Common Knowledge First

Seth Yalcin
yalcin@berkeley.edu

December 8, 2023

1

Between *common knowledge* and *common belief*, is one any sense the more important notion in the explanation of coordinated action?[1] Among those not already skeptical of one or both of these notions in the first place,[2] a not uncommon view is that the distinction between the two is, at least very often, not of great consequence. Readers are often warned that the 'knowledge' in 'common knowledge' isn't performing important work, and shouldn't be credited with much significance.[3] As Lederman observes in a survey article on the topic,

> Often, especially in economic theory (and almost without exception in economic theory from the 1980s and 1990s), the expression "common knowledge" is used generically to mean "common knowledge, common belief, or common certainty" (where someone is certain that p just in case they assign p probability 1). Thus authors speak of "common knowledge assumptions" where the assumptions may not require knowledge at all. [Lederman, 2017, 182]

[1] In this paper I assume the standard (sometimes called *hierarchical* or *iterative*) definitions of these two notions: p is *common knowledge* (*belief*) in a group g just in case everyone (in g) knows (believes) p, everyone knows (believes) everyone knows (believes) p, and so on. Fagin et al. [1995] is a textbook overview in the context of modal logic. See Lederman [2017], Vanderschraaf and Sillari [2022] for discussion of alternative definitions in the literature.

[2] See Fagin et al. [1999], Lederman [2018a,b] for some important challenges to the applicability of these notions, and Thomason [2021], Greco [2023] for some replies. In another vein, Monderer and Samet [1989], Paternotte [2017] reject common knowledge based on a general skepticism about knowledge of other minds. This paper is simply conditionalized on the hypothesis that common knowledge and common belief both exist and have nontrivial explanatory application. My central claim is one about the comparative explanatory power of these two notions. It would be a further project to compare common belief/knowledge-based explanations of coordination action in general with explanations that would eschew appeal to such iterated attitudes altogether. I take it that skeptics about common knowledge will be more interested in that question.

[3] See for instance Heal [1978, 116], Thomason [2021, 229].

Why this blasé attitude about whether what's at issue is *knowledge*?

There may be several motivations. In this paper I begin with what might be the most common motivation for thinking that common knowledge is not explanatorily more important, general, or fundamental than common belief. This is the motivation based on the simple observation that it is possible to have cases where a group manifests coordinated behavior by way of collectively assumed "information" that is not known. I will argue that in standard examples of such cases, the agents possess far more common knowledge than is recognized, and that as a result, the cases do not show what they are meant to show. They are not, after all, cases of coordination that aren't grounded in common knowledge.

This prompts the next question: what would it take for a case of coordination to be one fully grounded merely in common belief, with no relevant common knowledge? I suggest that the most uncontroversial examples of this would be in effect *thoroughly Gettierized* cases of common belief. I argue that such cases—common belief in the complete absence of relevant common knowledge—are a much rarer phenomenon than it is often assumed, and that uncontroversial examples of it do not in fact enable something recognizable as coordinated action at all.

Greco [2016] argues that knowledge robustly explains *success* in action in ways that belief cannot match. Coordination, I suggest, is a variety of success in collective action. Drawing on and extending Greco's line of thinking, I argue that common knowledge can explain success in collective action—coordination—with a robustness that mere common belief cannot match. My conclusion is that the notions are not on explanatory par. Common knowledge is the fundamental notion in the explanation of coordinated action. Common belief makes for robust explanations of coordination only insofar as it approaches common knowledge.

2

It is not hard to think of cases like this: some agents engage in coordinated action, but the coordination appears to be facilitated by common belief in a proposition that is false. If such cases are possible, then since false propositions can't be known, common knowledge in that proposition couldn't be what facilitates the coordination observed. Greco [2015] provides a clear summary of this motivation for downplaying the 'knowledge' in 'common knowledge', and for thinking common belief is often at least as explanatory. Here is the example he

uses to illustrate the idea that "it is possible to coordinate around information that is not known":

> Suppose we are making plans to see a movie, and I make the following proposal to a group of friends: "Let's meet at the theater fifteen minutes before the movie starts." Suppose further that, while we commonly take for granted that the movie starts at eight, it actually starts at nine. In such a case, one should predict that we'll all show up to the theater fifteen minutes before eight—in one sense, we'll succeed in coordinating... To the extent that one is sympathetic to common knowledge explanations, it's natural to think that explaining how we managed to coordinate will require appealing to some sort of common-knowledge-like attitude that we bear towards the claim that the movie starts at eight. But that attitude cannot literally be common knowledge—we don't commonly know that the movie starts at eight, as the movie does not start at eight. I take the lesson of the example to be that questions about common "knowledge" really arise for any sort of attitude—belief, acceptance, presupposition, etc.—that can figure in explanations of coordinated action. Moreover, if explaining how we attain common knowledge requires positing an iteration principle for knowledge... then we'll also have to posit iteration principles for other attitudes that can play similar roles in explanations of coordinated action. (757)

Greco doesn't claim this is an original argument; he remarks that this view is "standard in work on common knowledge" (757).

The argument is too quick. Greco's example is not an example of coordination that common knowledge cannot explain. It is of course correct that in the case, the agents do not have common knowledge that the movie starts at eight (since it starts at nine). Nevertheless, common knowledge still figures centrally in the most natural versions of cases like this. For the agents *do* have common knowledge of the fact *that they all believe (indeed, take themselves to know) that the movie starts at eight*. And this fact—a fact about the common knowledge possessed by the agents in the case—is sufficient to predict what coordination is observed (viz., the achievement of their collective aim to arrive at the same time).

Let me draw out the point out. We see an example like the movie case, and we know that we can't explain the coordinated action in question by appeal to the idea that:

> The group knows p,
>
> The group knows that the group knows p,

The group knows that the group knows that the group knows p ...

—for p is false. The standard reaction is to solve this problem by replacing *every* appearance of 'knows' in the hierarchy with 'believes':

The group believes p,

The group believes that the group believes p,

The group believes that the group believes that the group believes p ...

—by appeal to common belief. But this is an overreaction. Greco's group is epistemically compromised with respect to p, but they are not epistemically compromised with respect to the truth that they all believe p, or with respect to their higher-order knowledge of that fact. The example gives us reason to change 'knows' to 'believes' at the ground level of the hierarchy, but only there:

The group believes p,

The group knows that the group believes p,

The group knows that the group know that the group believes p ...

Thus such examples show that we may sometimes need to explain coordination in terms of common knowledge of what is believed. But they are not cases where common knowledge isn't what explains the level of coordination that is observed. Indeed, in the movie example we already notice that the *absence* of common knowledge about the movie time corresponds with the group's *inability* to collectively act in a certain way that the group would collectively prefer. The group wants to arrive, not just at the same time, but fifteen minutes before the movie time. Arriving at the same time just requires common knowledge of what everyone believes the movie time to be; but arriving fifteen minutes before the movie time requires common knowledge of the movie time.

3

I expect the following reply: "Granted that the common knowledge just cited *can* explain the coordination observed. Still, the common belief possessed by the agents is *also* sufficient to predict what coordination is observed; so at best we have an explanatory parity here."[4]

[4]Indeed, one might naturally go a step further: if common belief *can* fully explain the coordination observed, then since it is a strictly weaker property of the group (common belief in p is implied by, but doesn't imply, common knowledge that everyone believes p) common belief has claim to being the *more basic* explanatory notion.

Is this correct? To assess whether common belief can explain coordination "on its own", the natural sort of cases to consider are those that isolate it from relevant common knowledge, so that common knowledge couldn't possibly be what does the explaining. Such cases are unlike the sort Greco mentions. We need cases of common belief not understandable as founded in common knowledge. Call a case of common belief in p *pure* just in case for no n is it the case that the agents commonly know that they mutually believe^n p.[5] What we must look for are cases of coordinated action based on, explained by, pure common belief. For this it would suffice to have cases of coordinated group behavior where every level of the hierarchy is epistemically compromised, so that no level of mutual belief in the group constitutes knowledge. (At least, this needs to be so in connection with propositions relevant to the coordination we are purporting to explain.) These would be cases of common belief in p where the group believes (but doesn't know) that the group believes p, where the group believes (but doesn't know) that the group believes that the group believes p, and so on.

And here we confront the point that the most uncontroversial examples of pure common belief are generally the stuff of bad sitcom writing. Here is one such case:

> In a noisy and darkly lit bar, Alice sees from behind a person she mistakes for Carl. (He's sitting in Carl's usual seat, looks just like him from behind, and so on.) She leans in whispers the following into the stranger's ear: "It's me, Alice. I've gotta run now, but let's meet tomorrow at the cafe at noon." The stranger is deaf, so he doesn't respond, but he gives a thumbs up without looking at Alice. Alice interprets this as agreement about the plan, but in fact the stranger is just signaling to the bartender that he wants another shot. Meanwhile, Carl is at another bar and coincidentally undergoes a completely analogous confusion: he mistakes a stranger for Alice, says "Let's meet tomorrow at the cafe at noon", and experiences the illusion of assent from the stranger. Alice and Carl leave their respective bars none the wiser, with both expecting to meet the other the next day at the cafe at noon.

Consider Alice and Carl the next morning. They commonly believe they are meeting at the cafe at noon. But they don't commonly know they are meeting at the cafe at noon, and nor do they commonly know that they mutually believe^n they are meeting at the cafe at noon, for any n. So this is a case of pure common

[5] Where some agents *mutually believe* (or mutually believe^1) that p in case they all believe p; and they *mutually believe*^{n} that p ($n > 1$) just in case they mutually believe^{n-1} that they mutually believe p.

belief in p. It is a case of thoroughly Gettiered common belief. If there was an important sort of coordination to explain in this kind of case, it might serve to motivate an explanatory role for common belief *per se*.

But there is no meaningful coordinated behavior at all here. Alice and Carl do meet at the cafe, but this isn't the result of their actually being coordinated. It is the result of a weird accident, an incredibly unlikely sequence of events. I take that in the target sense of "coordination", "lucky coordination" is an oxymoron. In a group capable of coordinated behavior, the attitudes of the agents are not merely accidentally aligned, but robustly covariant over some relevant range of possibilities.[6] Since there is no coordination present in the above case, there is no coordination for common belief to explain. So we don't yet have motivation for the thought that common belief *per se* explains. To the contrary: the example shows that common belief is compatible with the absence of the possibility of coordinated action, and that its presence does not suffice to enable coordination, even when everything else is in place (the agents have suitably aligned preferences, they are positioned to perform relevant actions, etc.). To this extent, the example undermines the idea that common belief is what explains coordination where we do find it.

This is not to say that the facts of the case lead us to predict *nothing* about what will happen the next day. Of course, we predict that Alice and Carl will each show up at the cafe at noon. For we expect that they will each individually act in accord with their private belief and desires, and that the outcome resulting from those actions will coincide (luckily) with their common preferred outcome. But even these predictions, it is worth stressing, don't importantly depend on the fact that the agents have common belief. Alice's action of heading to the cafe is explained, not so much by the fact that she actually *has* common belief with Carl about their meeting at the cafe, but rather by the fact that she believes she does; and likewise *mutatis mutandis* for Carl.[7] We can make this more obvious by adjusting the case so that only one of them believes that they have common belief with the other. Suppose Alice is as before, but that Carl believes they will meet at the cafe at noon only because a palm reader has convinced him that Alice is fated to be there. Now there is no common belief, since Carl doesn't believe that Alice believes that Carl believes they are going to meet. But Alice's

[6] The *relevant* is important: one can artificially restrict one's attention to a space of possible scenarios where some extraordinary coincidence like the one above occurs, thereby creating the illusion of a sensitivity or interdependence in the way the agents act, an illusion of robust covariance in expectations and behavior, that resembles what actual coordination would make for. But an illusion of coordination isn't coordination.

[7] Indeed we could say that they each individually incorrectly believe that they have common *knowledge* that they are meeting.

heading to the cafe in this case is explained in the same way that it is explained in the first case, by appeal to her (now false) belief that she has common belief with Carl about meeting. The general point here is that while there are facts about what *individuals* do here that admit of explanation, in terms of the private beliefs (confusions) of those individuals, there is no interesting fact of successful collective action in the case to explain, and so no work for the notion of common belief to do.

Objection: "Look, I'll *give* you the word 'coordination'. Granted, they are not coordinated. Still, Alice and Carl each have a common preferred outcome—meeting at the same place and time—that depends on the actions of the other, in that the action that is best for one depends on the action taken by the other. And in the end, they each acted in such a way so that this common preferred outcome was realized. Are you saying that the fact that (in the original case) they had common belief about their meeting time and place *doesn't explain* how, in the end, they *succeeded* in realizing their common preferred outcome?"

That is what I am saying—or at any rate, I want to claim that it provides at best a shallow explanation, since there is no robust explanandum here in the first place. We should keep two points in mind. First, not everything that happens has a deep or interesting explanation; some things just happen coincidentally or by chance. Alice and Carl's meeting at the cafe is such an event. Second and related, there is a difference between prediction and explanation. Together with the other facts of the case, the fact of their common belief does give rise to the *prediction* that they will show up at the same time. Furthermore, their individuals beliefs and desires do *explain* why they each show up at the place and time that they each did. But there is no substantive explanation available for how they *succeeded* in realizing their common preferred outcome—their getting to the *same* place and time. That part was just luck. So again, common belief can't earn its keep by giving such an explanation.

4

Let me emphasize that the issue here is not (merely) about explaining collective action, but more specifically about explaining *success* in collective action. Greco [2016] argues that where action is guided by knowledge, the success of the action admits of deeper explanation than it otherwise would.[8] This point is important and relevant here, since coordination involves a kind of success in collective

[8] On the role of knowledge as compared to belief in the explanation of action, see also Williamson [2000], Gibbons [2001].

action. Let me draw out the point with some examples, focusing on belief versus knowledge explanation at the individual level. Then I will circle back to extend the points to common belief and common knowledge.

Bob plays a game to win. In the end, he does win. We ask: "What (causally) explains his winning?" Here are two accounts:

> Belief Explanation: Bob believed action X would result in success, and this belief caused him to take action X. In fact, X was the action that would result in success. So, Bob won.
>
> Knowledge Explanation: Bob knew action X would result in success, and this knowledge caused him to take action X. So, Bob won.

The explanation in terms of knowledge, I now argue, is importantly deeper.

A first point is that it simply *says more*, at least assuming knowledge implies belief. The explanation in terms of belief leaves open the question: "Why did Bob believe that X would result in success?" The explanation in terms of knowledge, while not fully answering that question, does begin to address it, by imposing nontrivial constraints on the form that the answer could take.

But it is not simply the fact that the explanation in terms of knowledge says more that makes it deeper explanation of the winning. Rather, it is the specific more that it adds. For we could add to the explanation in terms belief, filling in some details about etiology of Bob's belief, so that it starts to address the question *why* Bob believed as he did. But merely doing that won't necessarily suffice to make the explanations of the winning equally deep; it matters how the details are filled in.

To bring that out, suppose the game was a lottery. The explanation in terms of belief goes like this: "Bob believed that such-and-such were the winning numbers. This caused him to select those numbers. In fact, those were the winning numbers. So, Bob won." This is, I guess, an explanation of Bob's success, but intuitively it is not much of one. To look for something deeper, we ask: "Why did Bob believe that those were the winning numbers?" If the answer lies in his trust of astrology, then we conclude that Bob's winning doesn't admit of any deep explanation. It was just luck. So here, while we added more in answer to the question why Bob believed as he did, the more we added failed to make the explanation of the win any deeper. On the contrary, filling the details out in this way precludes the possibility of a non-shallow explanation of Bob's winning.

And this isn't to do with any lack of justification in Bob's astrology-based belief. It could just as well have been a Gettiered belief. Suppose instead that

Bob's belief about the numbers was justified: he was convincingly pranked by someone claiming to have inside information about the lottery outcome, but that person just made up the numbers, coincidentally hitting on the right ones. This change to the story doesn't make our explanation of Bob's winning any less shallow. His winning is still a matter of an amazing coincidence between his choosing those numbers and their being the correct numbers.

Things are different, however, when the success is explained in terms of knowledge. If Bob chose the numbers he did because he *knew* they were correct, there was no amazing coincidence between his choosing those numbers and their being correct. His winning ceases to surprise; it is explicable. I take it this is intuitively obvious. But what is it about knowledge *per se* that makes for a deeper explanation of successful action? The answer is that where there is knowledge-guided choice, there is a robust connection between the agent's choosing and the facts about the choice that make that choice the success-conducive one. Knowledge just is a robust sensitivity to the fact known, so that where an event of choosing is explained by the agent's knowledge of what would bring success, there is a robust connection between the choosing and the fact of success that results.[9]

It is a triviality about rational agency, and about how we are understanding "success", to say that a rational agent chooses the option that they believe is the one most conducive to success. But identifying the option with that property—the property of *actually being* the most conducive to success—is often hard. We often get it wrong.[10] What knowledge can supply is sensitivity to that property, and the ability to chose an option *because* it has that property. Where the knowledge-guided agent chooses, there is something nontrivial to say in explaining what success results. We can say that the option's *actually making for success* is *why* the knowledge-guided agent chose it. To put it simply, where the agent *knows* what option will make for success, the agent has the ability to *choose success*. Without such knowledge, the agent is positioned only to choose what they believe will make for success—but again, that's a triviality.

I suggest this all carries over relatively straightforwardly to the issue of the comparative explanatory power of common belief versus common knowledge in accounting for successful collective action. Much as Bob's astrology-induced belief about the winning numbers doesn't interestingly explain his winning the

[9] Of course, the target sense of robustness has been cashed out by philosophers in different ways, for instance via the idea that knowledge is *sensitive* (Nozick [1981]) or *safe* (Sosa [1999], Williamson [2000] Pritchard [2005]). Greco [2016] considers the matter from the perspective of the role robustness plays in explanation in general.

[10] To use the sense of 'wants' brought into focus by Jerzak [2019]: we often choose the option we don't want.

lottery, Alice and Carl's common belief about their meeting doesn't interestingly explain their success in meeting. I'll fix on Alice, but all of the following applies *mutatis mutandis* to Carl. The feature of Alice's action (of getting to the cafe by noon) conducive to success is the fact that this action will take Alice to a place and time that Carl will be. But the facts that make *that* the case—that make it true that Carl is going to be then and there—are not facts that play any role in causing, or otherwise explaining, the beliefs that guide Alice's selection of that very action. Alice is in the relevant sense *insensitive* to the feature of the action she chooses that in fact makes it conducive to success. We therefore cannot explaining her as choosing the action she does *because of* its having this success-promoting feature—just as we cannot explain Bob as choosing the numbers he chose *because* they were the winning ones. Since the beliefs guiding her choices of action aren't knowledge, and her choices are only accidentally connected to the success that results, the explanation of success her belief state affords is shallow at best.[11]

5

Am I conveniently ignoring games mixing chance and skill? Very often successful action admits of interesting explanation, though the agent did not know *for sure* that the action they selected would result in success. Good poker playing, for instance, is a matter of sensitivity to the probabilities. The different actions available have different odds of success, and the victory of a player is interestingly explainable (or not) depending on the sensitivity of their choices to the actual odds that those choices would bring success. But if that is right, it seems that the success of an action might admit of interesting explanation, though the agent didn't *know* their choices would assure victory. Don't we need here to talk in terms credence, or rational credence—things closer to the belief side of the fence? And on reflection, surely this is the norm in life, as the probabilists

[11] Greco's movie case can be modified to make a similar point. As he notes, in one sense the group in that case will succeed in coordinating (they will arrive at the same time), but in another sense they will fail, since they had aimed to collectively arrive fifteen minutes before the movie time. Imagine that the group's belief that the movie time starts at eight has the same mistaken basis, whatever it is, as it does in Greco's original case, but that, independently, the theater makes a late change the showing times, and coincidentally moves the screening in question from nine to eight. Now the group's common belief that the showing time is at eight is correct, and so they actually will collectively arrive fifteen minutes before the movie time. Is this enough to make it true that they coordinated around the movie time? Is there further robust fact of success in collective action to explain now, in terms of common belief? No—they just got lucky. The group's view about the movie time, and correspondingly their behavior, was not relevantly sensitive to the movie time.

emphasize: we often don't *know* what option will *assure* success, but we often have a sense of the odds, and to the extent we track the real odds and success results, the success is less surprising, and more explicable.

I think this all correct, but I want to suggest that it does not undermine the tie between knowledge and the explanation of successful action. A first point is that to the extent that knowledge of what action will *assure* success is out of the question in a given case, this does result in a degree of explanatory compromise. The expert poker player's acumen may go a long way explaining why she won, but a dose of chance will have to figure in the full explanation—and the bigger the dose, the weaker the explanation.

But second and more important, I claim the acumen in question is knowledge. The expert's alternative options for action in the game have their various actual probabilities for success. The expert is an expert because she is sensitive to those features of her options when she comes to choose. When she selects the option that is in fact most likely to bring about success, it is because it is the option has that property. Her choosing is caused by her knowledge of what is likely.

Saying this is compatible with saying that she chose the option she did because, of all the options, she had the highest credence in that option's bringing success. But I want to suggest that this choosing only makes for a non-shallow explanation of what success resulted insofar as the credence guiding the choice constitutes knowledge—insofar as it is *probabilistic knowledge* in the sense developed by Moss [2013, 2018]. As Moss brings out, justified credence and justified belief are alike in their potential to be Gettierized. Rational credences that correspond only accidentally with the true probabilities do not make for non-trivial explanations of the success of an action any more than cases where the agent lucks into a justified true belief of the ordinary nonprobabilistic variety. Suppose Bob makes his poker choices by aligning his credences with those recommended by a sophisticated AI system. It gives Bob knowledge by testimony. The system has always worked, but owing to a recent software update, it has gone haywire, unbeknownst to Bob. The probabilities of success it assigns to Bob's various options are in fact randomly caused. But in the game Bob just won, the randomly assigned probabilities of the AI system just happened to match the actual ones. Bob's credences were sensibility arrived at (and justified in at least one sense), and owing to them, he made choices that were, in fact, the ones most likely to make for success. But when he wins, the success can only be explained as a freak accident. Because his credences didn't constitute knowledge, he cannot be described as having chosen the options he did *because*

those options had the property of raising the actual probability of success. So although his actions made success more likely, they don't give a non-shallow explanation of the success that resulted.[12]

6

We have observed that the existence of common belief in a group is compatible with the group's having arrived at that common belief in a defective way. It is compatible with the group's members being almost entirely informationally isolated from each other. It is compatible with the etiologies of the beliefs of the individual members being completely different. I have argued that where common belief owes to random chance in this way, there is no possibility of group behavior intelligible as coordinated action, and that any explanation of how it is that the group succeeded in realizing a common preferred outcome will be shallow at best.

One might concede all this, and yet argue that none of it yet reveals common belief to be explanatorily deficient in *normal* cases, as follows. "Common belief is a dispositional property of a group, as ordinary belief is a dispositional property of the individual. But dispositions are often a matter of what happens, or tends to happen, in *normal* conditions, or in some conditions that are relevantly *favorable*. The solubility of a thing in water, for instance, is a matter of what will happen if you put it in water under normal conditions—not in any possible conditions. And this point is relevant to the causal-explanatory reach of the disposition. The solubility of a material in water might, for instance, causally explain the behavior of the material in some normal context, while entirely failing to explain the behavior of that same material under weird conditions. If under weird conditions the material's solubility in water doesn't cause the results that it normally would, this doesn't mean that, in the relevant favorable conditions, the material's solubility in water isn't after all interestingly explanatory of what happens. Likewise, common belief might causally result in, and robustly explain, successful coordination in *normal* or relevantly *favorable* cir-

[12] Monderer and Samet [1989] define a probabilistic counterpart of common (full) belief, as follows: some agents *commonly p-believe* q just in case they each believe that q is likely to be true with probability at least p, each believes that the other believes it with probability at least p, and so on. One might infer I must be claiming that this notion lacks explanatory power. But if we recognize with Moss [2018] that probabilistic knowledge is possible, then (insofar as interesting common p-belief is possible) I see no obstacle to the idea that common p-belief can itself sometimes constitute knowledge—that there can be *common p-knowledge*. The points we have stressed about the tie between knowledge and the explanation of successful action do not threaten the explanatory power of common p-knowledge, and so do not threaten the interest of the notion Monderer and Samet formalize.

cumstances, though it fails to produce coordination, or allow for a non-shallow account of success, in the sort of deviant cases emphasized above."

On this take, my complaints about the causal-explanatory power of (common) belief are alleged to be like the complaint that the flammability of a match doesn't help to explain why it lit when struck, because there are some abnormal conditions—such as when there's no oxygen, or when the match is wet—where the flammability wouldn't make for lighting.

We must agree that the causal potency of a disposition can be sensitive to the obtaining of some background conditions. But this abstract point by itself doesn't yet reveal how or why it would be that in the relevant favorable conditions, a group's having common belief *per se* would make for non-shallow explanations of the group's success in coordination. The points made in the last few sections were aimed at drawing out *why* mere (common) belief yields only a shallow explanations of success in (collective) action. If they were on track, they at least shifted the burden. If one wants to defend the view that common belief *per se* can explain coordination or successful collective action, one cannot settle for an abstract point about what is logically possible for dispositions. One has to explain what it is about belief in the normal or favorable conditions that makes for success.

And of course, one has to do while avoiding this answer: what it is about belief in the favorable conditions that enables it to robustly explain successful action is that in those conditions, such belief constitutes knowledge.

Meanwhile, there is a case to be made for that sort of answer. That is, there is a case for the view that (common) belief *does* explain, restricted to right class of favorable conditions, but only in virtue of the extent to which it amounts, in those conditions, to (common) knowledge. Such a view might itself naturally flow from an explanation of belief in terms of knowledge—a "knowledge first" view. A view along these lines is defended, for example, by Stalnaker.[13] He summarizes his view in the introduction to Stalnaker [2019b] like this:

> Knowledge whether ϕ, according to a slogan I like, is the capacity to make one's actions depend on whether ϕ. Knowledge is a matter of causal sensitivity to facts that are the subject matter of one's knowledge. My earlier gestures at explaining intentionality took a similar form: I took belief and desire to be the basic intentional states, but argued that belief states get their intentional content from the information that they tend to be sensitive to (under cer-

[13] See Stalnaker [2006, 2015, 2019a].

> tain normal conditions). Looking back from the later perspective of Timothy Williamson's general picture of epistemology, I came to appreciate that my account of intentionality is really a version of his "knowledge first" view: belief is what would be knowledge if the relevant normal conditions in fact obtained, or to put it the other way around, knowledge is full belief when it is non-defective. (1-2)

If something like this view is right, it would cohere with the view that (common) belief robustly explains successful (collective) action to the extent that is, or approaches, (common) knowledge.[14]

Ultimately, the close connection between coordinated action and common knowledge should not seem too surprising. Just a group's merely luckily acting in the way it would if it were were coordinated doesn't suffice for it's actually being coordinated, luckily believing in the way one would if one knew that p doesn't suffice for knowing p. Knowledge is some kind of robust, non-lucky sensitivity to the facts known, as coordinated action is some kind of robust, non-lucky sensitivity of the agents in the group to each other. This sensitivity is hard to understand as coordination at all if it is not founded on a genuine informational connection between the agents in the group—the sort of connection that makes for knowledge, and for common knowledge.

7

Lederman [2017] helpfully distinguishes the formal notion of common knowledge from the intuitive notion of a proposition's being *public*. He notes that "the default position in the literature" is that these notions coincide, with the former notion serving to clarify the latter—though in this context Lederman is using "common knowledge" in a general way that is meant to encompass common belief.

Let me close with a suggestion: if the intuitive notion of publicity tracks any iterated attitude, it tracks common knowledge specifically. In cases of pure

[14] There is an interesting point of contact here with Stalnaker's notion of "common ground". Stalnaker [2002] defines the notion of the *common ground* of a conversation in terms of common belief (what is common ground in a conversation is what is common belief about what is accepted). He adjusts this definition in Stalnaker [2014], where the common ground of a conversation is identified with what is commonly accepted. Partly for reasons reviewed in this paper, however, I would argue that the notion is better defined in terms of common knowledge. (It might be defined, for instance, as common knowledge of what is accepted for the purposes of the conversation.) For where there is something worth calling "common ground", there plausibly is genuine coordination, and (if the preceding has been on the right track) what makes for genuine coordination is knowledge.

common belief, like that of Alice and Carl above, it is not especially natural to describe the Gettiered common beliefs as *public* (for the group in question). Here is another example: we are completely in different parts of the country, but I suddenly hallucinate you approaching me, while you suddenly hallucinate me approaching you. Fill in the facts so that our simultaneously hallucinating is sheer coincidence, and so that neither of us realizes we are hallucinating. We each start talking with what we take to be the other. The details of the scenes we each hallucinate, and of the "conversations" we are having, are entirely different. In this case we will have the (pure) common belief that we are having a conversation. Is this proposition *public* for us? It seems strange at best to describe it that way. Plausibly the weirdness traces to the fact that the individual states that make for our common belief lack any common explanation. "Public" brings to mind a picture of the agents jointly confronting something in some kind of symmetrical way. Such an arrangement is the sort that makes for non-lucky coordination of states, the sort that makes for a common informational connection to each other and to the environment—that is, it is the sort that makes for common knowledge. I don't want to overstate the importance of predicting ordinary intuitions about "public", which are bound to be vague.[15] Still, it seems to me that our intuitions about the conditions under which some information is public largely support the idea that it is common knowledge, rather than common belief or any other iterated attitude, that makes for publicity. If that is right, publicity is perhaps another thing that common belief doesn't much help us to explain.

References

Ronald Fagin, Joseph Y. Halpern, Yoram Moses, and Moshe Y. Vardi. *Reasoning about Knowledge*. MIT Press, Cambridge, MA, 1995.

Ronald Fagin, Joseph Y. Halpern, Yoram Moses, and Moshe Y. Vardi. Common knowledge revisited. *Annals of Pure and Applied Logic*, 96(1-3):89–105, 1999.

John Gibbons. Knowledge in action. *Philosophy and Phenomenological Research*, 62 (3):579–600, 2001.

[15] What if instead we are together, and unbeknownst to us, we are subject to a common, externally produced illusion as of an elephant (a hologram, say)? Is it not *public* (for us) that there's an elephant in front of us in that case? Perhaps, in which case the intuitive notion of publicity is not strictly factive. Still, this is a case where there is common knowledge that we believe there is an elephant before us. This common knowledge is grounded in a symmetrical sensitivity to a shared environment, an environment where we (correspondingly) exhibit a robust coordination of states. Hence it is not a case where the sense of publicity at issue couldn't ultimately ground out in common knowledge.

Daniel Greco. Iteration principles in epistemology I: Arguments for. *Philosophy Compass*, 10(11):754–764, 2015.

Daniel Greco. Safety, explanation, iteration. *Philosophical Issues*, 26(1):187–208, 2016.

Daniel Greco. *Idealization in Epistemology: A Modest Modeling Approach*. Oxford University Press, 2023.

Jane Heal. Common knowledge. *The Philosophical Quarterly*, 28(111):116–131, 1978.

Ethan Jerzak. Two ways to want? *The Journal of Philosophy*, 116(2):65–98, 2019.

Harvey Lederman. Common knowledge. In Marija Jankovic and Kirk Ludwig, editors, *The Routledge Handbook of Collective Intentionality*, pages 181–195. Routledge, 2017.

Harvey Lederman. Two paradoxes of common knowledge: Coordinated attack and electronic mail. *Noûs*, 52(4):921–945, 2018a.

Harvey Lederman. Uncommon knowledge. *Mind*, 127(508):1069–1105, 2018b.

Dov Monderer and Dov Samet. Approximating common knowledge with common belief. *Games and Economic Behavior*, 1:170–190, 1989.

Sarah Moss. Epistemology formalized. *Philosophical Review*, 122(1):1–43, 2013.

Sarah Moss. *Probabilistic Knowledge*. Oxford University Press, 2018.

Robert Nozick. *Philosophical Explanations*. Belknap Press, 1981.

Cédric Paternotte. The fragility of common knowledge. *Erkenntnis*, 82:451–472, 2017.

Duncan Pritchard. *Epistemic Luck*. Clarendon Press, 2005.

Ernest Sosa. How to defeat opposition to moore. *Philosophical perspectives*, 13:141–153, 1999.

Robert Stalnaker. Common ground. *Linguistics and Philosophy*, 25(5-6):701–21, 2002.

Robert Stalnaker. On logics of knowledge and belief. *Philosophical Studies*, 128(1): 169–199, 2006.

Robert Stalnaker. *Context*. Oxford University Press, Oxford, 2014.

Robert Stalnaker. Luminosity and the KK thesis. In Sanford C. Goldberg, editor, *Externalism, Self-Knowledge and Skepticism*, pages 17–40. Cambridge University Press, 2015.

Robert Stalnaker. Contextualism and the logic of knowledge. In *Knowledge and Conditionals*, pages 129–148. Oxford University Press, 2019a.

Robert Stalnaker. *Knowledge and Conditionals: Essays on the Structure of Inquiry*. Oxford University Press, 2019b.

Richmond H Thomason. Common knowledge, common attitudes and social reasoning. *Bulletin of the Section of Logic*, 50(2):229–247, 2021.

Peter Vanderschraaf and Giacomo Sillari. Common Knowledge. In Edward N. Zalta and Uri Nodelman, editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Fall 2022 edition, 2022.

Timothy Williamson. *Knowledge and its Limits*. Oxford University Press, Oxford, 2000.