Gmail **Robin Engelhardt <robin.engelhardt@gmail.com>**

Re: Action Letter on Cognitive Science Submission 20-373R1

12 messages

Thomas Bolander <tobo@dtu.dk> Wed, May 18, 2022 at 2:28 PM
To: Robin Engelhardt <robin.engelhardt@gmail.com>, Thomas Nicolet <thomas1nicolet@gmail.com>

Kære Robin og Thomas

Jeg har været igennem et længere Corona-forløb hvor jeg i 3 måneder kun kunne arbejde on and off. Det varede fra slutningen af januar til slutningen af april. Derefter skulle jeg så prøve at komme tilbage på sporet og jeg har været nødt til at tage mig af de mest akutte ting først. Derfor har jeg haft syltet nedenstående. Jeg burde selvfølgelig bare have sendt den videre med det samme, men jeg kunne også godt ud fra en hurtig skimning se at det var dårlige nyheder, og så fik jeg det lidt lagt på is. Men anyway, nu er den her så, og jeg prøver også ved samme lejlighed at få læst det hele igennem og kommentere på det selv nedenfor.

> On 24 Mar 2022, at 03.50, Cognitive Science <em@editorialmanager.com> wrote:
>
> 20-373R1
>
> "The Curse of Shared Knowledge: Recursive Belief Reasoning in a Coordination Game with Imperfect Information"
>
> Thomas Bolander; Robin Engelhardt; Thomas Nicolet
>
> Cognitive Science
>
> Dear Mr Bolander,
>
> My sincere apologies that this decision took longer than usual, in part due to the pandemic circumstances.
>
> Thank you for submitting your revised manuscript to *Cognitive Science*. I have received comments from reviewers with expertise in the area and also read your revision carefully myself. Reviewer 1 also served in the previous round, and Reviewer 4 is a new reviewer for this round. The reviewers' comments are appended below. I would like to thank both reviewers for their time and helpful comments.
>
> I will not repeat all reviewer comments; please see their reviews for full details. I will merely highlight here those aspects that motivate my decision to reject this revised submission.
>
> In my previous action letter, I had asked to clarify the ecological validity and wider implications of this work in a way that is convincing also to skeptics, who e.g. question whether or not we (the cognitive science community) can learn much from the kind of toy scenario used in your experiment. Comments by Reviewer 1 make clear that they are unconvinced by the revision and feel it fails to address their main concerns. My own reading of the arguments also leaves me unconvinced, at best (more on this below). You seem to have taken the approach to arguing for ecological validity by listing presumed real-world analogues of the game scenario. This seems unsatisfactory and insufficient. It does not address the question from my previous action letter: "to what extent [are] the results … due to the artificiality of the task instructions?". The issue is not whether or not one can come up with real-world analogues, but whether or not the experimental task itself has ecological validity in the sense that it taps actual, natural cognitive processes that we as cognitive scientists aim to understand (cf. Reviewer 1's comments), and not mere artifacts.

OK, så det er åbenbart virkelig svært at overbevise folk i cognitive science og de har et helt andet forskningsparadigme end jeg er vant til, hvor der gælder nogen regler og normer som jeg ikke er bekendt med. Så er det jo også rigtig svært at gøre dem tilfredse. Hvis de synes spillet er for "tænkt" til at de kan tro på at det må igangsætte nogen former for naturlige kognitive processer, så skal vi måske satse på at formidle til et andet community, hvor den slags er et mindre problem?

> Even if the above were resolvable by further revision (of which I am not convinced), a more crucial problem would remain: Reviewer 4 has identified a substantive shortcoming in your methodology: i.e., the experimental set-up cannot (reasonably be expected to) induce “the necessary common knowledge between the pairs of participants about the game rules and pay-offs” (Reviewer 4). This common knowledge is necessary for the players to be able to solve the puzzle, which invalidates your formal analysis and interpretation of the findings. Given that this methodological limitation undermines the main conclusion and presumed theoretical contribution, I am unable to accept this manuscript for publication.
>
> I know that this decision will be disappointing given the work that you have put into this manuscript. I do hope that the reviewers’ comments will be useful to you in advancing your research, and if you choose to submit the manuscript elsewhere.
>
> If you indeed plan to revise this manuscript and submit elsewhere, besides fixing the methodological flaw and/or changing the conclusions, I feel compelled to urge you to please remove the example about “consent” from your manuscript. I view the casual treatment of this topic as if it were to illustrate “ecological validity” and/or real-world relevance of your paradigm to be inappropriate and potentially triggering for many readers (especially women).

OK, så skulle vi alligevel have droppet det eksempel, åbenbart. Vi var jo også usikre på det, og jeg mener da også vi spurgte Vincent om hans vurdering, men vi tænkte allesammen at det nok var OK. Jeg var selv ret usikker på om vi kunne risikere en stærk reaktion, men blev overbevist om at det var OK. Men så fik vi alligevel den stærke reaktion. Pokkers også!

Jeg tror det må være en følelsesmæssig reaktion, for jeg kan ikke se hvad der ellers skulle være “inappropriate” i det. Er det fordi vi håndterer det “casual”, altså at vi ikke tager problemet seriøst nok, men reducerer det til noget logisk? Jeg har svært ved at se et videnskabeligt problem her overhovedet, så jeg kan ikke se andet end potentielt nogen følelsesmæssige reaktioner på et emne som er ekstremt sensitivt for rigtig mange.

> For a more nuanced treatment of this topic (including the limits of a “consent” based view of sexual negotiation), I would kindly recommend: Kukla, R. (2018). That’s what she said: The language of sexual negotiation. Ethics, 129(1), 70-97.

Så vidt jeg kan vurdere ud fra abstracted, så går forfatteren i kødet på selve ideen om consent. Der kan man sige at vi måske blot antager at consent er en meningsfuld ting, mens at der måske er meget mere kompleksitet. Jeg ved det ikke. Men det er et minefelt, og måske skulle vi ikke have trådt ud i det overhovedet.

> I hope that the critical points of feedback, by reviewers and myself, are helpful to you for improving your methodology and arguments, and moving this work forward.
>
> I would like to take this opportunity to thank you for considering *Cognitive Science* as a possible outlet for your work, and hope that you will continue to do so in the future. We welcome any feedback that you may have about the journal's review process.
>
> Yours sincerely,
>
> Iris van Rooij
>
> Senior Editor
>
> ==========
>
> Reviewer #1: The authors have clarified their perspective in this version of the manuscript, but made few substantive changes to the paper overall. As a result, we are left with the following (if I'm reading this correctly). In theory, without explicit knowledge of when the other party arrived on campus, neither player has the definitive information they need to decide to go to the canteen. Yet, participants do elect to go and do so in a manner such that the earlier they arrive, the more likely they are to go to the canteen. Ultimately, I still feel that this paper is not helping us to understand the underlying psychology of this task. Essentially, the authors have pointed out that there is a paradox that there are situations in which there is an ambiguity that can never be resolved without explicit communication, yet people still navigate those situations.

Hvor siger vi det? Jeg mindes ikke at vi påstår at folk kan navigere bedre i disse situationer end teorien tilsiger, det er jo faktisk omvendt. I virkeligheden er der jo *ikke* ambiguity, man skal bare gå på kontoret, men folk agerer *som om* der er en ambiguity man kan overkomme. Eller også forstår jeg ikke revieweren. Eller begge dele.

> They do not provide much detail to help the reader to understand what people are doing.
>
> For example, one possibility is that people are not playing a logic game, but a probability game. That is, perhaps participants are assuming that there is a 50% of chance their partner arrived earlier or later than they did. One could construct a decision tree for each arrival time with a probability that their partner arrived at each time and a payoff and punishment matrix and determine the optimal choice to make at each arrival time given the potential range of outcomes. Then, it might be possible to compare this model to people's performance. Again, I don't care what kind of model the authors choose, but I find the paper unsatisfying without some work to determine what participants are doing in this task.

Vi kunne selvfølgelig godt have kigget på mixed strategies. I min oprindelige version af spillet undgik jeg at det skulle kunne give mening ved at miskoordinering gjorde at man med det samme tabte hele spillet, således at det var åbenlyst at man aldrig med nogen positiv sandsynlighed skulle vælge en handling som kunne lede til miskoordering (sålænge der eksisterer en anden handling som ikke leder til det, i hvert fald). Problemet er dog at hvad der er rationelt afhænger af ens antagelser om agenttypen af medspilleren. Og hvis man ikke antager at medspilleren kan lave dette ræsonnement, så kan man risikere at det er umuligt at udelukke miskoordinering. Hvis jeg ser det som sandsynligt at min medspiller vil gå i kantinen sålænge klokken blot er før 9.00, så er mit bedste svar på den strategi jo trods alt også at gå i kantinen før 9.00. Så har vi dog igen balladen med at det vi observerer heller ikke passer med denne strategi og antagelse. Så er der flip-a-coin strategien, hvor man 8.50 tilfældigt vælger at gå i kantinen eller på kontoret. Hvad der er det bedste svar på denne strategi kan jeg dog ikke lige regne ud i hovedet.

Men det er klart at jeg har været meget optaget af optimale strategier i spillet under antagelse af common knowledge af perfekt rationalitet, og ikke så meget på hvordan man ville spille hvis man *ikke* antog at modspilleren var perfekt rationel. Så vi ender lidt i samme type problemstilling som p-beauty contest, hvor det ikke bliver et spørgsmål om at gøre det mest rigtige under antagelse af at andre også har ressourcer til at gøre det mest rigtige, men hvor det handler om at gætte på de andres typer, herunder deres ressourcebegrænsethed. Og ja, vi ved jo faktisk dybest set ikke hvad der er foregået i folks hovedet. *Pånær* at vi dog spørger om ting som “har I common knowledge” og “hvis I kunne aftale et tidspunkt på forhånd”. Jeg synes faktisk svarene på de spørgsmål understreger vores pointe ret kraftigt. Hvis vi snakker om at aftale et tidspunkt på forhånd, så handler det vitterligt om hvorvidt der kan eksistere en pure strategi, hvor man også kender modspillerens strategi, som er tilstrækkelig til at “vinde” spillet. Så reviewerens argument falder om ikke andet lidt fra hinanden når man tænker på hvor overvældende folk mener at det findes et cut-off tidspunkt som er sikkert.

> Reviewer #4: The paper concerns an interesting problem about whether people can discern a difference between common knowledge and shared knowledge (up to some level n). Sections 1 and 2 are very readable and interesting and cover the related literature very well.
>
> The main problem in this paper lies in the features of the experiment introduced in Section 3, which is said to be inspired by the 'consecutive numbers riddle' of (Van Ditmarsch and Kooi 2015), but which unfortunately does not preserve some of the vital ingredients of that riddle, namely, common knowledge among the players of the rules of the game and common knowledge of rationality.
>
> Both of these items of common knowledge are relevant in the consecutive numbers problem, in the version cited by the authors. To explain the important difference between the consecutive numbers riddle and similar logic puzzles on the one hand, and the coordination game introduced in the submitted paper on the other hand, let me quote the whole riddle and most importantly the beginning of its explanation (Ditmarsch & Kooi 2015, p. 1):
>
> "Anne and Bill get to hear the following: "Given are two natural numbers. They are consecutive numbers. I am going to whisper one of these numbers to Anne and the other number to Bill." This happens. Anne and Bill now have the following conversation.
> Anne: "I don't know your number."
> Bill: "I don't know your number."
> Anne: "I know your number."
> Bill: "I know your number."
>
> First they don't know the numbers, and then they do. How is that possible? What surely is one of the two numbers?'
>
> The natural numbers are the numbers 0, 1, 2, 3, etc. Numbers are consecutive if they are one apart. It is important for the formulation of the riddle that
> Anne and Bill are simultaneously aware of this scenario, and also know that
> they both are aware of this scenario, etc. Therefore, they are being spoken
> to, instead of, for example, both receiving written instructions. It is therefore

too that the numbers are whispered into their ears—the whispering creates
common knowledge that they have received that information. We can imagine the setting of this riddle as Anne, Bill, and the speaker sitting round a table, such that the speaker has to lean forward to Anne in order to whisper to her, and subsequently has to lean forward to Bill and whisper to him."

This shows the precision with which common knowledge of the important facts is created in epistemic logic puzzles. Here follows the translation of Freudenthal's 'sum and product' puzzle as given in [1]:

"A says to S and P: I have chosen two integers x,y such that $1 < x < y$ and $x + y \leq 100$. In a moment, I will inform S only of $s = x + y$, and P only of $p = xy$. These announcements remain private. You are required to determine the pair (x,y).
He acts as said. The following conversation now takes place:
i. P says: 'I do not know it.'
ii. S says: 'I knew you didn't.'
iii. P says: 'I now know it.'
iv. S says: 'I now also know it.'
Determine the pair (x,y).

Notice that the set-up is described in such a way that it is clear that S and P have common knowledge of the ranges of the integers x and y, of the fact that only S hears their sum and only P hears their product from A, and of what is said in the conversation. All this common knowledge of the setup, plus common knowledge of the fact that S and P are perfect logical reasoners, is needed to solve the puzzle.

Something similar holds for games. Here follows a quote by Aumann about the assumption in game theory that rationality of the players, the rules of the game, and the set of players are commonly known ([2, p. 31]):
"The common knowledge assumption underlies all of game theory and much of economic theory. Whatever be the model under discussion, whether complete or incomplete information, consistent or inconsistent, repeated or one-shot, cooperative or non-cooperative, the model itself must be assumed common knowledge; otherwise the model is insufficiently specified, and the analysis incoherent."

Let us turn to Section 3 of the submission, with the above in mind.
Unfortunately, from the description of the experiment on p. 6, it is not made clear whether the facts from the introductory story ("Every morning …. your offices") are conveyed to the experimental subjects in such a way that they become common knowledge. Especially the facts about the arrival range between 8:10 and 9:10 and the fact that the subject and their colleague always arrive exactly 10 minutes apart, need to be common knowledge; and so do the goals of the coordination (when they should go to the canteen and when to their offices). It is also important that the pay-off structure (here explained on the second paragraph of p. 7) is made common knowledge between the two subjects.
However, there is nothing in the paper that discusses what measures the experimenters have taken (if any) to create any of the necessary common knowledge between the pairs of participants about the game rules and pay-offs.

Det er måske trods alt en valid indvending. Jeg tænker det er implicit at den anden person får den samme beskrivelse, men det er jo ikke noget vi siger eksplicit, trods alt. Så ret beset kan man ikke påstå det er common knowledge. I DTU-eksperimenterne er vi måske tættere på at have opnået common knowledge, men jeg ved ikke om vi kan påstå at vi har det. Jeg tror ærligt talt folk handler *som om* de har common knowledge omkring begge at have modtaget de samme instruktioner, men som en djævlens advokat kan jeg godt se at man kan finde på at grave i det.

It would be possible to create the necessary common knowledge in an experiment with two subjects who are jointly told all the needed information face-to-face by the experimenter, when both of them are jointly paying attention (see Clark & Marshall (1981) for an explanation of this way to instantly create common knowledge). It is not possible to create common knowledge about the game rules and pay-off structures between two participants in a Mechanical Turk experiment, nor by giving the introductory story to two subjects separately; and I suspect that some methods like these were used in the current experiment.

From the current description, we may not even conclude that the subjects know that their colleague has read the same introductory story as they did, so even shared second-order knowledge at seems to be missing.

This lack of common knowledge of the game setup and lack of common knowledge of rationality of the players also causes the game strategies described in Section 4 not to work. In l. 14 of Section 4, the crucial element is "and if you believe your colleague would reason as yourself", whereas in the described setup there is no reason to assume this.

The same problem causes the formal analysis of Appendix C to be incorrect. For example, it is said that "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". However, if the introductory story is not made common knowledge, the division of even/odd is not common knowledge either.

Det er meget sjældent at man beskriver reglerne for et spil forskelligt til to spillere. Enten lyver man så for den ene spiller, eller man undlader noget information. I begge tilfælde vil man fejle i sin beskrivelse af spillet. Dette er måske ikke common knowledge, men nok dog alligevel en gængs antagelse. Så igen, mit gæt er at man faktisk *vil* antage common knowledge, og det er højst usandsynligt at nogen spillere vil forestille sig at medspilleren måske har fået oplyst reglerne for spillet anderledes. Så igen: Ja, formelt set er der et problem, men næppe i praksis.

Løsningen på netop dette problem kunne være at lave nogen ekstra eksperimenter med public announcement af regler og payoff og så sammenligne med de andre resultater. Hvis disse resultater ligner hinanden, så falder hele reviewer 4s indvending til jorden. Vi kommer dog nok ikke til at kunne tilfredsstille Iris van Rooij som har skrevet mailen til os, og jeg tror måske vi er lidt bandlyst nu pga consent-eksemplet. Så vi skal nok under alle omstændigheder finde et andet sted at submitte. Men indvendingen omkring common knowledge af regler og payoff er selvfølgelig reel nok, og vil også være en indvending vi kunne høre fra logikere, så det hjælper ikke bare at resubmitte til et logik eller spilteori-tidsskrift. Reviewer #4 er jo sandsynligvis en logiker, det kunne fx være Hans van Ditmarsch.

> The analysis in Appendix C does mention the problem that the players should be perfectly rational (see last paragraph of Appendix C), but it would not be sufficient "if they knew they could expect the other player to play perfectly rational as well". In fact, the formal analysis presupposes *common knowledge* of rationality of the two players, while in the experiment nothing has been done to create such common knowledge;

Nej, og det kan vi nok heller ikke. Det er mennesker som spiller mod andre mennesker, så vi kan ikke gøre noget som ville garantere at den anden spiller spiller rationelt.

> this on top of the before-mentioned lack of common knowledge of the game rules and payoff structures.
>
> All this lack of common knowledge and possibly even shared second-order knowledge of the players' information about the game means that the results of the experiments need to be interpreted much more cautiously: there is no shared knowledge of a certain level to be contrasted to common knowledge.
>
> The participants are probably not reasoning about (levels of shared) knowledge at all, but rather about beliefs and probabilities.
> The description of the results in Section 5 appears quite appropriate to me: The subjects' possible reasoning is mostly described there in terms of "being more certain" and "you believe such a strategy to be optimal", and that seems right.
>
> However, the Discussion in Section 6 over-interprets the experimental results. The whole analysis illustrated in Figure 5 would only work if the game rules and payoffs and the subjects' rationality were common knowledge; and these items need to be second-order shared knowledge even for the first steps in the reasoning example on page 13: "if arriving at 8:40 or before the player will 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."
> Again, nowhere does the paper make clear what if anything has been done to create these prerequisites.

Dette er ikke en ny indvending, men bare en understregning af konsekvenserne af den oprindelige indvending. I samme øjeblik at spillet er common knowledge, så bliver disse påstande om shared knowledge også korrekte.

> The second paragraph on page 15 mentions computational overload due to the complexity of the game's payoff structure and gives a good explanation that the subjects probably did understand the payoff structure.
>
> The paragraph after that gives a good description of how subjects may actually reason ("…. participants do not believe that their colleague will reason in the same way as they do themselves, and instead try to guess what their colleague will choose"). Thus, participants may be reasoning about a sufficiently large probability of the other making a certain choice, instead of thinking about shared knowledge.

Ja, fair nok. Det peger også igen tilbage på det med agent-typer. Hvis man antager at den anden agent følger en simpel mixed strategy med faldende sandsynlighed for at gå i kantinen jo senere man kommer, så kan det være den strategi man reagerer på, og så tænker man også at man hellere selv må have en strategi med faldende sandsynlighed for at gå i kantinen. Kan man så påstå at de alligevel "tænker på shared knowledge"? Måske ikke. I hvert fald ikke på nogen eksplicit måde.

Så der er altså flere problemer:

1) Mangel på common knowledge om rammerne for spillet. Selvom revieweren siger dette ikke kan løses i Mechanical Turk, så er det jo ikke umuligt at lave en variant af spillet live, hvor vi opfylder disse krav. Og hvor vi så måske kan relatere til det større datagrundlag i Mechanical Turk.

2) Mangel på common knowledge af rationality. Det er svært at se hvordan man skulle kunne realisere den i et spil mellem mennesker. Og selv hvis man sagde: “du spiller mod en robot som er perfekt rationel og antager du er det samme”, så vil det jo næppe hjælpe. Jeg tænker næsten det er umuligt at gøre i et spil, så det eneste sted hvor vi reelt kan konkludere noget er i forhold til deres meta-overvejelser omkring spillet. “Hvis I kunne aftale et cutoff-tidspunkt” er jo en måde at sige på at man nu gør hele strategien både pure og common knowledge. Her siger flest 8:40, som viser at de er opmærksom på den anden agents perspektiv (eller ville 8:50 være OK), men ikke itererer videre end det. Så måske skulle der lægges mere vægt på disse dele, men spørgsmålet er hvilken rolle der så er tilbage for selve det spillede spil.

> Contrary to Footnote 5, it may well be that there is no "illusion of common knowledge" among the subjects, but that there is, for example, a common belief that both will decide to go the canteen with high probability.

Jo, men det er så modsagt af netop det med deres overbevisning om at 8:40 er safe. Medmindre high probability og safe betyder det samme, så der skulle vi måske også være lidt klarere i vores formuleringer. Vi burde spørge dem på en måde, så det er klart at safe betyder at der ikke er nogen risiko for miskoordinering (sandsynligheden for miskoordinering er 0).

> All this puts into question the main conclusion and even the title of the paper: it appears highly unlikely that there really is any shared knowledge of a relevant level, let alone a curse of shared knowledge.

Jeg synes det er dramatisk at konkludere fra at der kan være tvivl om hvorvidt der er common knowledge af reglerne (eller common belief) til at det er “highly unlike” at der er nogen shared knowledge overhovedet. Jo, OK, i det omfang at vi maksimalt kan have common belief omkring reglerne, så kan vi selvfølgelig også kun argumentere for shared belief og ikke shared *knowledge*. Men det lyder lidt på de tidligere dele af reviewet at revieweren mener at det er hele ideen om hvorvidt noget som helst er “shared” der er på spil her. At man måske slet ikke ræsonnere med nogen former for Theory of Mind ræsonnering overhovedet.

> In this light, the conclusion in Section 7 is much too strong, especially the end of the first paragraph (p. 17) and the second paragraph (p. 17 and 18).
>
> [1] van Ditmarsch, H. P., Ruan, J., & Verbrugge, R. (2008). Sum and product in dynamic epistemic logic. Journal of Logic and Computation, 18(4), 563-588.
>
> [2] R.J. Aumann. Game theory. In J. Eatwell, M. Milgate, and P. Newman, editors, Game Theory, The New Palgrave, pages 1-54. Macmillan, 1997.

Så det er lidt noget pis, for at sige det på godt gammeldags dansk. Vi bør lægge en slagplan, og vi kan selvfølgelig mødes og snakke om det, men jeg tror egentlig det måske er meget godt hvis vi stadig med hver især at tænke over det og så formidler vores tanker via mail. Jeg synes i hvert fald ikke det er let at konkludere noget og finde ud af hvad den bedste vej videre er. Så jeg ville være meget i tvivl om hvorvidt vi kunne mødes i fx en time og nå til nogen som helst meningsfuld konklusion. Men *efter* vi har brainstormet hver især kan det godt være at det ville være smart med et møde. Men umiddelbart tænker jeg min egen strategi vil være at vente på at høre hvad I har af kommentarer til det hele, og efterfølgende vil jeg så prøve at tænke dybt over det både i relation til mine egne tanker og i forhold til hvad I nu ender med at sige, og så se om det giver mig et eller andet form for nogenlunde klart billede af vejen videre herfra. Jeg har aldrig tidligere været i så meget modvind med en artikel, så det er lidt nyt for mig at skulle forholde mig til. Jeg har lidt sådan en “skomager bliv ved din læst”-følelse, hvor jeg tænker at jeg aldrig skulle være gået med på et projekt med empiriske eksperimenter og relationer til kognitionsvidenskab, men det er jo også noget fjolleri at tænke sådan: Hvis man ikke forsøger at bryde grænser til andre fagligheder, så bliver man også bare ret snæver og begrænset i sin forskning (hvilket så dog nok også desværre er sandt for 90+% af alle forskere).

Thomas

> ==========

==========

Submission details:

Mar 23, 2022

Journal Name: Cognitive Science

MS ID: 20-373R1

Article Type: Regular Article

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

Submitting Author Name: Thomas Bolander

Submitting Author Email: tobo@dtu.dk

Author Names: Thomas Bolander; Robin Engelhardt; Thomas Nicolet

All Author Emails: 1. Thomas Bolander
E-mail Address: tobo@dtu.dk
ORCID: 0000-0003-1551-1703

2. Robin Engelhardt
E-mail Address: robin@hum.ku.dk
ORCID:

3. Thomas Nicolet
E-mail Address: thomas1nicolet@gmail.com

ORCID:

All Author ORCID iDs:

Article 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.

Please be aware that if you ask to have your user record removed, we will retain your name in the records concerning manuscripts for which you were an author, reviewer, or editor.

Robin Engelhardt <robin.engelhardt@gmail.com> Thu, May 19, 2022 at 3:06 PM
To: Thomas Bolander <tobo@dtu.dk>
Cc: Thomas Nicolet <thomas1nicolet@gmail.com>

Det er en alt for deprimerende nyhed til at jeg kan svare på det hurtigt. Umiddelbart lyder det for mig som et bisset trick. Man kan altid påstå at common knowledge ikke er etableret, selv hvis det bliver råbt ind i ørerne på folk. Jeg skal lige fordøje det...

/r

Robin Engelhardt

Teaching Associate Professor

ph.d.

University of Copenhagen

Department of Food and Resource Economics

Rolighedsvej 23, 1958 Frederiksberg C

DIR +45 26 30 04 03

re@ifro.ku.dk

www.robinengelhardt.me

[Quoted text hidden]

Thomas Bolander <tobo@dtu.dk> Thu, May 19, 2022 at 3:42 PM
To: Robin Engelhardt <robin.engelhardt@gmail.com>
Cc: Thomas Nicolet <thomas1nicolet@gmail.com>

Sent from my iPhone

> On 19 May 2022, at 16.06, Robin Engelhardt <robin.engelhardt@gmail.com> wrote:
>
> Det er en alt for deprimerende nyhed til at jeg kan svare på det hurtigt.

Det var også sådan jeg havde det. I meget lang tid…

> Umiddelbart lyder det for mig som et bisset trick. Man kan altid påstå at common knowledge ikke er etableret, selv hvis det bliver råbt ind i ørerne på folk.

Det er også rigtigt. Men det er trods alt etableret med lidt større sandsynlighed hvis det bliver råbt i en masse mennesker tilstedeværelse og de mennesker måske endda har øjenkontakt og dermed kan vurdere med en vis sandsynlighed at de andre også har hørt det. Men ellers har du jo ret i at der intet findes som *i praksis* garanterer common knowledge, så det kan kun være et spørgsmål om at sandsynligheden for at der er etableret (og spillerne tror på) common belief er mindre i vores spil end hvis reglerne var annonceret i begge spilleres nærvær.

> Jeg skal lige fordøje det...

Ja, det er klart.

Thomas

[Quoted text hidden]

Robin Engelhardt <robin.engelhardt@gmail.com> Wed, Jun 1, 2022 at 11:07 AM
To: Thomas Bolander <tobo@dtu.dk>
Cc: Thomas Nicolet <thomas1nicolet@gmail.com>

Kære begge,

Nu har jeg sundet mig lidt, men jeg har ikke ændret mening: det er en eftertackling af en ekstra reviewer som leger dørvogter over for newbies. Kritikken er urimelig både hvad angår common knowledge om spillets regler og hvad angår spillernes rationalitet.

Jeg foreslår at vi finder en anden journal, hvor ventetiden er lav, f.eks.
https://www.mdpi.com/journal/information eller
https://www.mdpi.com/journal/behavsci

man plejer at få en decision efter 2-3 uger.

kh, robin

Robin Engelhardt

Teaching Associate Professor

ph.d.

University of Copenhagen

Department of Food and Resource Economics

Rolighedsvej 23, 1958 Frederiksberg C

DIR +45 26 30 04 03

re@ifro.ku.dk

www.robinengelhardt.me

[Quoted text hidden]

Thomas Nicolet <thomas1nicolet@gmail.com> Thu, Jun 2, 2022 at 7:16 PM
To: Robin Engelhardt <robin.engelhardt@gmail.com>
Cc: Thomas Bolander <tobo@dtu.dk>

Kære Robin og Thomas

Tak for tilbagemelding, og håber du er helt ok og oppe at køre igen Thomas.

Jeg kunne skrive en del kommentarer om reviewet, hvor og hvorvidt de tager fejl eller har ret, men det tror jeg vi er relativt afklaret med for nu. Jeg synes mange af pointerne er gode, men ikke afvisende.

Jeg har svært ved de her spørgsmål, som der tilsyneladende bliver lagt vægt på igen:

“to what extent [are] the results … due to the artificiality of the task instructions”

“ whether … the experimental task … taps actual, natural cognitive processes that we as cognitive scientists aim to understand”

At menneskers adfærd (og derfor kognition) i spil som vores blot er artefakter, og ikke faktiske, naturlige processer, virker radikalt, og det kunne man også skrive en masse om.

Vi baksede en del med det under økologisk validitet, og som beskrevet ovenfor, lyder det bare urimeligt, så jeg synes også det lyder fint at finde en anden journal. Vi kan evt. have et møde om det på et tidspunkt, få redigeret lidt og sendt afsted.

Omkring skomager-ved-egen-læst Thomas, så forstår jeg godt din følelse. Jeg kan sige at jeg har været super glad for at kunne arbejde sammen, og er sikker på at modvind også er essentielt for at forske bedre. Udover at hjælpe en med et kritisk syn på eget felt, tænker jeg også det kan give et værdifuldt og kritisk syn på kognitionsvidenskab. Hvis felter bliver for meget til siloer, bliver det jo kun løst af dem som møder modvind, når de bygger bro.

Bedste hilsner Thomas
[Quoted text hidden]

Robin Engelhardt <robin.engelhardt@gmail.com> Mon, Jun 20, 2022 at 3:34 PM
To: Thomas Nicolet <thomas1nicolet@gmail.com>
Cc: Thomas Bolander <tobo@dtu.dk>

Hvad siger I to til en stille og rolig frokostaftale i næste uge?

kh, robin
[Quoted text hidden]

Thomas Bolander <tobo@dtu.dk> Mon, Jun 20, 2022 at 4:06 PM
To: Robin Engelhardt <robin.engelhardt@gmail.com>
Cc: Thomas Nicolet <thomas1nicolet@gmail.com>

God idé. Onsdag eller torsdag passer mig bedst.

Thomas
[Quoted text hidden]

Thomas Nicolet <thomas1nicolet@gmail.com> Mon, Jun 20, 2022 at 4:12 PM
To: Thomas Bolander <tobo@dtu.dk>
Cc: Robin Engelhardt <robin.engelhardt@gmail.com>

Det lyder fint. Onsdag og torsdag er ok for mig.
T
[Quoted text hidden]

Robin Engelhardt <robin.engelhardt@gmail.com> Mon, Jun 20, 2022 at 4:17 PM
To: Thomas Nicolet <thomas1nicolet@gmail.com>
Cc: Thomas Bolander <tobo@dtu.dk>

super. Lad os tage onsdag. Kender I et godt sted? Jeg kan foreslå
kaffesalonen ved dr. louises bro.
khr
[Quoted text hidden]

Thomas Bolander <tobo@dtu.dk> Thu, Jun 23, 2022 at 12:36 PM
To: Robin Engelhardt <robin.engelhardt@gmail.com>
Cc: Thomas Nicolet <thomas1nicolet@gmail.com>

Det er meget fint. Kl 12?

Thomas

Sent from my iPhone

> On 20 Jun 2022, at 16.18, Robin Engelhardt <robin.engelhardt@gmail.com> wrote:
>
> super. Lad os tage onsdag. Kender I et godt sted? Jeg kan foreslå
[Quoted text hidden]

Thomas Nicolet <thomas1nicolet@gmail.com> Thu, Jun 23, 2022 at 3:25 PM
To: Thomas Bolander <tobo@dtu.dk>
Cc: Robin Engelhardt <robin.engelhardt@gmail.com>

Ja, det lyder meget fint, lad os sige det.

T
[Quoted text hidden]

Robin Engelhardt <robin.engelhardt@gmail.com> Thu, Jun 23, 2022 at 3:28 PM
To: Thomas Nicolet <thomas1nicolet@gmail.com>
Cc: Thomas Bolander <tobo@dtu.dk>

deal.
[Quoted text hidden]