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

Thomas Bolander, Robin Engelhardt, Thomas S. Nicolet

May 20, 2021

Abstract

Common knowledge is a necessary condition for safe group coordination. When common knowledge cannot be obtained, humans rely on their ability to identify shared knowledge to distinguish who knows what. But such shared knowledge attributions are limited in depth and therefore prone to coordination failures, because any finite-order knowledge attribution allows for an even higher order attribution that may change what is known by whom. In three separate experiments we investigate to which degree human participants (N=802) are able to recognize the difference between common knowledge and nth-order shared knowledge. We use a new two-person coordination game with imperfect information that is able to cast the recursive game structure and higher-order uncertainties into a simple, everyday-like setting. In the game, successful coordination to guarantee the highest possible payoff requires common knowledge among the participants. However, such common knowledge is impossible to achieve in the game, only nth-order shared knowledge can be achieved for varying values of n. Our results show that already from quite shallow depths of shared knowledge, the participants act as if they had common knowledge, and express the same certainty in their actions as if they had common knowledge. We call this phenomenon 'the curse of shared knowledge'. It occurs in spite of participants suffering huge payoff penalties when falsely assuming that coordination, and hence common knowledge, is guaranteed.

1 Introduction

Anybody who has been caught up in a sidewalk dance knows that human coordination without common knowledge about each others 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 (Fagin, Halpern, Moses, & Vardi, 1995). 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 (Aumann, 1976; Clark & Marshall, 1981; Lewis, 1969; Schelling, 1980). If the premise of "everyone knows" is not infinitely nested, but only nested to finite depth, we instead have *shared knowledge*. If there is no nested knowledge about knowledge at all and not everyone necessarily knows the fact, we have *private knowledge*.

Let us illustrate the difference between these notions with an example. Two friends, Agnes and Bertram, are taking different trekking routes to the top of a mountain. In the morning they agree that if the weather gets bad, they will go back down and sleep in the mountain hut at the base. Otherwise, they strongly prefer to stay overnight

at the top. The equipment essential for an overnight stay at the top has been divided between their backpacks. It is therefore crucial that if one of them decides to go to the top, the other one does the same. On the way to the top, they both observe a thunderstorm approaching, but are uncertain about whether the other person has seen it. At this point they both know the fact that "a thunderstorm approaches", but don't know whether the other knows. In this situation, we would say that Agnes and Bertram both have *private knowledge* that a thunderstorm approaches, and since they both know it, it is also *shared knowledge* between them. More generally, a fact is *private knowledge* in a group of agents if some non-empty subset of the agents know the fact. For the special case where everybody in the group knows it, we say that there is *shared knowledge to depth one* (or *first-order shared knowledge*) of the fact (Clark & Marshall, 1981).

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

How many messages back and forth does it take for Agnes and Bertram to coordinate going back to the base in the evening? At first it might seem that it is sufficient for both of them to know that at least one of them plans to go to the base. However, that is not so. After the first message has been received, both know that Bertram plans to go to the base, but he is still uncertain whether she knows. And if she doesn't, he might risk leaving her alone at the top. So shared knowledge to depth one is clearly not sufficient. Even shared knowledge to depth two is not sufficient. After she has confirmed receiving his original message, from her perspective it is still entirely possible that he doesn't know that she knows, and hence, he might decide to go to the top to not leave her alone. And in this case, she also has to go to the top in order not to leave *him* alone. This argument can be generalised to prove that even nth-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 nth-order knowledge for *all* $n \in \mathbb{N}$. Shared nth-order knowledge of some fact for all $n \in \mathbb{N}$ is called *common knowledge* of the fact. For a

---

[1]To derive a contradiction, suppose nth-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 nth 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_0th 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_0th 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_0th 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_0th message stil 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 (Fagin et al., 1995).

more formal definition of private, shared and common knowledge, see the Supplementary Material.[2]

In practice, shouldn't it be possible to coordinate meeting at the base after one or two messages being sent back and forth? Isn't it a purely mathematical problem with no practical implications for humans trying to coordinate their actions? One of the main goals of this paper is to argue that the answer to both questions is no. In order to make that argument, we have designed a coordination game that has a similar underlying mathematical structure as the example just given, but cast in a simpler, more everyday-like setting, where the higher-order uncertainty is established already when the game is initialised and not, as above, through a series of message passings. We originally developed the game to illustrate how the difference between shared and common knowledge can have a real impact on human behaviour. Humans generally find the concept of common knowledge hard to grasp, and even harder to grasp the practical relevance of, due to the unbounded nesting of knowledge. The game we developed intends to make 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 (van de Pol, van Rooij, & Szymanik, 2018). 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 (Bradbury, Vehrencamp, et al., 1998; Clark, 1996; Milgrom, 1981). Humans may obtain common knowledge via mutually accessible first-order sensory experiences (Bolander et al., 2015; Gintis, 2010; Lorini, Tummolini, & Herzig, 2005; Tomasello, 1995), eye contact (Friedell, 1969), public rituals and conventions (Lewis, 1969), or salient focal points (Schelling, 1957). 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 (Harari, 2014; Tooby & Cosmides, 2010). 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) (Premack & Woodruff, 1978), "mentalizing" (Frith & Frith, 2003), "mind reading" (I. Apperly, 2010; Vogeley et al., 2001), "mental models" (Johnson-Laird, 1983), "mind perception" (Gray, Jenkins, Heberlein, & Wegner, 2011), "perspective taking", and "social intelligence" (Baron-Cohen et al., 1999), which often are used interchangeably for studying the cognitive mechanisms of shared knowledge, but sometimes focus on slightly different ideas and associated meanings (Schaafsma, Pfaff, Spunt, & Adolphs, 2015), depending on the field of investigation.

Looking at the ToM-literature, conclusions about human belief reasoning abilities are rather heterogeneous (I. A. Apperly & Butterfill, 2009; Saxe & Young, 2013): Human reasoning about the reasoning processes of other humans is limited (Gopnik & Astington, 1988; Hedden & Zhang, 2002; Keysar, Lin, & Barr, 2003; Nagel, 1995; Pinker, 2003; Premack & Woodruff, 1978; Stahl & Wilson, 1995), contextual, and possibly domain specific (Heyes, 2014; Leslie & Thaiss, 1992; Saxe, 2009). Three-year-old children tend to fail in the well-known first-order false belief

---

[2]Supplementary Material is accessible at `http://www2.compute.dtu.dk/~tobo/cd_sm.pdf`.

tasks by falsely assuming that their private information is shared by others, while second-order false belief tasks are mastered around ages 5-7 (Perner & Wimmer, 1985; Sullivan, Zaitchik, & Tager-Flusberg, 1994). Adults may reliably master up to four orders (Kinderman, Dunbar, & Bentall, 1998), 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 (Birch & Bloom, 2007; Camerer, Loewenstein, & Weber, 1989). 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 along[3] 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 (De Freitas, Thomas, DeScioli, & Pinker, 2019). When humans compete or try to detect cheaters, higher-order belief reasoning seems to perform better than lower-order belief reasoning (Goodie, Doshi, & Young, 2012). In negotiations and other mixed motive situations, where innuendo, threads, bribes and other kinds of indirect propositions are common, humans are very good at the *strategic* use of higher-order belief reasoning, for instance as a means to *prevent* common knowledge in certain groups of agents, or as a means to form specific knowledge alliances (de Weerd, Verbrugge, & Verheij, 2017; Pinker, 2007; Pinker, Nowak, & Lee, 2008). ToM proficiency may also be facilitated by providing games with stepwise increase in ToM (Verbrugge et al., 2018).

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 (Curry & Chesters, 2012) suggests that higher-order ToM reasoning may improve coordination efforts, other work seems to suggest that coordination favours lower orders of ToM sophistication (Devaine, Hollard, & Daunizeau, 2014; de Weerd, Verbrugge, & Verheij, 2015). 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 (Bard et al., 2020; Bolander, 2018; Dissing & Bolander, 2020; Erb, 2016).

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 (De Freitas, Thomas, et al., 2019). Controlled pure coordination experiments in social settings on market collaboration (Thomas, DeScioli, Haque, & Pinker, 2014), the bystander effect (Thomas, De Freitas, DeScioli, & Pinker, 2016), indirect speech (Lee & Pinker, 2010), self-conscious emotions (Thomas, DeScioli, & Pinker, 2018), and charity (De Freitas, DeScioli, Thomas, & Pinker, 2019), consistently find that people indeed make strategically different choices under common knowledge conditions (presented in the form of public anouncements), compared to situations in which there is only private knowledge (in the form of private messages) or shared knowledge (private messages that elaborate on the depth of knowledge of other participants). Apart from seeing a clear benefit of common

---

[3]Especially 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'. (Academian, 2018).

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 (Thomas et al., 2014), which is in contrast to the assumptions of standard rational choice theory in which payoffs are expected to be maximized (Becker, 1976).

So 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 the two states of 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 the wild, in particular that they are prone to mistake nth-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 nth-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 (1989), a game version of the mountain trekking example presented above (and of the structurally equivalent 'coordinated attack problem' by Fagin et al. (1995)). The Rubinstein paper shows that 'almost common knowledge' in the sense of nth-order shared knowledge for some large n, leads to a very different expected player behaviour than 'absolute common knowledge'. Essentially his conclusion, translated into the context of the mountain trekking example, is that common knowledge will make the two mountain hikers both go to the base, whereas if there is only nth-order shared knowledge for some n, then both will meet at the top, independent of n and despite the bad weather condition (resulting in non-maximal payoffs). Rubinstein does a pure game-theoretic analysis of the game with no experiments, and only speculates what people playing the game might do. However, it is important to understand how humans in fact reason in such a situation as it may provide profound insights into human abilities to coordinate without communication. What does human intuition recommend? Which depth of shared knowledge (if any) would be enough to attempt risky coordination that would lead to maximal payoff if successful? How would humans be certain that the person they try to coordinate with thinks that the same depth is sufficient?

The electronic mail game and the mountain trekking example are complicated in terms of the dynamics of iterated message passing. In this paper, we devise a novel game in which the higher orders of shared knowledge are not achieved dynamically via actions, but are already present at the beginning of the game, using uncertainty about arrival times. This, we believe, makes the game easier to understand. Letting humans play our game, we have been able to address the previous questions in more detail. Our results 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. The reason, we believe, is that the sole presence of shared knowledge is enough to make participants try to coordinate, and that moderate depths of shared knowledge become effectively indistinguishable from common knowledge due to the recursive nature of the game. We call this effect "the curse of shared knowledge" because even small depths of shared knowledge raises the participant's expectation of being able to coordinate in spite of repeated payoff penalties for having miscoordinated before.

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. van Ditmarsch and Kooi (2015); van Emde Boas, Groenendijk, and Stokhof (1980). 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 (Meijering, van Maanen, van Rijn, & Verbrugge, 2010; Wason & Shapiro, 1971). The introductory story of the game goes as follows:

> "Every morning you arrive at work between 8:10 am and 9:10 am. You and your colleague will arrive
> by bus 10 minutes apart. Example: You arrive at **8:40 am**. Your colleague may arrive at **8:30 am**, or
> **8:50 am**. Both of you like to meet in the canteen for a cup of coffee. If you arrive before 9:00 am, you
> have time to go to the canteen, but you should only go if your colleague goes to the canteen as well.
> If you or your colleague arrive at 9:00 am or after, you should go straight to your offices."

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

    The game has three possible outcomes: 1) both choose the canteen which we refer to as *coordination into the canteen*; 2) both choose their respective offices which we refer to as *coordination into the offices*; 3) one chooses the canteen and the other chooses the office which we refer to as *miscoordination*. Penalties are tiered in such a way that a small penalty is deducted for successful coordination into the canteen (achieving the highest payoff), which is doubled for coordination into the offices (achieving the second-highest payoff), while the penalty for miscoordination or forbidden choices, i.e. going to the canteen at 9 am or after, is much larger (up to 921 times larger, meaning a significantly lower payoff than the previous two). See Appendix B, Payoffs and Penalties, for details about the payoff structure. In the instructions shown to the participants beforehand, we also include payoff examples of both successful and failed coordinations. Screenshots and full descriptions of the experimental setup can be found in the Supplementary Material.

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

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 getter a better payoff than if both choose office. Now consider a case where you are one of the players, and you arrive at 8:50. Then your colleague will be arriving at either 9:00 or 8:40. If your colleague arrives at 9:00, she has to choose office according to the previous argument, and then you would have to choose office as well to avoid the large penalty of miscoordination. However, if your colleague arrives at 8:40, you may both choose the canteen, and this will lead to the highest payoff. In other words, depending on the arrival time of your colleague, a piece of information that you don't have access to, the best choice is either office or canteen. So which one to choose?

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

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

exist at least one pair of arrival times for which the two players are miscoordinated.[4]

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

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

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

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

---

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

Table 1:

Experiment	N*	pairs*	R*	$\bar{r}^*$	Ruin* (%)	Payoff* (%)	$\bar{s}^*$ ($)
MTurk	680	340	10	6.3	52.8	23.6	-1.59
DTU1	80	40	30	27.0	17.5	27.0	-0.83
DTU2	42	21	30	24.1	31.0	24.1	-0.98

*N = number of subjects; pairs = number of pairs; R = maximum number of rounds; $\bar{r}$ = average number of rounds played; 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.
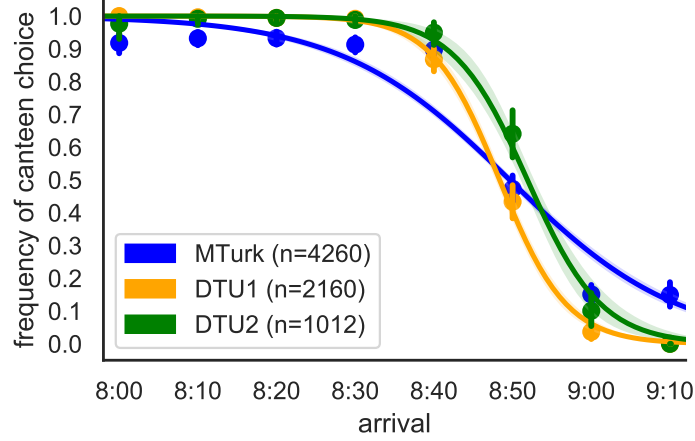

Figure 1: Frequency of canteen choices as a function of arrival times. Circles indicate the mean frequency of participants choosing the canteen at a certain arrival time with error bars. Colored lines are logistic regression lines with bootstrapped 95% confidence intervals (10.000 resamples) shown as translucent bands. Fitted parameters show significant differences for all three experiments.

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. This may come as no surprise, because these students were later into their studies, and had already been taught about social cognition.

Looking at Fig. 1, we see the frequency of participants' canteen choices as a function of their arrival time together with a fitted binary logistic regression line. While DTU students (orange and green) show similar steep profiles, MTurk participants have a slightly more gradual decline in canteen choices for increasing arrival times. However, the point at which there is a 50% probability of choosing the canteen or the office is close to 8:50 in all three experiments (see Appendix A.1). In Section 6, Discussion, we therefore combine all three experiments (Fig. 5) in order to understand the experimental results in terms of degrees of shared knowledge.

Fig. 2 shows the distribution of certainty estimates for each arrival time. It clearly shows that it is exceedingly rare for any of the participants to consider it problematic to go to the canteen when arriving early. Arriving at 8:30
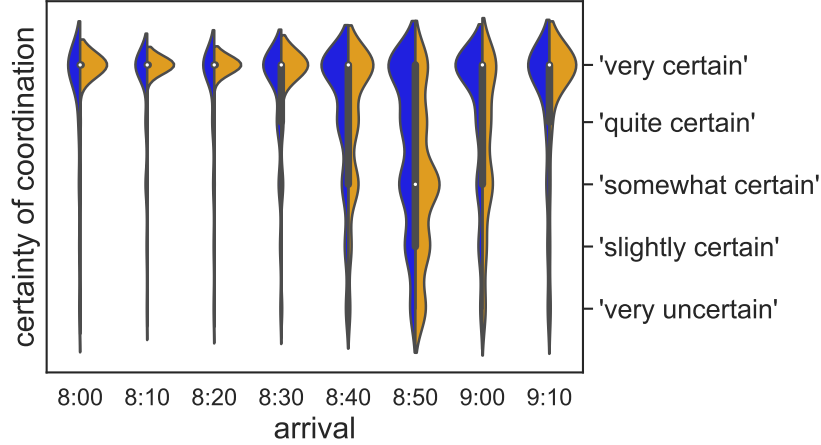
Figure 2: Violin plots of certainty estimates. In each round, participants were asked how certain they were of successful coordination with their colleague. Blue areas show the results from MTurk ($n = 4260$) and orange areas show the results from DTU1 and DTU2 combined ($n = 3172$). We predefined a 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).

or earlier is deemed sufficiently early to visit the canteen with very high confidence. The difference in certainty estimates between MTurk participants and DTU students show that the latter tend to be more certain that their co-players follow a similar strategy (higher certainty estimates for the early and late arrival times), and also that they are more aware of the danger of miscoordination (lower certainty estimates around the cut-off). This is in particular the case for the DTU1 experiment that has the steepest profile. Being more certain that your co-players follow a similar strategy probably indicates that you believe such a strategy to be optimal. So, interestingly, the DTU1 participants are both the ones that appear to be most aware of the danger of miscoordination, and at the same time those who most firmly believe a cut-off strategy is optimal, i.e., believing that the risk of miscoordination is unavoidable. The differences between the three experiments are however still relatively minor, and in the following we will combine data from all three experiments.

In Fig. 3, we see the number of successful pair coordinations into the canteen/office (green/purple) together with the number of miscoordinations (red) as a function of all possible arrival time combinations. The figure shows clearly that players are able to coordinate into the canteen more than 80% of the time if both of them arrive before 8:50. As soon as a 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 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
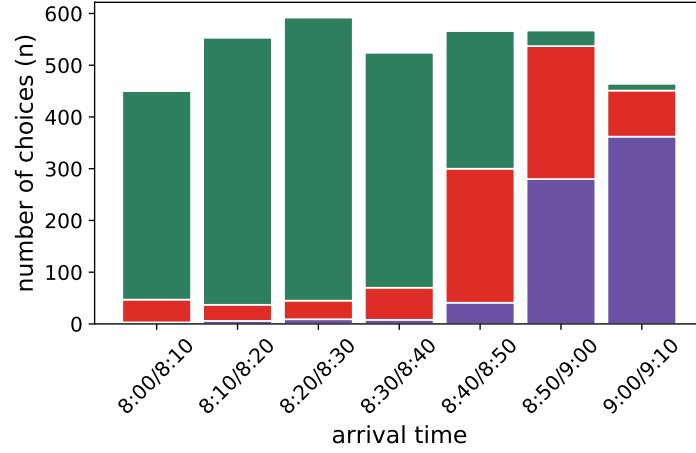
10

Figure 3: Number 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.
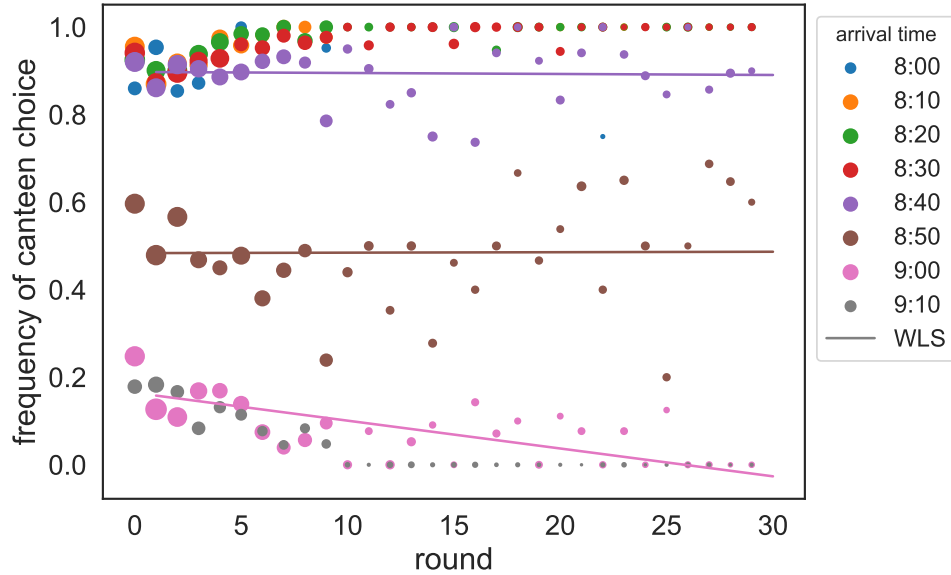

Figure 4: Mean frequencies of canteen choices for all possible arrival times as a function of the number of rounds played. The fitted straight lines are weighted linear squares (WLS) with the weights chosen to be the square root of the number of data points constituting the mean frequencies for each round, also shown by dot size.
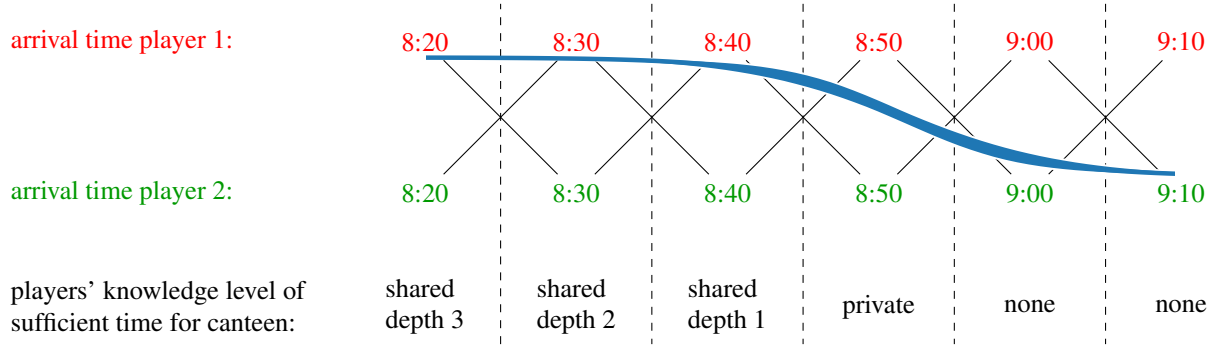
| arrival time player 1: | 8:20 | 8:30 | 8:40 | 8:50 | 9:00 | 9:10 |

| arrival time player 2: | 8:20 | 8:30 | 8:40 | 8:50 | 9:00 | 9:10 |

| players' knowledge level of sufficient time for canteen: | shared depth 3 | shared depth 2 | shared depth 1 | private | none | none |

Figure 5: 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 concerning whether there is sufficient time to go to the canteen, e.g. when arriving at $8:40$ there is shared knowledge to depth 1 of this fact, but not shared knowledge to depth 2. In blue, a binary logistic regression model was used to predict the probability of a participant going to the canteen (upper limit) or to the office (lower limit) at the shown arrival times. The width of the regression line indicates the 95% confidence interval using 10.000 bootstrapped resamples of all choices in all three experiments ($n = 7432$).

and pairs having had one or more miscoordinations (see the supplementary data analysis in the Supplementary Material), shows somewhat decreasing certainty estimates around the critical arrival times, especially for DTU students. 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 the Supplementary Material), but this does not translate into better payoffs as later miscoordinations just move to earlier arrival times. So even though participants learn that their choices are risky, they don't see any way to improve their strategy. Specifically, they never converge to the optimal all-office strategy, and also not to the alternative canteen-before-9 strategy (cf. Theorem 4 in Appendix C). This apparent lack of behavioral change in higher-order social reasoning games is also shown by Verbrugge and Mol (2008).

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 am. With the aim of achieving the highest possible payoff, each participant can be expected to consider her own arrival time and try to assess whether there is still time to meet in the canteen. When a participant arrives strictly before 9 am, i.e. at $8:50$ or earlier, she has private knowledge that she arrives sufficiently early to go to the canteen. If participants only make choices based on their private knowledge, we should then expect participants to always go to the canteen at $8:50$. This is not what we see, cf. 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 am, 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, again, an arrival pair (t_1, t_2) denotes that player 1 arrives at time t_1 and player 2 at time t_2). Note that for an arrival pair $(8:50, 8:40)$, there is shared knowledge of there being sufficient time to go to the canteen, but only player 2 knows this fact: Player 2 knows that also player 1 must have arrived before $9:00$, but player 1 doesn't know this about player 2.

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

The participants seem to clearly be able to distinguish between private and shared knowledge, which is supported by their significantly different choices at 8:50 and 8:40 (see again Fig. 5). However, it is less clear whether they are able to robustly distinguish different levels of shared knowledge, and whether they are able to distinguish that from common knowledge. Indeed, most participants relatively robustly choose the canteen at 8:40 and any time before that, despite the difference in depth of shared knowledge in those possible arrival times. The certainty estimates are however slowly decreasing from 8:10 to 8:50 in all three experiments (see Fig. 2), showing that the participants are not completely ignorant to the differences. This could suggest that many participants are aware that it is less safe to go to the canteen based on nth-order shared knowledge than $(n + 1)$st-order shared knowledge. However, very few seem to draw the conclusion that it is never safe to go to the canteen. Our game theoretic analysis showed that they ought to only choose the canteen when there is common knowledge that it is safe, which in this case actually means never.

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

Probing theses 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. 6A show that most answers range from 8:30 to 8:50 (approximately 75% of all answers), giving support to the verdict that participants are not able to continue taking the perspective of each other recursively, or at least that they believe that shared knowledge of some modest finite depth is sufficient for the canteen choice to be safe. Rather, they stop after one or two, possibly three, iterations, thus believing that as long as they arrive
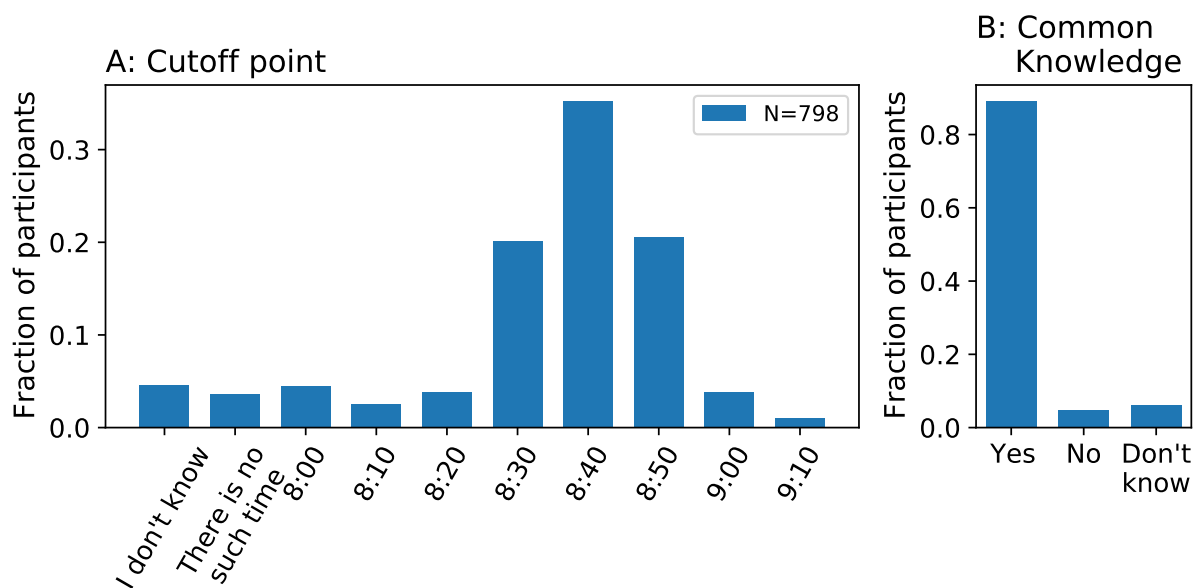
Figure 6: A) Frequencies of answers to the question: "Imagine you could have agreed beforehand with your colleague about a point in time where it is safe to go to the canteen. What time would that be?" Due to the pragmatics of language, we assume that an answer like 8:30 entails the belief that all earlier arrival times would also be deemed safe. B) Frequencies of answers to the question: "Imagine you arrive at 8:10. Is it common knowledge between you and your colleague that it is safe to go to the canteen, that is, that you both arrived before 9:00?"

sufficiently early, they can be sure to coordinate safely in the canteen. Notice that the correct answer "there is no such time" is chosen by less than 4% of all participants, close to the margin of random error.

This indicates that participants indeed do believe that there exists a strategy that includes canteen choices without the risk of miscoordination, although in reality they would need common knowledge in order for the canteen choice to be without risk. Participants might of course not necessarily have a precise idea of the technical notion of common knowledge and how it differs from nth-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. In our experiments, we see the player behavior stabilizing already at relatively modest depths of shared knowledge, both in terms of action choices and certainty estimates. And that player behavior does not match what we would expect to see if they only believed to have nth-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. 6B, the results show that 89% of all participants responded that it was common knowledge that both players arrived before 9:00 when they themselves had arrived at 8:10. The answers may signify that indeed they believe there to be common knowledge in the strict technical (logical) sense. But of course

14

the answers could also pertain to the everyday linguistic usage of the term 'common knowledge', which is less strict.

7 Conclusion

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

Our experience from playing the Canteen Dilemma with many people and explaining to them its unintuitive result, is that many players simply do not accept the argument that there is no time early enough for them to coordinate safely into the canteen. On top of this, the certainty of participants that they will coordinate into the canteen at early arrival times indicates that when participants fail at nth-order reasoning, they do not default to agnosticism, but the opposite. That is, when there is a sufficiently large depth of shared knowledge about a fact, it is possible that such a fact is mistaken for proper common knowledge. An interesting avenue for future research would be to investigate if there may be any social and psychological benefits of having an *illusion of common knowledge*, such as a higher willingness to cooperate. Thus, what we have called the 'curse' of shared knowledge in the Canteen Dilemma, may turn out to be a blessing in other settings.

An obvious question is how often this illusion of common knowledge occurs in real life. For instance, in the real-world version of the Canteen Dilemma scenario, the two colleagues would be likely to simply coordinate their actions via cell phone ("I'll arrive at 8:50 today. Are you up for a cup of coffee in the canteen?"). This suggests that the advent of modern technology could have made the information asymmetry inherent in shared knowledge situations less widespread. However, modern technology has also given us the Internet and social media, where the flow of information is much more complex, creating more intricate cases of information asymmetry than ever before.

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

References

Academian. (2018, August). *Unrolling social metacognition: Three levels of meta are not enough.* Lesswrong,. Retrieved from `https://www.lesswrong.com/posts/K4eDzqS2rbcBDsCLZ/unrolling-social-metacognition-three-levels-of-meta-are-not`

Apperly, I. (2010). *Mindreaders: the cognitive basis of" theory of mind"*. Psychology Press.

Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological review*, *116*(4), 953.

Aumann, R. J. (1976). Agreeing to disagree. *The annals of statistics*, 1236–1239.

Bard, N., Foerster, J. N., Chandar, S., Burch, N., Lanctot, M., Song, H. F., . . . others (2020). The hanabi challenge: A new frontier for ai research. *Artificial Intelligence*, *280*, 103216.

Baron-Cohen, S., Ring, H. A., Wheelwright, S., Bullmore, E. T., Brammer, M. J., Simmons, A., & Williams, S. C. (1999). Social intelligence in the normal and autistic brain: an fmri study. *European journal of neuroscience*, *11*(6), 1891–1898.

Becker, G. S. (1976). *The economic approach to human behavior*. University of Chicago press.

Birch, S. A., & Bloom, P. (2007). The curse of knowledge in reasoning about false beliefs. *Psychological Science*, *18*(5), 382–386.

Bolander, T. (2018). Seeing is believing: Formalising false-belief tasks in dynamic epistemic logic. In *Jaakko hintikka on knowledge and game-theoretical semantics* (pp. 207–236). Springer.

Bolander, T., van Ditmarsch, H., Herzig, A., Lorini, E., Pardo, P., & Schwarzentruber, F. (2015). Announcements to attentive agents. *Journal of Logic, Language and Information*, 1–35. doi: 10.1007/s10849-015-9234-3

Bradbury, J. W., Vehrencamp, S. L., et al. (1998). Principles of animal communication.

Camerer, C., Loewenstein, G., & Weber, M. (1989). The curse of knowledge in economic settings: An experimental analysis. *Journal of political Economy*, *97*(5), 1232–1254.

Chen, D., Schonger, M., & Wickens, C. (2016). otree—an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, *9*(C), 88–97.

Clark, H. H. (1996). *Using language*. Cambridge university press.

Clark, H. H., & Marshall, C. R. (1981). Definite knowledge and mutual knowledge.

Curry, O., & Chesters, M. J. (2012). 'putting ourselves in the other fellow's shoes': The role of 'theory of mind'in solving coordination problems. *Journal of Cognition and Culture*, *12*(1-2), 147–159.

De Freitas, J., DeScioli, P., Thomas, K. A., & Pinker, S. (2019). Maimonides' ladder: States of mutual knowledge and the perception of charitability. *Journal of Experimental Psychology: General*, *148*(1), 158.

De Freitas, J., Thomas, K., DeScioli, P., & Pinker, S. (2019). Common knowledge, coordination, and strategic mentalizing in human social life. *Proceedings of the National Academy of Sciences*, *116*(28), 13751–13758.

Devaine, M., Hollard, G., & Daunizeau, J. (2014). Theory of mind: did evolution fool us? *PloS One*, *9*(2), e87619.

de Weerd, H., Verbrugge, R., & Verheij, B. (2015). Higher-order theory of mind in the tacit communication game. *Biologically Inspired Cognitive Architectures*, *11*, 10–21.

de Weerd, H., Verbrugge, R., & Verheij, B. (2017). Negotiating with other minds: the role of recursive theory of mind in negotiation with incomplete information. *Autonomous Agents and Multi-Agent Systems*, *31*(2), 250–287.

Dissing, L., & Bolander, T. (2020). Implementing theory of mind on a robot using dynamic epistemic logic. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*.

Erb, B. (2016). Artificial intelligence & theory of mind.

Fagin, R., Halpern, J. Y., Moses, Y., & Vardi, M. Y. (1995). *Reasoning about knowledge*. MIT Press.

Friedell, M. F. (1969). On the structure of shared awareness. *Behavioral Science*, *14*(1), 28–39.

Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *358*(1431), 459–473.

Gintis, H. (2010). Rationality and common knowledge. *Rationality and Society*, *22*(3), 259–282.

Goodie, A. S., Doshi, P., & Young, D. L. (2012). Levels of theory-of-mind reasoning in competitive games. *Journal of Behavioral Decision Making*, *25*(1), 95–108.

Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child development*, 26–37.

Gray, K., Jenkins, A. C., Heberlein, A. S., & Wegner, D. M. (2011). Distortions of mind perception in psychopathology. *Proceedings of the National Academy of Sciences*, *108*(2), 477–479.

Harari, Y. N. (2014). *Sapiens: A brief history of humankind*. Random House.

Hedden, T., & Zhang, J. (2002). What do you think i think you think?: Strategic reasoning in matrix games. *Cognition*, *85*(1), 1–36.

Heyes, C. (2014). Submentalizing: I am not really reading your mind. *Perspectives on Psychological Science*, *9*(2), 131–143.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness* (No. 6). Harvard University Press.

Keysar, B., Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition*, *89*(1), 25–41.

Kinderman, P., Dunbar, R., & Bentall, R. P. (1998). Theory-of-mind deficits and causal attributions. *British Journal of Psychology*, *89*(2), 191–204.

Lee, J. J., & Pinker, S. (2010). Rationales for indirect speech: The theory of the strategic speaker. *Psychological review*, *117*(3), 785.

Leslie, A. M., & Thaiss, L. (1992). Domain specificity in conceptual development: Neuropsychological evidence from autism. *Cognition*, *43*(3), 225–251.

Lewis, D. (1969). *Convention: A philosophical study*.

Lorini, E., Tummolini, L., & Herzig, A. (2005). Establishing mutual beliefs by joint attention: towards a formal model of public events. In *Proc. of cogsci* (pp. 1325–1330).

Meijering, B., van Maanen, L., van Rijn, H., & Verbrugge, R. (2010). The facilitative effect of context on second-order social reasoning. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 32).

Milgrom, P. (1981). An axiomatic characterization of common knowledge. *Econometrica: Journal of the Econometric Society*, 219–222.

Nagel, R. (1995). Unraveling in guessing games: An experimental study. *The American Economic Review*, *85*(5), 1313–1326. Retrieved from `http://www.jstor.org/stable/2950991`

Palfrey, T. R., & Wang, S. W. (2009). On eliciting beliefs in strategic games. *Journal of Economic Behavior & Organization*, *71*(2), 98–109.

Perner, J., & Wimmer, H. (1985). "john thinks that mary thinks that. . . " attribution of second-order beliefs by 5-to 10-year-old children. *Journal of experimental child psychology*, *39*(3), 437–471.

Pinker, S. (2003). *The language instinct: How the mind creates language*. Penguin UK.

Pinker, S. (2007). *The stuff of thought: Language as a window into human nature*. Penguin.

Pinker, S., Nowak, M. A., & Lee, J. J. (2008). The logic of indirect speech. *Proceedings of the National Academy of sciences*, *105*(3), 833–838.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and brain sciences*, *1*(4), 515–526.

Rubinstein, A. (1989). The electronic mail game: Strategic behavior under" almost common knowledge". *The American Economic Review*, 385–391.

Saxe, R. (2009). Theory of mind (neural basis). *Encyclopedia of consciousness*, *2*, 401–410.

Saxe, R., & Young, L. (2013). Theory of mind: How brains think about thoughts. *The Oxford handbook of cognitive neuroscience*, *2*, 204–213.

Schaafsma, S. M., Pfaff, D. W., Spunt, R. P., & Adolphs, R. (2015). Deconstructing and reconstructing theory of mind. *Trends in cognitive sciences*, *19*(2), 65–72.

Schelling, T. C. (1957). Bargaining, communication, and limited war. *Conflict Resolution*, *1*(1), 19–36.

Schelling, T. C. (1980). *The strategy of conflict*. Harvard university press.

Seidenfeld, T. (1985). Calibration, coherence, and scoring rules. *Philosophy of Science*, *52*(2), 274–294.

Shoham, Y., & Leyton-Brown, K. (2008). *Multiagent systems: Algorithmic, game-theoretic, and logical foundations*. Cambridge University Press.

Stahl, D. O., & Wilson, P. W. (1995). On players' models of other players: Theory and experimental evidence. *Games and Economic Behavior*, *10*(1), 218–254.

Sullivan, K., Zaitchik, D., & Tager-Flusberg, H. (1994). Preschoolers can attribute second-order beliefs. *Developmental Psychology*, *30*(3), 395.

Thomas, K. A., De Freitas, J., DeScioli, P., & Pinker, S. (2016). Recursive mentalizing and common knowledge in the bystander effect. *Journal of Experimental Psychology: General*, *145*(5), 621.

Thomas, K. A., DeScioli, P., Haque, O. S., & Pinker, S. (2014). The psychology of coordination and common knowledge. *Journal of personality and social psychology*, *107*(4), 657.

Thomas, K. A., DeScioli, P., & Pinker, S. (2018). Common knowledge, coordination, and the logic of self-conscious emotions. *Evol Hum Behav*, *39*, 179–190.

Tomasello, M. (1995). Joint attention as social cognition. *Joint attention: Its origins and role in development*, 103–130.

Tooby, J., & Cosmides, L. (2010). Groups in mind: The coalitional roots of war and morality. *Human morality and sociality: Evolutionary and comparative perspectives*, 91–234.

van de Pol, I., van Rooij, I., & Szymanik, J. (2018). Parameterized complexity of theory of mind reasoning in dynamic epistemic logic. *Journal of Logic, Language and Information*, *27*(3), 255–294.

van Ditmarsch, H., & Kooi, B. (2015). One hundred prisoners and a light bulb. In *One hundred prisoners and a light bulb* (pp. 83–94). Springer.

van Emde Boas, P., Groenendijk, J., & Stokhof, M. (1980). The conway paradox: Its solution in an epistemic framework. In *Proceedings of the third amsterdam montague symposion* (pp. 159–182).

Verbrugge, R., Meijering, B., Wierda, S., van Rijn, H., Taatgen, N., et al. (2018). Stepwise training supports strategic second-order theory of mind in turn-taking games. *Judgment and decision making*, *13*(1), 79–98.

Verbrugge, R., & Mol, L. (2008). Learning to apply theory of mind. *Journal of Logic, Language and Information*, *17*(4), 489–511.

Vogeley, K., Bussfeld, P., Newen, A., Herrmann, S., Happé, F., Falkai, P., . . . Zilles, K. (2001). Mind reading: neural mechanisms of theory of mind and self-perspective. *Neuroimage*, *14*(1), 170–181.

Wason, P. C., & Shapiro, D. (1971). Natural and contrived experience in a reasoning problem. *Quarterly Journal*

of Experimental Psychology, 23(1), 63–71.

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 who would meet in the office all the time. This turned out not to be the case. Experiments on Amazon Mechanical Turk had a total of 714 participants (including dropouts, see Supplementary Material), while the two classroom experiments at the Technical University of Denmark (DTU1 and DTU2) had a total of 106 and 50 participants, respectively[5]

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 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 (Chen, Schonger, & Wickens, 2016).

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

A.1   Logistic regression

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

---

[5]Complete anonymized data files and all code can be downloaded from `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 am or later). Office coordinations are designed to have larger penalties than canteen coordinations, but smaller penalties than miscoordinations in order to make sure that coordination remains the main objective of the game. Penalties are defined as negative utility values in the following way. First, we define the chosen action a_i by player i, $i = 1, 2$, to take binary values encoding the canteen option ($a_i = 0$) and the office option ($a_i = 1$), and define their respective certainty estimates $e_i \in \{0.5, 0.625, 0.75, 0.875, 0.99\}$. We can then express the utility u received by player 1 as $u(e_1, a_1, a_2) = (1 - |a_1 - a_2| + a_1 a_2) ln(e_1) + 2|a_1 - a_2| ln(1 - e_1)$, and symmetric for player 2. If any of the players choose the canteen at 9 am or after, the utility becomes $u(e_1, a_1, a_2) = 2ln(1 - e_1)$ 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 $u(e_1, a_1, a_2) = 2ln(1 - e_1) = -2.77$. If she goes to the office just like her colleague, her utility is $u(e_1, a_1, a_2) = 2ln(e_1) = -0.58$. It should be noted that the logarithmic scoring rule used here is not strictly proper since office and canteen coordinations are penalized differently. Nevertheless, we find a good match between estimates and actual choices at arrival times different from those that are prone to miscoordinations, as seen in Fig. 2, indicating that loss minimization remained a central concern and that participants made their choices and estimates as honestly as possible (Palfrey & Wang, 2009; Seidenfeld, 1985).

C   Formal Analysis

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

The game starts by nature choosing an element $\mathbf{t} \in T$. Nature is not a strategic player, so we assume that $\mathbf{t}$ is chosen uniformly at random, which is exactly how $\mathbf{t}$ is chosen in our experiments. The participants do not know

that the arrival times are chosen uniformly at random, as this is left implicit in the description of the game. The following analysis of optimal strategies in the game could potentially change if arrival times were chosen according to a highly skewed probability distribution.

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

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

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

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

Note that the defined strategies are memoryless (Markov strategies), that is, a player's choice only depends on the observed arrival time in the current round, not the history of arrival times and chosen actions in earlier rounds. Human players playing the game should not be expected to necessarily play memoryless strategies, as they might seek to 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.

Given strategies s_1 and s_2, the pair $\mathbf{s} = (s_1, s_2)$ is called a *strategy profile*. Given an arrival pair $\mathbf{t} = (t_1, t_2)$, we use $\mathbf{s}(\mathbf{t})$ as a shorthand for $(s_1(t_1), s_2(t_2))$. Hence $\mathbf{s}(\mathbf{t})$ denotes the choices made by players 1 and 2 when their strategies are given by $\mathbf{s}$ and their arrival times are given by $\mathbf{t}$. The payoff of those choices is then $u_{\mathbf{t}}(\mathbf{s}(\mathbf{t}))$. We will use $u_{\mathbf{t}}(\mathbf{s})$ as an abbreviation of $u_{\mathbf{t}}(\mathbf{s}(\mathbf{t}))$, i.e., $u_{\mathbf{t}}(\mathbf{s})$ is the utility received by player 1 and 2 when they play by strategy profile $\mathbf{s}$ in the game with arrival pair $\mathbf{t}$. The *expected utility* $EU(\mathbf{s})$ of a strategy profile $\mathbf{s}$ is the average of the payoffs (Shoham & Leyton-Brown, 2008):

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

Note again that player 1 and 2 get the same payoff (common-payoff game), so there is only one expected utility value to be computed. A strategy profile $\mathbf{s}'$ *Pareto dominates* another strategy profile $\mathbf{s}$ if $EU(\mathbf{s}') > EU(\mathbf{s})$ (Shoham & Leyton-Brown, 2008). A strategy profile is *Pareto optimal* if there does not exist another strategy profile dominating it. A strategy profile $\mathbf{s}'$ *weakly Pareto dominates* another strategy profile $\mathbf{s}$ if $EU(\mathbf{s}') \geq EU(\mathbf{s})$. The game is cooperative (between player 1 and 2), so both players should seek to play a Pareto optimal strategy

profile. Also, since it is a common-payoff game, all Pareto optimal strategy profiles have the same expected utility (Shoham & Leyton-Brown, 2008).

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

$$T_1 = \{(8\text{:}10 + 20x, 8\text{:}20 + 20y) \in T \mid y \leq x \leq y + 1\}$$
$$T_2 = \{(8\text{:}20 + 20y, 8\text{:}10 + 20x) \in T \mid y \leq x \leq y + 1\}$$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

necessarily have $t' < 9:00$, since otherwise $\mathbf{s}$ would not be Pareto optimal according to Lemma 1. From this it follows that we for all arrival pairs $(t_1, t_2) \in T_1$ must have

1. If $t_1, t_2 < t'$, the two players coordinate into the canteen, receiving the highest possible payoff.

2. If $t_1 < t' < t_2$ or $t_2 < t' < t_1$, the two players are miscoordinated (one chooses canteen, the other office), receiving the lowest possible payoff.

3. If $t_1, t_2 > t'$, the two players coordinate into their offices, receiving a payoff strictly between the lowest and highest.

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

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 adapt any of these two strategies, but if two perfectly rational players were to play the game, and if they knew they could expect the other player to play perfectly rational as well, of course the optimal strategy would be played. And, as earlier mentioned, in our particular version of the game, the optimal strategy is the all-office strategy.

D Author Contributions and Acknowledgments