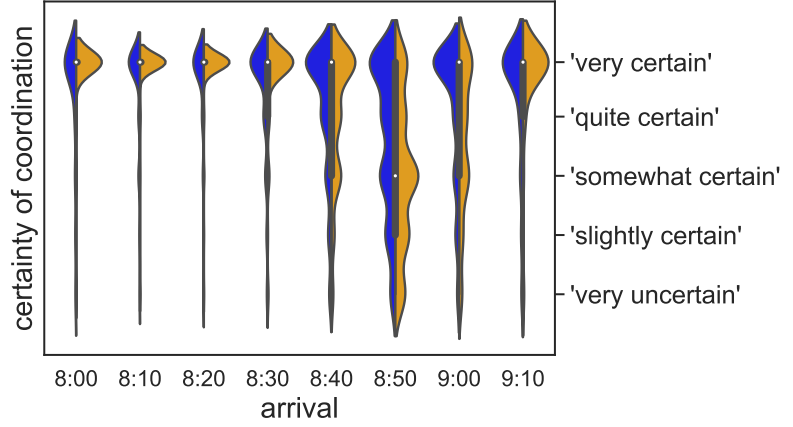


Fig 2. Violin plots of certainty estimates. In each round, participants were asked how certain they were of successful coordination with their colleague. Blue areas show the results from MTurk ($n = 4260$) and orange areas show the results from DTU1 and DTU2 combined ($n = 3172$). We predefined a five 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 Appendix B). The white dots correspond to the median certainty estimate.

however still relatively minor. In the following we will combine data from all three experiments.

In Fig. 3, we see the percentage of successful coordinations into the canteen or office (green or purple) together with the number of miscoordinations (red) as a function of all possible arrival time combinations. The figure shows that players are able to coordinate into the canteen more than 86% of the time if both of them arrive before 8:50. As soon as a pair has a player who arrives at 8:50, however, the result changes drastically. Suddenly almost half of such pairs 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.

Fig. 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 choices. This indicates that there is no behavioral change during the game. 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 Fig. 1 into two bins, corresponding to pairs having had no miscoordination and pairs having had one or more miscoordinations (see the supplementary data analysis in S1 Supplementary Material), shows somewhat decreasing certainty estimates around the critical arrival times. However, this does not affect their actual choices. MTurk participants do choose the canteen a little less often after a miscoordination (see Fig. 6 in S1 Supplementary Material), but this does not translate into better payoffs as later miscoordinations just move to earlier arrival times. So even though participants learn

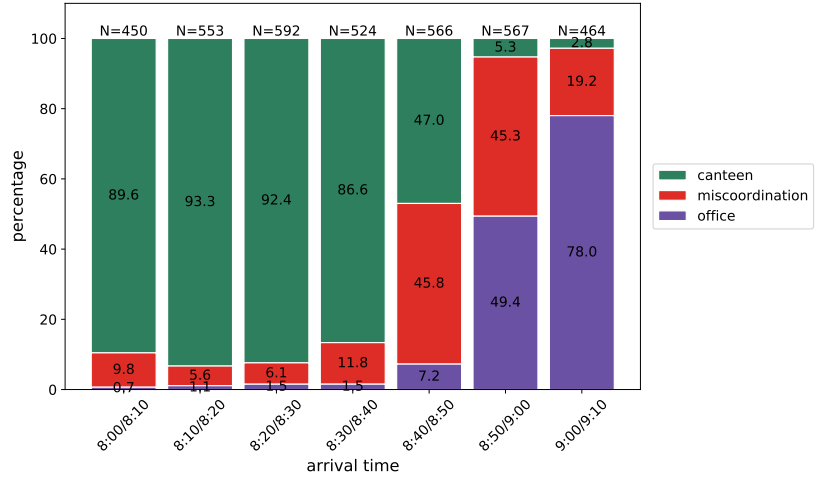


Fig 3. Percentage of coordinations and miscoordinations as a function of arrival times. Green means coordinating into the canteen, purple means coordinating into the office, and red means miscoordination. We use the notation 8:00/8:10 to denote the union of the arrival pairs (8:00, 8:10) and (8:10, 8:00), 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.

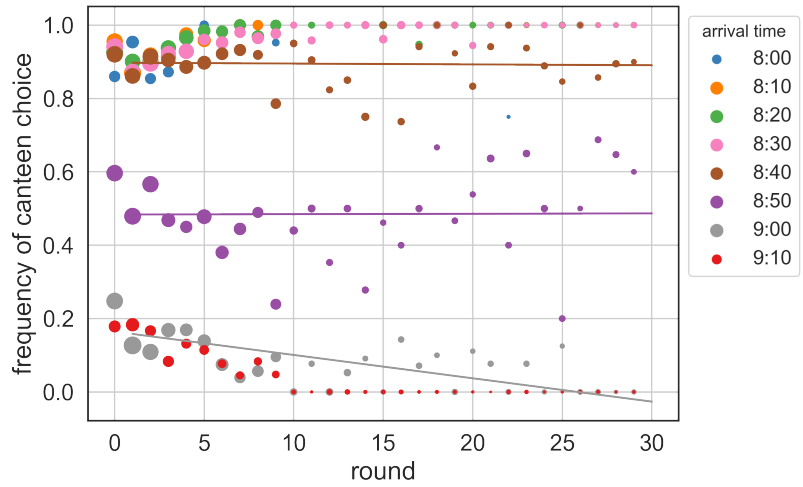


Fig 4. Mean frequencies of canteen choices for all possible arrival times as a function of the number of rounds played. The three fitted horizontal lines, corresponding to arrival times 8:40, 8:50, and 9:00 respectively, represent weighted linear squares (WLS) with the weights chosen to be the square root of the number of data points constituting the mean frequencies for each round, also shown by dot size.

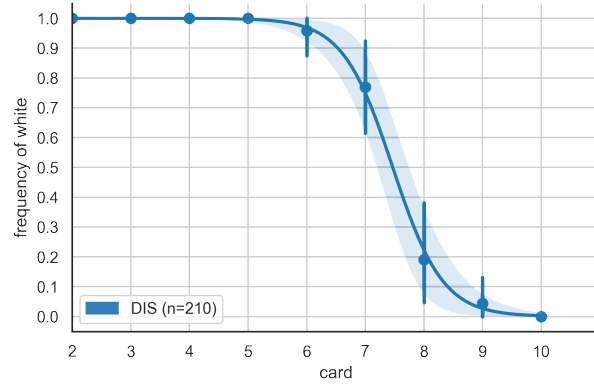


Fig 5. Logistic regression of the relationship between card number and choice of a white stone in the qualitative experiment at DIS. The resulting curve shows a very similar overall trend as in the main experiments shown in figure 1.

that their choices are risky, they do not see any way to improve their strategy (at least not in the short term). Specifically, they never converge to the optimal all-office strategy, and also not to the alternative canteen-before-9 strategy (cf. Theorem 5 in Appendix C). This apparent lack of behavioral change in higher-order social reasoning games was also identified by Verbrugge [72].

Fig. 5 shows the results for the alternative version of the game conducted at DIS, where players are dealt cards with numbers from 2 to 9, and choose either a black or a white stones (cf. Section 3). Choosing a white stone corresponds to choosing office in the original version of the game, and getting a card with the number n corresponds to arriving at $9:00 + 00:10(n - 9)$ in the original game. So e.g. receiving card 7 corresponds to arriving at 8:40, receiving card 8 corresponds to arriving at 8:50, and receiving card 9 corresponds to arriving at 9:00. The results in Figure 5 are very similar to the results of the original version of the game, Figure 1. The curve has the same shape, and with a slope that is most similar to the DTU experiments (which makes sense, as here the participants are also students in computer science and AI). The choices at 6 (8:30) are more than 90% white (canteen), and the choices at 9 (9:00) are more than 90% black (office), similar to the DTU1 experiments. The choices at 8 (8:50) are around 20% white (canteen), which is lower than for any of the original experiments, suggesting a potentially better understanding of the risk of miscoordination. Alternatively, it might be due to the fact that in this version of the game, miscoordination leads to immediately losing all earnings, so the penalty of miscoordination is even clearer and higher in this version of the game (this was a deliberate choice in our game design, in order to make sure that the observed results of the original game were not due to unawareness of the high penalty of miscoordination).

The DIS classroom experiment was followed by a number of qualitative interviews with the students, divided into groups of 4. These confirmed that the players considered the rules to be common knowledge. When asked about the existence of safe strategies guaranteeing coordination (always choosing the same stone), most players expressed the false belief that a cut-off strategy existed that would guarantee coordination on white stones for sufficiently low numbers (which we proved false in footnote 3). Most players believed that the cut-off would be around 6 or 7, although some realized during the interview that such a strategy can never guarantee coordination. Only one student was skeptical about the existence of a safe strategy for coordinating on white from the very beginning of the interview.

6 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 successfully coordinating into the canteen before 9:00. 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 9:00, 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. Fig. 1. Thus, other considerations in addition to the player's private knowledge must play a role in their decision making.

When both participants know they 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 an *arrival pair* (t_1, t_2) denotes that player 1 arrives at time t_1 and player 2 at time t_2 , see Appendix C, Formal Analysis, for further details). 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 there 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 Fig. 6. Note that in general, if a player arrives at time $8:50 - 0:10n$, $n > 0$, 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 9:00 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 of having sufficient time to meet for coffee 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 Fig. 6). 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 Fig. 2), showing that the participants are not completely ignorant to the differences. This could suggest that many participants believe that it is less safe to go to the canteen based on n th-order shared knowledge than $(n + 1)$ st-order shared knowledge. Very few seem to draw the conclusion that it is never safe to go to the canteen, or, if they do, they at least don't expect the other player to be able to draw the same conclusion.

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 also be unsafe, etc.? One reason may be that the all-office strategy is cognitively difficult to comprehend given the limited ability of humans to take the perspective of each other recursively. One might also speculate that the players suffer from a computational overload from the complexity of the game's rules and payoff structure. While this may be a problem, we do not think it has much significance for the following reasons. If

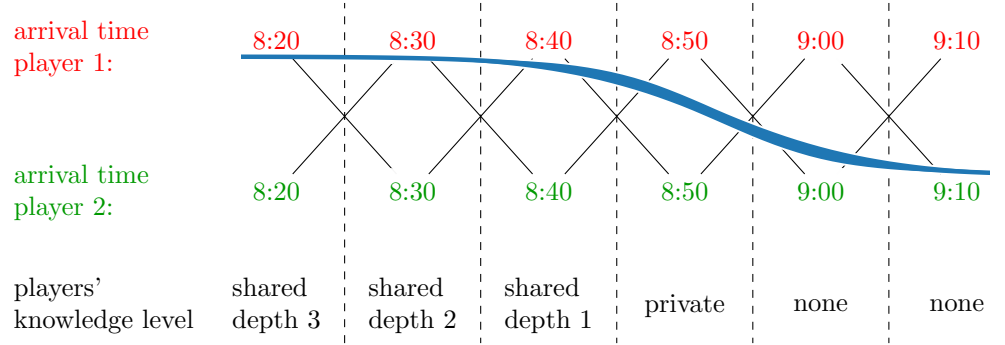


Fig 6. The solid diagonal lines express indistinguishability for the players, e.g. the arrival time 8:40 for player 1 has a line to both of the arrival times 8:30 and 8:50 for player 2, since these are the two arrival times for player 2 that player 1 will consider possible when herself arriving at 8:40. Below each possible arrival time, we have marked the highest level of knowledge of there being 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 ($n = 7432$).

computational overload was sufficient to distract our participants, we might expect more varied suboptimal choices. While we do see some bluntly irrational choices of canteen at 9:00 (given the rules and payoff structure), the vast majority choose office at 9:00 and 9:10, and canteen at earlier arrival times, with many still choosing office at 8:50. Besides this, we see participants be much less certain of their canteen choices at 8:50 or 8:40 than at earlier arrival times.

Another reason for not considering earlier office choices as viable may be that while the optimality of an all-office strategy are understood, the participants do not believe that this is common knowledge, and hence they do not rely on the other participants to choose the same strategy. That is, they either think their co-player does not realize there is no safe cut-off, or that their co-player thinks this about them, and so on. In this way, the strategic reasoning involved becomes akin to the Keynesian beauty contest [31, 73]. Under this assumption, we should not necessarily expect participants to play an all-office strategy, since that is only proven optimal when assuming there to be common knowledge of perfect rationality (see Appendix C). A possible alternative strategy may be the following mixed strategy: 1) always go to the canteen before 8:50, 2) always go to the office after 8:50, and 3) go to the canteen with probability 0.5 at 8:50. If both players follow this strategy, 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 Fig. 3. If this is the strategy followed, and if players assume their co-players to follow the same strategy, they should also be aware of the 50% probability of miscoordination at 8:50. In Fig. 2, we indeed see a much lower coordination certainty estimate at 8:50. However, if players were indeed playing this strategy and assuming their co-players to do the same, their certainty estimates ought to be even much lower at 8:50. These certainty estimates also ought to be lower if we assume that players are aware of the possibility of their co-players not being perfectly rational and their chosen strategy hence to be somewhat unpredictable.

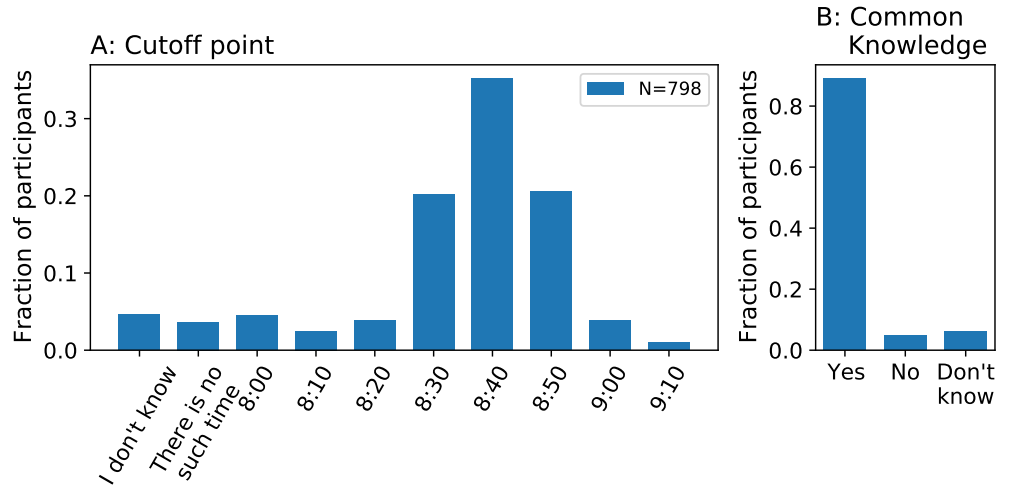


Fig 7. A) Frequencies of answers to the question: “Imagine you could have agreed beforehand with your colleague about a point in time where it is safe to go to the canteen. What time would that be?” Due to the pragmatics of language, we assume that an answer like 8:30 entails the belief that all earlier 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?”

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 Fig. 7A 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. The interviews with students in the DIS experiment indicates the same. One student claims for example that “(from) two to five, there is no risk involved”. This also reflected in the results, as Fig. 5 shows, no students plays black (equivalent to an ‘office’ choice) for numbers 5 and below. One interview from the DIS experiment also ends with a student asking “Did you find an optimal number?”.

The answers to the post-game question above, as well as the DIS interviews, indicates that the prevalence of canteen choices is not due to uncertainty about the rationality of other players, but instead due to a more immediately available belief, that there is a point where canteen choices (or white stones in the DIS experiment) can be played without the risk of miscoordination. In reality, they would need common knowledge in order for the canteen choice to be without risk, and the players seem to behave as if they have such common knowledge. Participants might of course not necessarily have a precise idea of the technical notion of common knowledge and how it differs from n th-order shared 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—at least in cases where the difference between these states of knowledge is relatively clear (for instance when the common knowledge situation is achieved via a public announcement). 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. We also see that player behavior does not match what we would expect to see if they only believed to have n th-order shared knowledge, but instead matches what we would expect to see if they indeed wrongly inferred common knowledge.

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 Fig. 7B, 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.

7 Conclusion and future work

We developed a new coordination game, the *Canteen Dilemma*, to examine higher-order social reasoning in humans. Our experimental results indicate a substantial degree of miscoordination, even in scenarios where coordination *is* achievable (albeit only for the second-highest payoff). This miscoordination appears to stem from what we term the “curse of shared knowledge”: an illusion of common knowledge in situations where only shared knowledge exists, and only to a limited depth. While the game results alone do not conclusively demonstrate that this illusion of common knowledge occurs, they show that players behave as though they have common knowledge. Moreover, variations to the game and responses to post-game questions strongly suggest the presence of this illusion. For instance, an overwhelming majority of participants report having common knowledge about the safety of arriving at the canteen at 8:10 (Fig. 7B). Additionally, qualitative interviews following the DIS classroom experiment further supported this interpretation.

While we cannot claim to have definitely proven the existence of a ‘curse of shared knowledge’, our experiments support it as a plausible conjecture that warrants further investigation. This result is somewhat surprising, as it initially appears to contrast with existing literature on shared and common knowledge. Previous findings [44, 57–61] suggest that people make strategically different choices depending on whether they have shared or common knowledge.

The key difference between our study and these previous works lies in the experimental designs: Earlier studies explicitly inform participants about their level of knowledge—whether it is shared or common—before asking them to make decisions. In contrast, our study requires participants to deduce their level of knowledge independently, without explicit clarification. Thus, while prior research investigates decision making when knowledge levels are explicitly stated, our work examines whether

individuals can distinguish between shared and common knowledge when this distinction is not made explicit.

Our findings suggest they cannot; at least not in our game. It is possible that specific aspects of our experimental design make it particularly challenging to distinguish between shared and common knowledge. However, this raises a broader question: If people make different strategic choices when being explicitly informed about their knowledge level in controlled experiments, does this have implications for real-world decision making where such explicit clarification is unlikely or at least probably relatively rare?

For individuals to make rational decisions based on whether they are in a shared or common knowledge situation, two conditions have to be met: 1) the ability to accurately identify their knowledge situation and 2) the ability to deduce the strategic implications of these situations. While prior research suggests that people can reliably accomplish the second task, our findings imply they may struggle with the first. Future work should explore human decision making in more ecologically valid, real-life scenarios where shared and common knowledge are relevant, but participants are not explicitly informed about their knowledge levels.

If our experiments indeed reveal a curse of shared knowledge, as we conjecture, this is a somewhat surprising result. In most cases, agents' bounded rationality and lack of logical omniscience lead them to draw fewer conclusions than they otherwise might [1, 74]. However, in our case, these limitations appear to have the opposite effect: They cause individuals to conflate shared knowledge with common knowledge, leading them to conclude more than is warranted (e.g. believing common knowledge exists when it does not).

From studies on coordination, particularly the coordinated attack problem [1], we know that successful coordination generally requires common knowledge [75], which is often unattainable. Common knowledge is relatively rare, typically arising only in situations of joint attention, such as during public announcements. Despite this, humans frequently attempt to coordinate actions and decisions based solely on shared knowledge, accepting the inherent risk of miscoordination. In most real-life situations, the penalties for miscoordination are likely lower than in the Canteen Dilemma, making the risk more acceptable (though also in the coordinated attack problem, the penalty for miscoordination is severe). What the Canteen Dilemma highlights is that, in many cases, we may not even recognize the risk of miscoordination. This suggests that humans might not consciously choose to rely on lower-order shared knowledge for coordination, but instead often assume that coordination is guaranteed. This assumption might stem from an inability to distinguish lower-order shared knowledge from common knowledge. To our knowledge, this point has not been previously made.

A Experimental design and data collection

A large total sample size of $N = 870$ was chosen to get robust conclusions from the statistical analysis while giving room for high variability in behavior. In fact, we thought that there would be a minority of pairs opting for the all-office strategy. This turned out not to be the case. Experiments on Amazon Mechanical Turk had a total of 714 participants (including dropouts, see S1 Supplementary Material), while the two classroom experiments at the Technical University of Denmark (DTU1 and DTU2) had a total of 106 and 50 participants, respectively⁴.

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 [76].

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 S1 Supplementary Material.

A.1 Mixed effect logistic regression

The experimental results were also analyzed using a generalized mixed-effects logistic regression model in R (*glmer*), with the arrival time t and experimental condition exp (MTurk, DTU1, DTU2) as fixed effects including interactions and random effects for each pair of players, because we expect pairs to behave differently, depending on which levels of shared knowledge they individually take into account in their reasoning and collectively try to agree upon. Exploratory analysis including the round number as an additional fixed effect did not show any significant contributions, meaning that players did not significantly change their strategy during the game.

Using a generalized linear mixed-effects model with the arrival time and the three experimental conditions as fixed effects and with player-pairs as random effects, we obtain a logistic regression fit in Fig. 8, showing the probability of players choosing the canteen as a function of their arrival time. Results from the mixed-effects model are almost indistinguishable from the simple logit model without random effects used in Fig. 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 Fig. 6 is

⁴Complete anonymized data files and all code can be downloaded from https://github.com/gavstrik/Paper_canteen_dilemma.

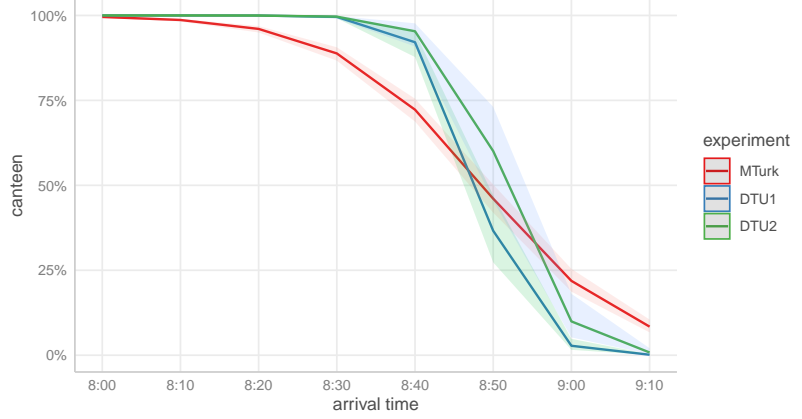


Fig 8. Percentage of canteen choices as a function of arrival times. The colored lines are logistic regression lines from a mixed-effects model with 95% confidence intervals shown as translucent bands. Fitted parameters show significant differences in the slope and intercept between MTurk and DTU experiments ($p < .0001$).

obtained by combining observations from all three experiments and using a simple logistic regression model without random effects. The high number of observations imply small confidence bands. Hence, conclusions from the models can be viewed as robust. Code and data are available on github (https://github.com/gavstrik/Paper_canteen_dilemma).

B 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 9:00 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 9:00 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 choses the canteen, $a_1 = 0$. She estimates the probability that her colleague also will to 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 strictly proper. It is therefore not surprising that we find a

good match between estimates and actual choices at arrival times different from those that are prone to miscoordinations, as seen in Fig. 2, indicating that loss minimization remained a central concern and that participants made their choices and estimates as honestly as possible [68–70].

C Formal Analysis

We now give a detailed formal analysis of the game. We want to keep things as general as possible, so we don’t assume a particular payoff structure, but any payoff structure satisfying certain constraints, and we want to consider arbitrary strategies, also mixed ones. This makes the analysis a bit non-trivial, although the underlying intuitions are rather straightforward. We of course haven’t expected our human subjects to have made a deep analysis corresponding to the one below before starting to play the game. That should also not be necessary: The crucial aspect is to realize that there is no time point at which it is safe to go to the canteen (and hence avoid the high penalty), and this only requires to reason that 8:50 is not safe, and that then also 8:40 is not safe, and that then also... etc. What we do below is just to verify in a generalised setting that this reasoning is indeed formally correct and game theoretically valid and acting according to it leads to the highest expected payoff.

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. The payoff for player 1 is always equal to the payoff for player 2 (common payoff game). We disregard 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 $|t_1 - t_2| = 0:10$. 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 $t_{\max} \geq 9:00$. Since we are counting in tens of minutes, by slight abuse of language we will call an arrival time *even* if it is of the form $9:00 + 0:20x$ for some integer x , and *odd* if it is of the form $9:00 + 0:10 + 0:20x$. The game then naturally splits into two disjoint subgames: in one of these subgames, player 1 always arrives at an odd time, and player 2 at an even time; in the other subgame, the arrival time of the two players are reversed. The two subgames are disjoint in the sense that it is always common knowledge among the players which of the two arrived at an even time. We can hence restrict focus to one of the two subgames. We will pick the subgame where player 1 arrives at an odd time. To simplify further, we can without loss of generality assume that both t_{\min} and t_{\max} are also odd times. All reasoning performed in the following generalises immediately to the other subgame and to other values of t_{\min} and t_{\max} .

We let T denote the set of arrival times and \mathbf{T} the set of arrival pairs. The game starts by nature choosing an element $\mathbf{t} \in \mathbf{T}$. 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 $\mathbf{t} \in \mathbf{T}$ 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 before 9:00, denoted u_c , is always better than successful coordination into the offices, denoted u_o , which again is always better than being miscoordinated, denoted u_m . Hence we suppose given fixed utility values $u_c > u_o > u_m$ such that for all $\mathbf{t} = (t_1, t_2) \in \mathbf{T}$, $u_{\mathbf{t}}(o, o) = u_o$, $u_{\mathbf{t}}(o, c) = u_{\mathbf{t}}(c, o) = u_m$ and

$$u_{\mathbf{t}}(c, c) = \begin{cases} u_c & \text{if } t_1, t_2 < 9:00 \\ u_m & \text{otherwise} \end{cases}$$

A *strategy* is a mapping from arrival times to actions, that is, a mapping $s : T \rightarrow \{c, o\}$. Normally one would define two strategies, one for each player. That is however not necessary in our game. By assumption, player 1 can only observe odd arrival times, so if t is odd, $s(t)$ denotes the strategic choice of player 1 at t , otherwise of player 2 at t . In this way s encodes a *strategy profile* (a strategy for each player). Since each player only observe their own arrival time, their strategic choice can only depend on their observed arrival time, which is why s is a mapping from T rather than from \mathbf{T} .

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 adapt 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.

The payoff of a strategy s at an arrival pair (t_1, t_2) is then defined as $u_{(t_1, t_2)}(s) = u_{(t_1, t_2)}(s(t_1), s(t_2))$. The *expected utility* $EU(s)$ of a strategy s is the average of the payoffs [77]:

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

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 s' *Pareto dominates* another strategy s if $EU(s') > EU(s)$ [77]. A strategy is *Pareto optimal* if there does not exist another strategy dominating it. The game is cooperative (between player 1 and 2), so both players should seek to play a Pareto optimal strategy. Also, since it is a common-payoff game, all Pareto optimal strategies have the same expected utility [77].

Note also that our strategies are pure. To make things as general as possible, one would normally also consider mixed strategies, i.e., probability distributions over strategies (or, equivalently, mappings of arrival times into probability distributions over actions). However, that is not necessary here as our game is a common payoff game. As explained in [78], common payoff games always have at least one pure Pareto optimal strategy. Furthermore, any Pareto optimal mixed strategy will be a combination of pure Pareto optimal strategies. To see the latter, suppose, for $i = 1, 2$, that player i has a mixed strategy $m_i \in \Pi(S_i)$, where S_i is the set of possible pure strategies of player i , and $\Pi(X)$ denotes the set of probability distributions over X [77] (in our particular game, S_1 would assign choices to only the odd arrival times, and S_2 to the even). Suppose further that the strategy profile (m_1, m_2) is Pareto optimal. Let $S'_i \subseteq S_i$ denote the subset of strategies that are assigned positive probability in m_i . Each pure strategy profile $(s_1, s_2) \in S'_1 \times S'_2$ has an expected utility. Consider a pure strategy profile $(s_1, s_2) \in S'_1 \times S'_2$ that has maximal expected utility among all policy profiles in

$S'_1 \times S'_2$. Suppose there exists another strategy profile $(s'_1, s'_2) \in S'_1 \times S'_2$ with strictly lower expected utility. If that's the case, then the pure strategy profile (s_1, s_2) has higher expected utility than the mixed strategy profile (m_1, m_2) , which contradicts the assumption. Hence all strategy profiles $(s'_1, s'_2) \in S'_1 \times S'_2$ must have the same expected utility, hence all be pure Pareto optimal strategies. This means that we can focus on finding pure Pareto optimal strategies, since any potential Pareto optimal mixed strategies will be a combination of such Pareto optimal pure strategies.

We will now try to determine the possible candidates for being Pareto optimal strategies.

Lemma 1. *If a strategy s is Pareto optimal then $s(t) = o$ for all $t \geq 9:00$.*

Proof. We prove the contrapositive. Consider a strategy s with $s(t) = c$ for some $t \geq 9:00$. We can assume t to be odd, the other case being proved similarly. This implies $t \geq 9:10$. Now define a strategy s' which is identical to s except $s'(t') = o$ for all $t' \geq 9:00$ (meaning that for arrival times at 9:00 or after, both agents go to the office). We want to show that s' Pareto dominates s , implying that s is not Pareto optimal. First note that $u_{(t,t+0:10)}(s') = u_o$ by definition of s' . For this arrival pair, it is impossible to receive the utility u_c , so u_o is the highest possible payoff for this arrival pair. Hence any strategy not specifying (o, o) at this arrival pair will have a strictly lower (expected) payoff. Now note that since $s(t) = c$, then s does *not* specify (o, o) at $(t, t + 0:10)$, and hence we get $u_{(t,t+0:10)}(s') > u_{(t,t+0:10)}(s)$. This proves the existence of an arrival pair for which s' has a strictly higher utility than s . To prove $EU(s') > EU(s)$, we hence only need to prove that $u_{\mathbf{t}'}(s') \geq u_{\mathbf{t}'}(s)$ for all $\mathbf{t}' = (t'_1, t'_2) \in \mathbf{T}$. When $t'_1, t'_2 < 9:00$ this is trivial, as we then have $s'(t'_i) = s(t'_i)$ by definition of s' . When $t'_j \geq 9:00$ for some j , $s'(t'_j) = o$ by definition. If $u_{\mathbf{t}'}(s) = u_m$, we immediately get $u_{\mathbf{t}'}(s') \geq u_{\mathbf{t}'}(s)$, since u_m is the lowest payoff. If $u_{\mathbf{t}'}(s) = u_o$ then also $u_{\mathbf{t}'}(s') = u_o$, since s' has an o in all the places where s has. \square

Lemma 2. *Let s be a Pareto optimal strategy. For all arrival times t , if $s(t) = o$ then $s(t + 0:10) = o$.*

Proof. We prove the contrapositive. Suppose s is a strategy with $s(t) = o$ and $s(t + 0:10) = c$ for some t . We need to prove that then s is not Pareto optimal, i.e., we need to find a strategy s' that Pareto dominates s . If $s(t') = c$ for some $t' \geq 9:00$, the existence of a strategy dominating s follows immediately from Lemma 1. We can hence in the following assume that $s(t') = o$ for all $t' \geq 9:00$. Since $s(t + 0:10) = c$, we can thus also conclude that $t + 0:10 < 9:00$. Now define s' to be identical to s except that we let $s'(t') = c$ for all $t' \leq t$. We want to show that $EU(s') > EU(s)$. First consider the arrival pair $\mathbf{t} = (t, t + 0:10)$. We immediately get $u_{\mathbf{t}}(s') = u_c > u_m = u_{\mathbf{t}}(s)$. To prove $EU(s') > EU(s)$, it hence now only remains to prove that $u_{\mathbf{t}'}(s') \geq u_{\mathbf{t}'}(s)$ for all $\mathbf{t}' = (t'_1, t'_2) \in \mathbf{T}$. The only non-trivial case is when $t'_i \leq t$ for some i (in all other cases, $s'(\mathbf{t}') = s(\mathbf{t}')$). If $t'_i \leq t$ then either we also have $t'_{3-i} \leq t$ or else $t'_i = t$ and $t'_{3-i} = t + 0:10$. The latter case has already been covered, so we only need to consider the case of $t'_1, t'_2 \leq t$. Then $s'(t'_1) = s'(t'_2) = c$, by definition of s' , so $u_{\mathbf{t}'}(s') = u_c$. Since u_c is the highest possible payoff, we immediately get $u_{\mathbf{t}'}(s') \geq u_{\mathbf{t}'}(s)$, and we're done. \square

Lemma 3. *Let s be a Pareto optimal strategy. For all arrival times $t < 8:50$, if $s(t) = c$ then $s(t + 0:10) = c$.*

Proof. We again prove the contrapositive. So suppose s is a strategy with $s(t) = c$ and $s(t + 0:10) = o$ for some $t < 8:50$. We need to show that s is not Pareto optimal. We can assume t to be odd, the other case being symmetric. Consider the strategy s' that is like s except $s(t + 0:10) = c$. Note that $u_{\mathbf{t}}(s') = u_{\mathbf{t}}(s)$ when \mathbf{t} doesn't contain $t + 0:10$.

To prove that s' dominates s , we hence only need to prove that
 $u_{(t,t+0:10)}(s') + u_{(t+0:20,t+0:10)}(s') > u_{(t,t+0:10)}(s) + u_{(t+0:20,t+0:10)}(s)$. We have
 $u_{(t,t+0:10)}(s) = u_m$ and $u_{(t,t+0:10)}(s') = u_c$, so we need to prove
 $u_c + u_{(t+0:20,t+0:10)}(s') > u_m + u_{(t+0:20,t+0:10)}(s)$. Since $s(t+0:10) = o$, we get
 $u_{(t+0:20,t+0:10)}(s) < u_c$ (the highest payoff of coordinating into the canteen cannot be
reached). Since the lowest payoff is u_m , we now get
 $u_c + u_{(t+0:20,t+0:10)}(s') \geq u_c + u_m > u_{(t+0:20,t+0:10)}(s) + u_m$, as required. \square

Definition 4. 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 $t \geq t'$. The cut-off strategy with cut-off t_{min} is also called the
all-office strategy.

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 9:00. We now get the
result on optimal strategies claimed in the informal discussion.

Theorem 5. Any Pareto optimal strategy is either the all-office strategy or the cut-off
strategy with cut-off 9:00.

Proof. Let s be a Pareto optimal strategy. Then by Lemma 1, $s(t_{max}) = o$. Let t_{cut}
denote the earliest time with $s(t_{cut}) = o$. Then $s(t) = c$ for all $t_{min} \leq t < t_{cut}$. Suppose
first that $t_{cut} > t_{min}$. Then $s(t_{min}) = c$. By Lemma 3, we then get that $s(t) = c$ for all
 $t \leq 8:50$. By Lemma 1, we furthermore have that $s(t) = o$ for all $t \geq 9:00$. In other
words, s is the cut-off strategy with cut-off 9:00. Suppose alternatively that $t_{cut} = t_{min}$.
Then $s(t_{min}) = o$, and by Lemma 2, we then get $s(t) = o$ for all $t \geq t_{min}$. This is the
all-office strategy. \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,
and expect the other player to expect the same, and so on (i.e., we have common
knowledge of perfect rationality [79]), 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.

Supporting information

1007

S1 Supplementary Material. Supplementary material for the paper. This PDF file includes additional information about the experiments, the game, and the results.

1008

1009

1010

Author Contributions and Acknowledgments

1011

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

1012

1013

1014

1015

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.

1016

1017

1018

1019

1020

1021

References

1. Fagin R, Halpern JY, Moses Y, Vardi MY. Reasoning About Knowledge. MIT Press; 1995.
2. Lewis D. Convention: A Philosophical Study; 1969.
3. Clark HH, Marshall CR. Definite knowledge and mutual knowledge. 1981;.
4. Schelling TC. The strategy of conflict. Harvard university press; 1980.
5. Aumann RJ. Agreeing to disagree. The annals of statistics. 1976; p. 1236–1239.
6. Herzig A, Maffre F. How to share knowledge by gossiping. In: Multi-Agent Systems and Agreement Technologies. Springer; 2015. p. 249–263.
7. van de Pol I, van Rooij I, Szymanik J. Parameterized complexity of theory of mind reasoning in dynamic epistemic logic. Journal of Logic, Language and Information. 2018;27(3):255–294.
8. Bolander T, Lequen A. Parameterized complexity of dynamic belief updates: A complete map. Journal of Logic and Computation. 2023;33(6):1270–1300.
9. Milgrom P. An axiomatic characterization of common knowledge. Econometrica: Journal of the Econometric Society. 1981; p. 219–222.
10. Clark HH. Using language. Cambridge university press; 1996.
11. Bradbury JW, Vehrencamp SL, et al. Principles of animal communication. 1998;.
12. Tomasello M. Joint attention as social cognition. Joint attention: Its origins and role in development. 1995; p. 103–130.
13. Lorini E, Tummolini L, Herzig A. Establishing mutual beliefs by joint attention: towards a formal model of public events. In: Proc. of CogSci; 2005. p. 1325–1330.

14. Bolander T, van Ditmarsch H, Herzig A, Lorini E, Pardo P, Schwarzentruher F. Announcements to Attentive Agents. *Journal of Logic, Language and Information*. 2015; p. 1–35. doi:10.1007/s10849-015-9234-3.
15. Gintis H. Rationality and common knowledge. *Rationality and Society*. 2010;22(3):259–282.
16. Friedell MF. On the structure of shared awareness. *Behavioral Science*. 1969;14(1):28–39.
17. Schelling TC. Bargaining, communication, and limited war. *Conflict Resolution*. 1957;1(1):19–36.
18. Tooby J, Cosmides L. Groups in mind: The coalitional roots of war and morality. Human morality and sociality: Evolutionary and comparative perspectives. 2010; p. 91–234.
19. Harari YN. *Sapiens: A brief history of humankind*. Random House; 2014.
20. Premack D, Woodruff G. Does the chimpanzee have a theory of mind? *Behavioral and brain sciences*. 1978;1(4):515–526.
21. Frith U, Frith CD. Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences*. 2003;358(1431):459–473.
22. Vogeley K, Bussfeld P, Newen A, Herrmann S, Happé F, Falkai P, et al. Mind reading: neural mechanisms of theory of mind and self-perspective. *Neuroimage*. 2001;14(1):170–181.
23. Apperly I. *Mindreaders: the cognitive basis of "theory of mind"*. Psychology Press; 2010.
24. Gray K, Jenkins AC, Heberlein AS, Wegner DM. Distortions of mind perception in psychopathology. *Proceedings of the National Academy of Sciences*. 2011;108(2):477–479.
25. Baron-Cohen S, Ring HA, Wheelwright S, Bullmore ET, Brammer MJ, Simmons A, et al. Social intelligence in the normal and autistic brain: an fMRI study. *European journal of neuroscience*. 1999;11(6):1891–1898.
26. Schaafsma SM, Pfaff DW, Spunt RP, Adolphs R. Deconstructing and reconstructing theory of mind. *Trends in cognitive sciences*. 2015;19(2):65–72.
27. Apperly IA, Butterfill SA. Do humans have two systems to track beliefs and belief-like states? *Psychological review*. 2009;116(4):953.
28. Saxe R, Young L. Theory of Mind: How brains think about thoughts. *The Oxford handbook of cognitive neuroscience*. 2013;2:204–213.
29. Gopnik A, Astington JW. Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child development*. 1988; p. 26–37.
30. Stahl DO, Wilson PW. On players' models of other players: Theory and experimental evidence. *Games and Economic Behavior*. 1995;10(1):218–254.
31. Nagel R. Unraveling in Guessing Games: An Experimental Study. *The American Economic Review*. 1995;85(5):1313–1326.

32. Hedden T, Zhang J. What do you think I think you think?: Strategic reasoning in matrix games. *Cognition*. 2002;85(1):1–36.
33. Keysar B, Lin S, Barr DJ. Limits on theory of mind use in adults. *Cognition*. 2003;89(1):25–41.
34. Pinker S. The language instinct: How the mind creates language. Penguin UK; 2003.
35. Leslie AM, Thaiss L. Domain specificity in conceptual development: Neuropsychological evidence from autism. *Cognition*. 1992;43(3):225–251.
36. Saxe R. Theory of mind (neural basis). *Encyclopedia of consciousness*. 2009;2:401–410.
37. Heyes C. Submentalizing: I am not really reading your mind. *Perspectives on Psychological Science*. 2014;9(2):131–143.
38. Perner J, Wimmer H. “John thinks that Mary thinks that...” attribution of second-order beliefs by 5-to 10-year-old children. *Journal of experimental child psychology*. 1985;39(3):437–471.
39. Sullivan K, Zaitchik D, Tager-Flusberg H. Preschoolers can attribute second-order beliefs. *Developmental Psychology*. 1994;30(3):395.
40. Kinderman P, Dunbar R, Bentall RP. Theory-of-mind deficits and causal attributions. *British Journal of Psychology*. 1998;89(2):191–204.
41. Camerer C, Loewenstein G, Weber M. The curse of knowledge in economic settings: An experimental analysis. *Journal of political Economy*. 1989;97(5):1232–1254.
42. Birch SA, Bloom P. The curse of knowledge in reasoning about false beliefs. *Psychological Science*. 2007;18(5):382–386.
43. Academian. Unrolling social metacognition: Three levels of meta are not enough.; 2018. Lesswrong,. Available from: <https://www.lesswrong.com/posts/K4eDzqS2rbcBDsCLZ/unrolling-social-metacognition-three-levels-of-meta-are-not>.
44. De Freitas J, Thomas K, DeScioli P, Pinker S. Common knowledge, coordination, and strategic mentalizing in human social life. *Proceedings of the National Academy of Sciences*. 2019;116(28):13751–13758.
45. Goodie AS, Doshi P, Young DL. Levels of theory-of-mind reasoning in competitive games. *Journal of Behavioral Decision Making*. 2012;25(1):95–108.
46. Pinker S. The stuff of thought: Language as a window into human nature. Penguin; 2007.
47. Pinker S, Nowak MA, Lee JJ. The logic of indirect speech. *Proceedings of the National Academy of sciences*. 2008;105(3):833–838.
48. de Weerd H, Verbrugge R, Verheij B. Negotiating with other minds: the role of recursive theory of mind in negotiation with incomplete information. *Autonomous Agents and Multi-Agent Systems*. 2017;31(2):250–287.

49. Verbrugge R, Meijering B, Wierda S, van Rijn H, Taatgen N, et al. Stepwise training supports strategic second-order theory of mind in turn-taking games. *Judgment and decision making*. 2018;13(1):79–98.
50. Curry O, Chesters MJ. ‘Putting Ourselves in the Other Fellow’s Shoes’: The Role of ‘Theory of Mind’in Solving Coordination Problems. *Journal of Cognition and Culture*. 2012;12(1-2):147–159.
51. Devaine M, Hollard G, Daunizeau J. Theory of mind: did evolution fool us? *PloS One*. 2014;9(2):e87619.
52. de Weerd H, Verbrugge R, Verheij B. Higher-order theory of mind in the tacit communication game. *Biologically Inspired Cognitive Architectures*. 2015;11:10–21.
53. Erb B. *Artificial Intelligence & Theory of Mind*. 2016;.
54. Bolander T. Seeing Is Believing: Formalising False-Belief Tasks in Dynamic Epistemic Logic. In: Jaakko Hintikka on Knowledge and Game-Theoretical Semantics. Springer; 2018. p. 207–236.
55. Bard N, Foerster JN, Chandar S, Burch N, Lanctot M, Song HF, et al. The hanabi challenge: A new frontier for ai research. *Artificial Intelligence*. 2020;280:103216.
56. Dissing L, Bolander T. Implementing Theory of Mind on a Robot Using Dynamic Epistemic Logic. In: *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*; 2020.
57. Thomas KA, DeScioli P, Haque OS, Pinker S. The psychology of coordination and common knowledge. *Journal of personality and social psychology*. 2014;107(4):657.
58. Thomas KA, De Freitas J, DeScioli P, Pinker S. Recursive mentalizing and common knowledge in the bystander effect. *Journal of Experimental Psychology: General*. 2016;145(5):621.
59. Lee JJ, Pinker S. Rationales for indirect speech: The theory of the strategic speaker. *Psychological review*. 2010;117(3):785.
60. Thomas KA, DeScioli P, Pinker S. Common knowledge, coordination, and the logic of self-conscious emotions. *Evol Hum Behav*. 2018;39:179–190.
61. De Freitas J, DeScioli P, Thomas KA, Pinker S. Maimonides’ ladder: States of mutual knowledge and the perception of charitability. *Journal of Experimental Psychology: General*. 2019;148(1):158.
62. Becker GS. *The economic approach to human behavior*. University of Chicago press; 1976.
63. Rubinstein A. The Electronic Mail Game: Strategic Behavior Under” Almost Common Knowledge”. *The American Economic Review*. 1989; p. 385–391.
64. van Emde Boas P, Groenendijk J, Stokhof M. The conway paradox: Its solution in an epistemic framework. In: *Proceedings of the third Amsterdam Montague Symposium*; 1980. p. 159–182.
65. van Ditmarsch H, Kooi B. One hundred prisoners and a light bulb. In: *One Hundred Prisoners and a Light Bulb*. Springer; 2015. p. 83–94.

66. Meijering B, van Maanen L, van Rijn H, Verbrugge R. The facilitative effect of context on second-order social reasoning. In: *Proceedings of the Annual Meeting of the Cognitive Science Society*. vol. 32; 2010.
67. Wason PC, Shapiro D. Natural and contrived experience in a reasoning problem. *Quarterly Journal of Experimental Psychology*. 1971;23(1):63–71.
68. Good IJ. Rational decisions. In: *Breakthroughs in statistics*. Springer; 1992. p. 365–377.
69. Seidenfeld T. Calibration, coherence, and scoring rules. *Philosophy of Science*. 1985;52(2):274–294.
70. Palfrey TR, Wang SW. On eliciting beliefs in strategic games. *Journal of Economic Behavior & Organization*. 2009;71(2):98–109.
71. McCutcheon RG. In favor of logarithmic scoring. *Philosophy of Science*. 2019;86(2):286–303.
72. Verbrugge R, Mol L. Learning to apply theory of mind. *Journal of Logic, Language and Information*. 2008;17(4):489–511.
73. Keynes JM. *The General Theory of Employment Terest and Money*. Macmillan and Company; 1936.
74. Halpern JY, Pucella R. Dealing with logical omniscience: Expressiveness and pragmatics. *Artificial intelligence*. 2011;175(1):220–235.
75. Moses Y. Relating knowledge and coordinated action: The knowledge of preconditions principle. In: *TARK 2015*. vol. 215 of EPTCS; 2016. p. 231–245.
76. Chen D, Schonger M, Wickens C. oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*. 2016;9(C):88–97.
77. Shoham Y, Leyton-Brown K. *Multiagent systems: Algorithmic, game-theoretic, and logical foundations*. Cambridge University Press; 2008.
78. Leyton-Brown K, Shoham Y. *Essentials of game theory: A concise multidisciplinary introduction*. Springer Nature; 2022.
79. Aumann RJ. Backward induction and common knowledge of rationality. *Games and Economic Behavior*. 1995;8(1):6–19.