Uncommon Knowledge*

Harvey Lederman

Tuesday 15th November, 2016
Forthcoming in *Mind*

Abstract

Some people commonly know a proposition just in case they all know it, they all know that they all know it, they all know that they all know that they all know it, and so on. They commonly believe a proposition just in case they all believe it, they all believe that they all believe it, they all believe that they all believe that they all believe it, and so on. A long tradition in economic theory, theoretical computer science, linguistics and philosophy has held that people have some approximation of common knowledge or common belief in a range of circumstances, for example when they are looking at an object together, or when they have just discussed something explicitly in conversation. In this paper, I argue that people do not have any approximation of common knowledge or common belief in these circumstances. The argument suggests that people never have any approximation of common knowledge or common belief.

*Frank Arntzenius, Alex Paseau, Richard Pettigrew, and Timothy Williamson provided detailed comments on earlier drafts of this paper. Audiences at CMU, NYU, USC, MIT, Berkeley and Columbia gave helpful feedback on the main ideas. Lengthy correspondence with Dan Greco and Bernhard Salow, and many many conversations with Jane Friedman brought important issues into focus; discussions with David Bennett, Andreas Ditter, Peter Fritz, Ben Holguin, Daniel Rothschild, Kyle Thomas, Andrei Ungureanu, and Dan Waxman helped in a number of places. I am grateful to all of these people, to two anonymous referees and to the editors of *Mind*, for their time, thought and energy. Thanks finally to Jeremy Goodman, who has pushed me to sharpen my ideas on key points throughout the life of the project.

1 Introduction

Two friends are walking together on a crowded street. As they walk, they pass a street carnival. One points it out to the other, and they stand looking at it for some time. As they are looking at it, one of them recalls to herself the sensation of wonder she experienced at a similar street carnival when she was a child.

These occurrences exhibit a contrast in what we might describe as "publicity". The fact that there is a street carnival before them is public for these two friends; it is out in the open between them. By contrast, the fact that one of them has just been remembering her childhood experience is private; it is not out in the open between them.

As the friends walk away from the carnival, the one who has been reminiscing begins to describe her memory to the other. As a result of her remarks, it becomes public that she has just been recalling her childhood experience; that fact is now out in the open between them.

There is an alluring first thought about these two examples of publicity, a thought which has been widely accepted in economic theory, in theoretical computer science, in linguistics, and in philosophy itself. The idea is that something is public to some people just in case their minds – at least in relevant respects – are open to one another. While looking at the street carnival, the pedestrians know that there is a street carnival before them. They also each know *that they each know* that there is a street carnival before them. If they reflected on the situation, they would also each know that they each know that they each know (three occurrences of "know") that there is a street carnival before them. Now perhaps as a matter of fact, the sequence does not continue beyond this – they may not know that they each know that they each know that they each know (four occurrences of "know") that there is a street carnival before them. But, the story goes, if they fail to know this, it is not for any deep reason; there are no in principle barriers to their knowing it, or indeed, to their knowing anything described by adding a finite number of repetitions of "they each know that" in front of "there is a street carnival before them". In this sense, the people's minds – at least regarding their knowledge of the street carnival, and their knowledge of one another's knowledge of the street carnival – are taken to be open to one another.

Some people *commonly know* a proposition just in case they all know it, they all know that they all know it, they all know that they all know that they all know it, and so on. They *commonly believe* a proposition just in case they all believe it, they all believe that they all believe it, they all believe that they all believe that they all believe it, and so on. The alluring first thought can be rephrased as the idea that something is public to some people just in case there are no in principle barriers to their achieving common knowledge or at least common belief of it. There are a number of different ways of making this metaphorical statement precise; this paper will be occupied with all of them, with common knowledge, as well as its relatives.

I will argue that all of these related theories of publicity are mistaken: in a paradigm case of publicity, there *are* in principle barriers to the achievement of common knowledge and common belief. The main argument can be seen as generalizing observations made by Fagin *et al.* (1995, 11.4). A variant of the argument suggests that people cannot have common knowledge of anything at all. The new arguments dramatize a flaw at the core of related theories of publicity. The idea that something is public to some people just in case their minds are open to one another, while initially attractive, on inspection is seen to demand that people be able to access others' minds as if they were their own. But – supposing they do not enjoy some special psychic connection – even those for whom a great deal is public do not enjoy this free access to one another's minds.

The plan of the paper is as follows. Section 2 states what I count as views which analyze publicity in terms of common knowledge and its relatives. Section 3 then presents the example of SAILBOAT, and develops an argument which shows that it is a counterexample to these analyses of publicity. Section 4 shows that the main arguments are not "paradoxical", by defending the viability of the position that people never have common knowledge or its relatives (further discussion can be found in Lederman (forthcomingb)).

An appendix provides technical details in support of the main text. A.1 presents a formalization of the arguments based on SAILBOAT. A.2 shows formally that, unlike the premises of a structurally similar argument due to Timothy Williamson (1992, 2000, Ch. 5), the premises of my argument are consistent with subjects' having complete knowledge of their *own* minds. A.3 considers whether one could resist the main argument by endorsing a "Lockean" theory of belief (sometimes called "p-belief"), according to which belief is identified with probability above some threshold p.

2 Common knowledge and its relatives

In discussing common knowledge and its relatives, it will be convenient to have some abbreviations. Some people *mutually know* (or: mutually know[1]) that p just in case they all know that p. Since everyone in my family knows that I have two siblings, the members of my family mutually know^1 that I have two siblings. Progressing further, some people mutually know^2 that p just in case they all know that they all know that p. Since everyone in my family knows that everyone in my family knows that I have two siblings, the members of my family mutually know^2 that I have two siblings. And we can continue: some people mutually know^3 that p just in case they all know that they all know that they all know that p. Once again, since everyone in my family knows that everyone in my family knows that everyone in my family knows that I have two siblings, the members of my family mutually know^3 that I have two siblings.

Extending this pattern, in general some people mutually know^n that p just in case they mutually know that they mutually know^{n-1} it. We can then provide a compact definition of common knowledge: some people commonly know that p just in case for all n, they mutually know^n it.[1] Common belief can also be defined in a similar way. Some people *mutually believe* (or: mutually believe^1) that p just in case they all believe it. Some people mutually believe^n that p just in case they all believe that they mutually believe^{n-1} it. Some people commonly believe that p just in case for all n, they mutually believe^n it.

Earlier, I used the word "public" to describe the contrast between the private recollections of the pedestrian on the one hand, and the fact that there is a street carnival on the other. Some authors instead use "common knowledge" for this pretheoretic notion of publicity (e.g. Heal (1978), Barwise (1988)). For these authors it makes sense to consider different definitions of common knowledge. In the social sciences, however, "common knowledge" has come to be the standard term for the precise, technical notion defined above. Following in this tradition, I will use the terms "common knowledge" and "common belief" only in the technical sense.[2]

[1]Some authors, following Schiffer (1972) use "mutual knowledge" for what I call "common knowledge" (in some cases "mutual knowledge" is used instead for common knowledge between two people). The definitions I am using have now become standard; see, e.g. Fagin *et al.* (1995). When speaking English, it is natural to use a quantificational idiom to define common knowledge, as I have done in the main text. But in epistemic logic common knowledge is typically formalized as an infinitary conjunction, and this will be the way I speak of it officially in more formal contexts in the paper.

[2]The noun phrase "common knowledge" is of course used in everyday English, but the various formal analyses of the vague notion of "common knowledge" (that is, "public information") were never intended as analyses of this English phrase. The core sense of "common knowledge" in English appears to be roughly "known by a typical member of a relevant community". Strikingly, something may be "common knowledge" in the English sense without being mutual knowledge (or even mutual belief) in the technical sense. Wikipedia gives the following examples of common knowledge: "Julius Caesar was a Roman"; "Dallas is in Texas"; "A tall spire sits atop the Empire State Building"; "German is the primary language in Germany". Americans score notoriously poorly on such general knowledge questions; perhaps

To describe the pretheoretic notion of publicity, by contrast, I will use the expression “public information”. Officially, I will use unary sentential operators such as “some people have public information that” to express claims about public information; thus, for example, the pedestrians have public information that there is a street carnival before them. This use of “have public information that” is intended merely as a convenient abbreviation; in using it I am not implying that there is a psychologically unified phenomenon which is common to all of the examples of public information. Indeed, although the proponents of common knowledge and its relatives are committed to the idea that there is a psychologically unified phenomenon exhibited in these examples, I myself will later deny that there is. The arguments which follow will rely only on modest claims about the target notion of public information, which all proponents of common knowledge or its relatives would accept. In particular, they will depend only on judgments of what is public in cases which are closely related to the first examples in the paper: cases of public information acquired on the basis of vision (as in the case of the street carnival) and of public information acquired on the basis of audition plus comprehension (as in the conversation the pedestrians had while walking away).

A simple analysis of public information says that some people have public information that p just in case they commonly know that p. A standard objection to this analysis is that, because having common knowledge or common belief would require that people know or believe an infinite set of propositions, the claim that people have common knowledge or common belief is psychologically implausible. The Stanford psycholinguist Herbert Clark, for example, writes: “CG-iterated [that is, common belief] obviously cannot represent people’s mental states because it requires an infinitely large mental capacity...” (1996, 95-6). But this objection does not carry the weight it has been thought to carry. For all $n > 0$, I know that there are not exactly n unicorns. I know an infinite set of propositions concerning the non-existence of unicorns; any theory of belief and knowledge must account for this datum. The way in which subjects who had common knowledge and common belief would know or believe infinitely many propositions does not appear to be importantly different from the way in which I know infinitely many propositions about the non-existence of unicorns. So contrary to what Clark and others have suggested, the claim that people do have common knowledge and common belief is compatible with a range of theories about the character of knowledge and belief.

But whatever one thinks about the plausibility of the idea that people have

none of these claims is mutual knowledge even among high-school graduates in the United States.

common knowledge, in what follows I will be arguing against a much weaker, more widely accepted claim. This claim is that the publicity in my initial examples is explained by some relative of common knowledge; in particular, the claim is that public information satisfies one of the following:

IDEAL COMMON KNOWLEDGE: Necessarily, if some agents have public information that they are ideal reasoners, then if they have public information that p, they commonly know that p.

IDEAL COMMON BELIEF: Necessarily, if some agents have public information that they are ideal reasoners, then if they have public information that p, they commonly believe that p.

Throughout the paper, the English "if...then" in displayed premises and definitions, as well as in semi-formal argumentation, should be read as standing for the material conditional. In the principles just stated, "necessarily" has scope over the rest of the sentence, and is intended as a strong alethic modality, such as metaphysical necessity. Note that I count analyses which satisfy IDEAL COMMON BELIEF as invoking "relatives of common knowledge".

These theses are weak because they apply only to subjects who have public information that they are ideal; they do not require that ordinary people have common knowledge. But the theses are still not vacuous. Proponents of the idea that people exhibit some approximation of common knowledge are plausibly committed to the claim that there could be agents who have public information that they are ideal. So long as this is possible, the principles have some bite.

Every systematic study of public information I'm aware of can be understood as subscribing to IDEAL COMMON KNOWLEDGE or IDEAL COMMON BELIEF.[3] In

[3]Margaret Gilbert comes closest to advocating precisely the first of these theses (Gilbert (1989, p. 186-197)). But as I understand them, they are entailed by a range of approaches to common knowledge which do not explicitly have the "infinitely iterated" form, for example: Lewis (1969) (cf. Cubitt & Sugden (2003) for discussion); Harman (1977, p. 422) (cf. Harman (1974, p. 225)); Heal (1978); Milgrom (1981); Clark & Marshall (1981); Mertens & Zamir (1985); Barwise (1988). Lismont & Mongin (2003) give a clear, simple presentation of some definitions in this family. None of these alternative definitions entails that if normal people have public information that p, then they have common knowledge that p. But all of the authors just cited explicitly claim as an important consequence of their analyses that, under suitably idealized conditions, if the agents in question have public information that p, they have common knowledge that p. Some of these definitions may not satisfy IDEAL COMMON KNOWLEDGE as stated above but only a variant condition where "have public information that they are ideal" is replaced with "have common knowledge that they are ideal." The argument below would also work against views which entail this variant condition. Since all of these authors emphasize that for ideal agents their definitions collapse into common knowledge,

the next section, I'll present a counterexample to these claims. I'll develop my arguments in detail only as applied to IDEAL COMMON KNOWLEDGE. Indeed for the rest of the paper, in the main text I'll speak primarily about knowledge, and very little about belief. Slight variants of my arguments apply to IDEAL COMMON BELIEF, as well, but because of constraints on space I'll only sketch those variants.

3 Sailboat

> SAILBOAT: Roman and Columba are ideal reasoners playing in a game show. Each contestant has a single button on a console in front of him or her. They have an unobstructed view of each other's faces, and of an area in the middle of the stage, where the hosts will place a sailboat. First, the hosts will bring out a toy sailboat (the "test") with a 100cm tall mast. They will then replace it with a sailboat chosen randomly from an array of sailboats of various sizes. If the mast of the new sailboat is taller than the test and both players press their respective buttons, they receive \$1,000 each. If the mast is not taller than the test, and both press, or if only one person presses their button, the person or people who pressed must pay the show \$100. Today, the mast of the chosen boat is 300cm tall.

Since the mast is obviously taller than the test, it is a paradigm example of publicity:

PUBLIC > 100: Roman and Columba have public information that the mast is taller than 100cm.

But I will argue that if Roman and Columba have public information that they are ideal reasoners, and public information of various facts about their visual systems, then they will not commonly know that the mast is taller than 100cm. Thus theories which satisfy IDEAL COMMON KNOWLEDGE are inconsistent with PUBLIC > 100. Since the example is a paradigm example of publicity, it is a counterexample to this theory.

The basic idea behind the argument is simple. Let "for all Roman knows, p" abbreviate "Roman does not know that not p" (and similarly for Columba).

they should admit the possibility that ideal agents could have common knowledge that they are ideal, and not merely public information that they are.

There is some amount of variation in how things appear visually to people on a given occasion. Thus, if the mast looks to be a certain height $k$cm to Roman, then for all Roman knows the mast looks to be a little shorter, $k - 1$cm, to Columba. But Roman knows that Columba is in a similar situation: he knows that *if* the mast looks to be $k - 1$cm to Columba, then for all Columba knows it looks to be $k - 2$cm to Roman. Given Roman's knowledge of this conditional, then: for all Roman knows, for all Columba knows, it looks to be $k - 2$cm to Roman. The zig-zag continues: Roman knows that Columba knows that if the mast looks to be $k - 2$cm to Roman, then for all Roman knows it looks to be $k - 3$cm to Columba. So for all Roman knows, for all Columba knows, for all Roman knows, it looks to be $k - 3$cm to Columba. And so on. We can repeat these steps as many times as we like. Under plausible assumptions it follows that Roman and Columba do not commonly know that the mast looks to be taller than 100cm to either of them, and, moreover, that they do not commonly know that the mast is taller than 100cm.

The argument proceeds by "zig-zagging" from the perceptual appearances of one person to the perceptual appearances of the other, relying at each step on one person's ignorance of the other's mind. Crucially, every stage of the argument is consistent with the subjects' having perfect knowledge of their own minds; the descending step of the zig-zag merely invokes one subject's ignorance about the *other's* mind. This marks a basic conceptual difference between my argument and a structurally similar argument due to Timothy Williamson (1992, 2000, Ch. 5). I discuss the relationship between the two arguments in more detail at the end of section 3.2.

The formal argument dramatizes a simple, intuitive point. The idea that in the opening examples the pedestrians' minds are open to one another can seem plausible if we restrict ourselves to the objects of categorical perception, that is, features which are perceived as invariant even under small changes in appearance, such as being a face, or a street carnival, or a mast. It is a gripping thought, for example, that both Roman and Columba know the scene unfolding around them, and that it is part of this scene that Roman knows that the mast is a mast. But this scene that is open to view doesn't include precise facts about Roman's private mental life. Just as one pedestrian's recollection of her childhood experience was not out in the open for her friend to see, Roman's perceptual experiences are not part of what's out in the open for Columba to see. Since these facts aren't open to view, neither are facts about Roman's exact knowledge of the height of the mast. So even though Columba may know everything which is open to view to her, she still may not know exactly what Roman knows about the height of the mast. The comparatively small inexactitude in what Columba knows about what Roman knows about the height of the mast is magnified in the extreme when we move to higher and higher levels of mutual knowledge.

I will now present this argument precisely.

3.1 The premises

Roman and Columba are ideal reasoners. A credible public announcement has been made to both of them stating that they are ideal, so that they have public information that they are ideal reasoners. Thus, given IDEAL COMMON KNOWLEDGE, if they have public information that p, they have common knowledge that p.

The argument against IDEAL COMMON KNOWLEDGE now has three premises. The first premise concerns the agents' logical abilities. It merely makes explicit one aspect of the way in which Roman and Columba are ideal reasoners. Since the fact that they are ideal is public information, it is, according to IDEAL COMMON KNOWLEDGE, common knowledge. Thus:

> CK PERTINENT CLOSURE: Roman and Columba commonly know that for any finite set of pertinent sentences S, if for each 's' in S, Roman knows that s and if 'p' is pertinent and follows by logic from S, then Roman knows that p,

where "pertinent" means "anything used in the argument which follows." The same principle is assumed to hold, *mutatis mutandis*, for Columba. Appendix A.1 is dedicated to stating the exact closure properties used here. For now I note only that they are "multi-premise" closure principles for knowledge (and belief). In the main text, I will not discuss the possibility of rejecting these closure principles: I will assume that even if they should be rejected in full generality, they are reasonable idealizations in the present case. In Appendix A.3, however, I consider the issue in more detail. I there discuss one principled way of rejecting multi-premise closure, based on a "Lockean" theory of belief, where belief is identified with confidence above some threshold p. I argue that while this may allow for a response to one version of the present argument, the response concedes too much. Roughly, the problem is that arguments motivating Lockean common belief ("common p-belief") depend on the possibility of non-trivial common certainty, and the retreat to common p-belief leaves untouched a version of my argument which shows that Roman and Columba do not have any common certainty.

The second premise of the argument concerns the relationship between the players' knowledge of the height of the mast and how the mast looks to them. When people look around the world, there is a way that objects look to them. There is a way the mast looks to Roman and a way it looks to Columba. The way the mast looks to each of the players includes its looking to be a certain height. Moreover, what a person can know about the height of an object by looking at it on a given occasion plausibly depends systematically on how tall the object looks to be to that person on that occasion.[4]

In some cases of known perceptual illusion, we are inclined (at least to try) not to believe that the world is the way it appears to us. But the case of SAILBOAT is not a case of known illusion. If the mast looks to be $r$cm tall to one of the players, the player doesn't believe that the mast is not $r$cm tall. Relatedly, if the mast looks to be $r$cm tall to someone, then for all that person knows it is $r$cm tall. Prior to the game, the game show hosts instruct Roman and Columba about these matters, teaching them publicly how the way an object looks affects what they can know about its height. Thus these facts about visual perception become public information for both Roman and Columba. Since these facts are public information, and since Roman and Columba have public information that they are ideal reasoners, IDEAL COMMON KNOWLEDGE implies:

> CK NO KNOWN ILLUSION: For all r, Roman and Columba commonly know that if the mast looks to be $r$cm tall to one of them, then for all that person knows, it is $r$cm tall.

In CK NO KNOWN ILLUSION, it is assumed that how tall something looks to be to a person can be indexed by a precise real value in centimeters. This indexing requires some correspondence between real values in centimeters and the way objects look. But this correspondence may map pairs of distinct real numbers to one and the same way the object looks.[5] For example, it may be that if an object looks to be 299cm tall to Columba it also looks to be 300cm tall to her. Some might find it more natural to describe this situation using intervals of real numbers: for example in this case it might be that the mast looks to be between 290cm and 310cm to Columba. If one prefers to use intervals in this way, then in what follows, one can reinterpret ⌜the mast looks to be $r$cm tall to Columba⌝ as ⌜$r$cm is in the interval of heights the mast looks to be to Columba⌝.

[4]For helpful taxonomy and discussion of English expressions related to "look", see: Breckenridge (2007).

[5]The correspondence could also map different "looks" to the same height.

I will assume that one can compare how tall the mast looks to be to different people: the mast may look to be taller to Roman than it looks to be to Columba.[6] I myself find it plausible that the ⌜looks to be $r$cm tall⌝ ascriptions of folk-psychology are comparable in this way. But the expression "looks to be" can also be understood as a placeholder for any mental state or even brain state which it makes sense to describe as accurately or inaccurately registering the heights of objects in the environment. There are many ways of filling in this placeholder which should make the possibility of interpersonal comparisons uncontroversial. I will describe one such alternative here, to illustrate this point. In principle it is possible to determine empirically, for each height of the mast and each person, a probability distribution over brain states that the person could be in when looking at the mast. Presumably within each individual we could (moreover) identify some relevant features of these brain states which correspond to their response to the mast (as opposed to, for example, wondering whether they left the stove on). Our original probability distribution induces a distribution on these relevant features, as well. We can then use this induced probability distribution to put the relevant features into correspondence with the heights of the mast. For example: if one of these relevant features of the brain state is most likely to be instantiated given that the mast is 300cm, we may index that feature by "looks to be 300cm". (If there are ties, we fix some unambiguous way of breaking them.) Once these features of brain states are put into correspondence with actual heights, one could use the features themselves to stand in for "looks to be" in the premises of the argument. CK NO KNOWN ILLUSION would then be roughly equivalent to the extremely plausible claim that if the subject exhibits the feature which is most probable (if there is one) when the mast is $r$cm tall, then for all they know, the mast is $r$cm tall. Going one step further: on the assumption that a relevant notion of perceptual content supervenes on these relevant features of the brain states, one could put perceptual contents directly into correspondence with heights of external objects via the features of the brain states.

In the remainder of the paper, I'll continue to speak in terms of how things look to be. But I invite readers who are skeptical of my assumptions about how things look to be to replace this locution with an alternative such as the one just described. The argument can be run given a large number of reasonable choices of mental states or brain states to stand in for "looks to be". The motivation for CK NO KNOWN ILLUSION is the idea that for all the subject knows, the relevant aspects of her mind or her brain have accurately registered the heights

[6] I also assume we can make comparisons of the degree of difference between different pairs of such looks.

of objects in the environment.[7]

The third and final premise of the argument concerns the agents' knowledge of one another's minds. Each player's knowledge of how things look to the other person is systematically related to how tall the mast looks to be to him- or herself. For example, if the mast looks to be $r$cm tall to Roman, then for all Roman knows it also looks to be $r$cm tall to Columba. But there is also considerable variation in how the same object looks to different subjects, and even how things look to the same subject on different occasions. So if the mast looks to be $r$cm to Roman then for all Roman knows it looks to be somewhat shorter, $r - \epsilon$cm (where ϵ may be very small) to Columba.

Psychophysicists use the notion of a "just noticeable difference" (JND) to parametrize subjects' accuracy in assessing intensities and magnitudes. For example, a JND at x% for a particular continuous magnitude is the difference between two values at which subjects are x% accurate in assessing which of the two is greater. Standard values of x in this context are $75 - 80$%. The exact empirical values for the scenario described above will not be essential to my argument; so long as the value for one JND is positive and more or less constant over the range of heights I consider, the argument will go through. But for simplicity I will present the argument using a concrete, constant value for one JND; I will suppose that subjects are 80% accurate at assessing whether the mast or the "test" is taller when the objects differ in height by 3%.[8]

Exhibiting 80% accuracy in a particular kind of judgment is generally in-

[7] Those who are still mistrustful of "looks to be" (even when understood as a placeholder) may find it easier to consider the remainder of the argument in terms of an alternative version of the scenario described above. We can imagine that Roman and Columba are tasked with writing down estimates in centimeters of the height of the mast; they receive some small monetary prize for writing down the value which is closest to the true height. Moreover, they are told that they will win $1000 if and only if (a) the mast is taller than the test mast; (b) their partner writes down a number taller than the test value; and (c) they both press their buttons. If any of these three conditions fails and they press their own button, they lose $100. In this variant on the original set-up, the relevant version of CK NO KNOWN ILLUSION would be:

> CK SINCERE GUESS: For all r, Roman and Columba commonly know that if one of them writes down that the mast is $r$cm tall, then for all that person knows, it is $r$cm tall.

The argument of the next subsection could also be run using this alternative set-up, and some may prefer that alternative version. I state the remaining premise of this version of the argument below in 10. Thanks to Ben Holguin here.

[8] As far as I am aware there have not been studies of a paradigm exactly matching ours, but in a somewhat similar paradigm Ernst & Banks (2002) found that subjects had a 4% JND at 84%, while Lu *et al.* (2009) report an approximately 2.3% JND at 75%. The figure in the main text is at least in the ballpark of the empirical values suggested by these studies.

consistent with knowing the content of the judgment. For example, one doesn't know that a coin biased 80% to heads will come up heads when flipped. So this figure for a JND at 80% suggests that if the mast looks to be $r$cm tall to Roman, then for all Roman knows it is 3% shorter than r, that is, for all Roman knows it is $r - .03r = .97r$cm tall. This seems natural given that 20% of the time Roman will judge an $r$cm mast to be shorter than a $.97r$cm mast.

But if the mast is in fact $.97r$cm tall (which, for all Roman knows is the true situation) then for all Roman knows, it looks that very height to Columba. Columba doesn't always get things exactly right, but for all Roman knows, today the mast looks to her to be exactly the height it in fact is. So:

INTERPERSONAL IGNORANCE: For all r, if it looks r to one of the agents, then for all that agent knows, it looks $.97r$ to the other.[9]

As I have said, the argument which follows could be run using any $0 \leqslant \epsilon < 1$ in place of .97 (and thus any $0 < j \leqslant 1$ in place of 3%), but I'll continue to use this concrete estimate. We suppose that prior to the game, the game show hosts consult with psychophysicists and then publicly teach the players relevant facts about the relationship between perceptual appearances, rates of error, and what they can know about one another's minds. After this public instruction, these facts become public information for Roman and Columba. Since these facts are public information and since Roman and Columba have public information that they are ideal reasoners, IDEAL COMMON KNOWLEDGE implies:

CK INTERPERSONAL IGNORANCE: For all r, Roman and Columba commonly know that if the mast looks $r$cm tall to one of them, then for all that player knows, it looks $.97r$ to the other.[10]

[9]Two observations: First, in general, the behavior of the psychometric function is less regular with extreme stimuli: in our case, very tall objects, or very short objects. (This is best documented for extreme intensities, although it seems also to hold for qualitative differences: see the textbook of Gescheider (1997, Ch. 1) for discussion.) For example the value of a JND seems to become larger for weak stimuli, so that fewer "steps" are required to move an equal percentage of the stimulus. The argument here won't require assumptions about the behavior of the psychometric function anywhere near heights of 0cm: the argument can still be run if we replace "For all r" with "For all $r > 100$" throughout. Second, the precise value of a JND can be sensitive to what might seem to be irrelevant aspects of the situation. For example, it may be that the value for one JND would be smaller if subjects were asked to say whether two objects standing side-by-side were the same height. But even in that situation the value would be positive, and the argument can still be run.

[10]For the alternative version of the scenario in n. 7, it is plausible that both players will write down natural numbers (or at least, not write down long decimals). Thus a plausible analogue of this premise would be:

The premise CK INTERPERSONAL IGNORANCE does *not* require that the agents be ignorant of their own minds. The premise doesn't say anything about what Roman knows about how things look to him or about what Columba knows about how things look to her. It only describes constraints on what Roman knows about Columba, or what Columba knows about Roman.

In CK INTERPERSONAL IGNORANCE, the figure .97, corresponding to an estimated 3% for a JND, occurs inside the scope of "it's common knowledge that". It is not just public information that visual "looks" are subject to variation of some kind or other. There is also a particular value such that it's public information that looks are subject to variation parametrized by that value.

3.2 The argument

Using CK PERTINENT CLOSURE, CK NO KNOWN ILLUSION and CK INTERPERSONAL IGNORANCE, and supposing that there is some way the mast looks to one of the agents, we can show that if PUBLIC > 100 is true then IDEAL COMMON KNOWLEDGE is false. Slightly more carefully, PUBLIC > 100 can be written as:

PUBLIC > 100: For all $m \leqslant 100$, Roman and Columba have public information that the mast is not $m$cm tall.

For concreteness we suppose

$L(300, R)$: The mast looks to be 300cm tall to Roman.

From this supposition, we argue by induction to the negation of IDEAL COMMON KNOWLEDGE.

From $L(300, R)$, it follows by CK INTERPERSONAL IGNORANCE and the fact that if the agents commonly know that p, then p that:

$(.97^1)$: For all Roman knows, the mast looks to be $.97^1 \cdot 300$cm tall to Columba.

This is the base case of an induction. It is the first "zig". Next we turn to an induction step. The induction hypothesis says that for some k our premises have allowed us to zig-zag k times. Using "for all the agents mutually know^n, p" as an abbreviation for "the agents do not mutually know^n that not p", this induction hypothesis is:

CK SINCERE GUESS: For all $r \geqslant 100$, Roman and Columba commonly know that if one of them writes down that the mast is $r$cm tall, then for all that person knows, the other player has written down $r - 1$cm.

($.97^k \cdot 300$): For all the agents mutually know^k, the mast looks to be $.97^k \cdot 300$cm tall to one of them.

For the induction step, we have to show that we can always "zag" one step further. In other words we need to show that the hypothesis, together with the premises of the argument, implies:

($.97^{k+1} \cdot 300$): For all the agents mutually know^{k+1}, the mast looks to be $.97^{k+1} \cdot$ 300cm tall to one of them.

Establishing this induction step takes a little work.

Given CK PERTINENT CLOSURE and the fact that if the agents commonly know that p, then p, it follows that the agents' knowledge is closed under pertinent consequence. Given closure under pertinent consequence, it is straightforward to show that if a subject knows that (if p, then q), then if for all the subject knows, p, then for all the subject knows, q.[11] CK PERTINENT CLOSURE allows us to derive an analogous principle for mutual knowledge^n:

CONDITIONAL POSSIBILITY: If the agents mutually know^n that (if p then q), then if for all they mutually know^n, p, then for all they mutually know^n, q.

Common knowledge is officially defined as an infinite conjunction. Eliminating the conjunction in CK INTERPERSONAL IGNORANCE gives us:

INTERPERSONAL IGNORANCE-N-K: For all n and k, the agents mutually know^n that (if the mast looks to be $.97^k \cdot 300$cm tall to one of them, then for all that person knows, it looks to be $.97^{k+1} \cdot 300$cm tall to the other).

Together with CONDITIONAL POSSIBILITY and the induction hypothesis ($.97^k \cdot 300$), INTERPERSONAL IGNORANCE-N-K implies that, for all the agents mutually know^k, for all one of them knows, it looks $.97^{k+1} \cdot 300$cm tall to the other. But given CK PERTINENT CLOSURE, this implies the principle ($.97^{k+1} \cdot 300$), which completes the induction.

The induction establishes that for all natural numbers $k > 0$, for all the agents mutually know^k, the mast looks to be $.97^k \cdot 300$ to one of them. We can choose k to be whatever we like; in particular, we can choose $k = 37$. Since $.97^{37} \cdot 300 \approx 97.202 < 100$, we have:

[11] A precise statement of this and the next claim, along with the closure conditions used to prove them, are given in appendix A.1.

LOOKS < 100: For all the agents mutually know^{37}, the mast looks to be $.97^{37} \cdot 300 < 100$cm to one of them.

Here we are already close to a *reductio*. At the outset one might have thought that Roman and Columba had public information that the mast *looked* to be taller than 100cm to each of them. But things get even worse. CK NO KNOWN ILLUSION, and CONDITIONAL POSSIBILITY, together with LOOKS < 100 entail that:

NOT CK > 100: For all the agents mutually know^{3}8, the mast is $.97^{37} < 100$cm tall,

which, together with PUBLIC > 100, implies the negation of IDEAL COMMON KNOWLEDGE. This completes the argument. SAILBOAT is a counterexample to IDEAL COMMON KNOWLEDGE.

The premise CK INTERPERSONAL IGNORANCE concerns each subject's ignorance of the other's mind. We could have avoided making this assumption directly and instead invoked principles about the subjects' ignorance of the exact height of the mast itself. For example, the following two principles, together with CK PERTINENT CLOSURE, can be used to run an argument which is structurally parallel to the one just described:

CK WORLDLY IGNORANCE: For all r, Roman and Columba commonly know that if the mast looks to be $r$cm tall to Roman, then for all Roman knows, it is in fact $.97r$cm tall (and similarly for Columba).

CK POSSIBLE ACCURACY: For all r, Roman and Columba commonly know that if the mast is $r$cm, then for all Roman knows, it looks to be $r$cm tall to Columba (and vice versa).

The motivation given for the premises of the main argument also motivate these alternatives.

Yoram Moses and Joseph Halpern first observed that if agents' knowledge of time is modeled at a fine level of precision, it becomes impossible for them to achieve common knowledge or common belief using time-sensitive message-passing procedures (Moses (1986), Halpern & Moses (1990), which became Fagin *et al.* (1995, Ch. 6, Ch. 11) and Fagin *et al.* (1999).)[12] The zig-zag argument can be seen as an elaboration of their basic idea. The most important difference from

[12]Halpern and Moses' argument was made famous to economists by the related "electronic mail game" of Rubinstein (1989).

a formal perspective is that my argument is conducted in the object-language, using premises about the agents' knowledge. Unlike these earlier arguments, the object-language argument does not depend on choices about how to model the agents' uncertainty; for example it does not require directly motivating assumptions about worlds or accessibility relations. Relatedly, the object-language argument shows that the formal result is not an artifact of unintended idealizations implicit in the standard models for knowledge and belief. In particular, as I will show in Appendix A.1, the premises do not require that the agents be logically omniscient, in the sense of knowing all tautologies of propositional logic.

The zig-zag argument is also structurally similar to an argument given by Timothy Williamson (1992, 2000, Ch. 5). But here the similarity is merely structural. The present argument does not rely on the claim that knowledge requires a margin-for-error. As Williamson shows, if knowledge requires a margin-for-error then in examples such as SAILBOAT there will be some proposition such that an agent with inexact perceptual knowledge will know it, although she fails to know that she knows it. Inspection of the premises of my argument (especially CK INTERPERSONAL IGNORANCE) already suggests that unlike Williamson's premises, they do not impose limitations on the agents' self-knowledge. Indeed, the most popular ways of resisting Williamson's argument are consistent with the informal motivation for the premises of my argument.[13] Even so, one might still reasonably be concerned that the premises of my argument ultimately do entail (by some subtle and difficult to discover proof) that the agents suffer from limitations on their self-knowledge. Appendix A.2 shows that this concern is not borne out, by giving a model in which the premises of the argument hold, but in which it is common knowledge among the agents that if one of them knows a proposition, she knows that she knows it. This model also demonstrates that my argument does not rely on the claim that knowledge requires a margin-for-error in Williamson's technical sense.

[13] Weatherson (2004), Berker (2008), Greco (2014a,b) and Stalnaker (2009, 2015) have suggested in different ways that there may be a "constitutive connection" between a person's knowing or believing that p, and that person's knowing or believing that she knows or believes that p. This idea is the basis of their rejection of the margin-for-error premise. But these authors do not suggest that there is a similar constitutive connection between one person's beliefs and another person's beliefs, as would be required to reject CK INTERPERSONAL IGNORANCE. And indeed this further claim is implausible. For example, it may be that for Roman to know that the mast is taller than 100cm just is for Roman to know that Roman knows that the mast is taller than 100cm. But surely it is not true that for Roman to know that the mast is taller than 100cm just is for Columba to know that Roman knows this. Columba may not know that Roman exists.

3.3 Categorical perception

I have argued that Roman and Columba don't commonly know certain facts about a continuous magnitude, the height of the mast. But the examples in the introduction did not concern continuous magnitudes. They concerned objects of categorical perception such as being a street carnival. Does the argument affect the agents' common knowledge of any facts of this kind?

It does. Principles analogous to the ones introduced above for variation in perception of heights are plausible for variation in other continuous magnitudes, such as width, curvature and color. If we run the same argument using variations in height, width and curvature simultaneously, we can zig-zag across slight variations in appearance to derive the result that Roman and Columba do not commonly know that the sailboat looks to be a sailboat to either of them. For all Roman knows, the boat looks to be slightly wider and slightly distorted in shape to Columba. For all Roman knows, for all Columba knows, it looks to be slightly wider again and slightly more distorted to Roman. And so on and so forth. At the end of this series, for all they commonly know, perhaps it looks to be simply a giant piece of canvas, hanging from a wooden pole, and not a sailboat at all. Given an analogue of CK NO KNOWN ILLUSION, for all they commonly know in this case, the sailboat is a giant piece of canvas.

The point applies to other modalities as well. In cases of public announcements, we can zig-zag across variations in auditory "appearances" of volume, pitch and modulation to derive the result that it is not common knowledge that any words were uttered as opposed to merely animal sounds. The argument thus shows that Roman and Columba do not have common knowledge of some objects of categorical perception (something's being a street carnival, or a sound's being a word) by showing that they do not have common knowledge of various facts about continuous magnitudes.

3.4 Uncommon knowledge

The argument can thus be generalized to show that, under the right circumstances, Roman and Columba may fail to have common knowledge that p, even though they have public information that p, and they perceive that p as invariant under small changes in appearance. But, we might ask, are there any p such that common knowledge that p is immune to extensions of the argument?

I will now suggest that there are not. As a warm-up to this argument, observe that the following five claims are inconsistent:

CK PERTINENT CLOSURE: Roman and Columba commonly know that for any finite set of pertinent sentences S, if for each 's' in S, Roman knows that s and if 'p' is pertinent and follows by logic from S, then Roman knows that p (and similarly for Columba);

CK WORLDLY IGNORANCE: for some constant parameter describing variations in Roman and Columba's visual perception, Roman and Columba commonly know that if Roman looks to be a certain way to Columba, then for all Columba knows, Roman's true hue, size and shape differ by the relevant constant parameter from how they appear to her;

CK POSSIBLE ACCURACY: Roman and Columba commonly know that if Roman is in fact a given hue, size and shape, then for all Roman knows, Roman looks to be that exact hue, size and shape to Columba; and

LOOKS LIKE A ROCK: Roman and Columba commonly know that if Roman looks to be a rock to Columba then for all she knows he is a rock;

ROCKS ARE IGNORANT: Roman and Columba commonly know that rocks don't know anything.

The argument for the inconsistency can be sketched schematically as follows. Let's abbreviate the predicate which expresses the property of having Roman's actual hue, size and shape by "F", and let d be a function from properties to properties which takes a hue-size-and-shape property and produces a property which is distorted by the relevant constant parameter "in the direction" of the appearance of a rock. Thus for example, one application of d distorts Roman's hue to be slightly grayer, distorts his height to be slightly smaller, distorts his shape to be squatter and lumpier, and so on. Given that Roman is F, by CK POSSIBLE ACCURACY, for all Roman knows, he looks to be F to Columba. By CK WORLDLY IGNORANCE, if for all Roman knows, he looks to be F to Columba, then for all Roman knows, for all Columba knows, Roman is in fact $d(F)$. This reasoning can be repeated. Roman knows that Columba knows that if Roman is in fact $d(F)$ then for all Roman knows, Roman looks to be $d(F)$ to Columba. So for all Roman knows, for all Columba knows, Roman looks to be $d(F)$ to Columba. But Roman knows that Columba knows that Roman knows that if Roman looks to be $d(F)$ to Columba then for all Columba knows he is in fact $d(d(F))$. And now it should be clear that we are off to the races.

We can thus show that for all Roman and Columba commonly know, Roman looks to be a rock to Columba. By LOOKS LIKE A ROCK and CK PERTINENT CLOSURE, moreover, Roman and Columba don't commonly know that Roman is not a rock: for some n, for all they mutually know^n, Roman is a rock. But now by ROCKS ARE IGNORANT and CK PERTINENT CLOSURE it follows that for some n, for all they mutually know^n, Roman doesn't know anything. And since they don't mutually know^n that Roman knows anything at all, they don't mutually know^{n+1} anything, and hence don't commonly know anything, either.

This argument shows only that the above five claims are inconsistent. But the premises can be altered slightly to show that for any p, Roman and Columba do not have common knowledge that p. There is some k, which is the exact number of applications of the distortion function d needed to arrive at a sufficiently "distant" appearance that Roman looks to be a rock. If we replace "commonly know" in the above five claims with "mutually know^k", the premises would no longer be inconsistent. But they would still entail that Roman and Columba do not mutually know^{k+1} anything at all, whether it is the height of a mast, the fact that there is a street carnival, or even the fact that one of them is conscious.

The zig-zag argument is easiest to grasp when there is only one source of perceptual information about the relevant claim. In more complex, realistic cases, however, some premise of the original zig-zag argument may be false because of what people know by other modalities, memory or testimony. It may be true that for all someone knows on the basis of vision, an object they are looking at might be a little bit shorter or a little bit grayer than it seems to be. But this variation in the object's appearance might conflict with what the person knows on the basis of what they have been told about how the object looks.

But even if we consider everything a subject knows via all modalities, there will still be what we might call a "total perceptual JND", a variation in the sum total of perceptual appearances consistent with everything learned in any way (by any modality) whatsoever. These total perceptual JNDs will no doubt be smaller and more oddly behaved than the ordinary JNDs measured by psychophysicists in familiar paradigms. But it is plausible that they exist: the information we possess on the basis of perception – even all perception taken together – isn't perfectly accurate at very precise levels of detail. If the nature of this variation becomes public information among ideal agents, then we can exploit variations in these total JNDs to move from the agents' actual perceptual appearances to quite "distant", strange appearances. If, as seems plausible, we can reach distant appearances which are inconsistent with others' being thinkers or existing at all, then it will follow that the subjects won't have common knowledge of anything at all.

Throughout this section, I've presented the arguments as they apply to IDEAL COMMON KNOWLEDGE. Parallel arguments can also be given for IDEAL COMMON BELIEF. As I show in appendix A.1, the arguments do not depend essentially on the factivity of knowledge – the fact that if one knows that p, then p. If we substitute "justifiedly believe" and its cognates for "know" and its cognates in the premises of the zig-zag argument, they remain at least equally plausible. The resulting premises can be used to run the main argument and the extensions of it. A further assumption allows to complete the argument against IDEAL COMMON BELIEF, namely, that if an ideal reasoner believes p, they justifiedly believe p. For then the argument that Roman and Columba do not have common justified belief that the mast is taller than 100cm would show that they do not have common belief that it is taller than 100cm, either.

The arguments of this section do not demonstrate that people never have common knowledge or its relatives. They require demanding auxiliary premises about what is public information among the relevant people. Plausibly people do not often have public information of this kind, if they ever do. But the arguments show that the motivation for the claim that we have common knowledge or its relatives is based on a basic mistake. The motivation for this claim was that in paradigm cases of public information our minds could be open to one another: there is no barrier to our attaining ever higher levels of mutual knowledge, whether or not we do attain them. When we consider only the objects of categorical perception, this "alluring first thought" can seem plausible. But the stylized example of SAILBOAT, and the formal argument based on it illustrate an intuitive, quite general problem with this idea. In the perception of continuous magnitudes, our minds are simply not open to one another in the way the alluring first thought claims that they are. As we have just seen, this basic problem with continuous magnitudes infects all claims whatsoever, and thus casts doubt on the idea that people ever have common knowledge or its relatives.

4 Coordination without common knowledge

The arguments of the previous section at least present a puzzle about how ideal agents who have inexact perceptual systems could have common knowledge, and thus how people could have common knowledge or any of its relatives. But I am inclined to take the arguments to be more than a puzzle. In my view, the arguments motivate exploring the hypothesis that people never have common knowledge or its relatives.

The arguments are not general paradoxes which afflict every theory of public information. A family of principles governing public information which escape the arguments can be produced using the following schema:

NECESSARY MK^n: Necessarily, if some agents have public information that p, then they have mutual knowledge^n that p.

For any $n < 37$, given the concrete values I assumed in the main argument in Section 3.2, a theory of public information which satisfies NECESSARY MK^n need not entail the negation of PUBLIC <100. Plausibly the argument in Section 3.4 requires considerably more steps than 37, so that argument could also not be run against NECESSARY MK^n when $n < 37$.

But even granting that some reasonable views escape the arguments of the previous section, one might wonder whether the hypothesis that people never have common knowledge or its relatives could be viable. A first concern is that if we abandon common knowledge and its relatives we cannot in the end give a satisfying theory of public information. Proponents of common knowledge typically begin their discussions by introducing readers to a putative natural class of examples where people have public information.[14] They then pose the question of what psychological features these examples share, and answer this question by invoking common knowledge or its relatives. The first concern is that, if we give up on this answer, we will no longer be able to give any satisfying answer at all.

But the concern is seen to be ill-founded once we recognize that we should anyway reject the presupposition of this question, namely, that there are some relevant psychological features which unite the examples of public information. When a simple account of these unifying psychological features was in view, this presupposition seemed reasonable. But once we reject the claim that people have common knowledge and its relatives even in apparently clear examples of public information, the claim loses its air of plausibility. There is little *prima facie* reason to think that there is one particular pattern of attitudes which people exhibit in situations as different as listening to a conversation and looking at their surroundings on a casual stroll. The motivating examples for the notion of public information are so limited that it is not even clear how to identify new examples of the phenomenon, or how to classify even mildly "hard" cases, as would be needed to assess a conceptual analysis of the notion. We cannot turn to natural language for help here, since "public information" (or "common knowledge" in its "informal" sense) is a technical term. People can invent technical terms as they like, but we should not expect that invented technical terms will invariably track interesting facts about psychology. We should also not expect that they will always admit of interesting conceptual analysis.

[14]Usually a different term, most often "common knowledge" itself, is used. See n. 2 and surrounding main text, above.

It is worth emphasizing that this view of public information is consistent with my use of the notion in the arguments against common knowledge. Any proponent of common knowledge or its relatives will hold that when something is out in the open, or overt, or public, the people in question have common knowledge or its relatives. In particular they will accept that this is so in the examples from the introduction and in SAILBOAT. The minimal assumptions I made about public information are thus admissible in an *ad hominem* argument, independently of my own view of the psychological disunity of the examples of public information, and my suspicion that the notion does not admit of interesting conceptual analysis.

Still, there is a second, independent concern about the viability of the hypothesis that people never have common knowledge or its relatives: that these states play an important role in explaining social behavior, so that rejecting them would leave us unable to explain such behavior. Some have claimed, for example, that common knowledge is needed to explain the fact that it would be rational for Roman and Columba to press their buttons in SAILBOAT or at least in situations closely related to it.[15] If this claim were correct, it would imply that the hypothesis that people never have common knowledge or its relatives is untenable.

But the claim is incorrect, as I will now show, by explaining the rationality of coordination in situations such as SAILBOAT without appealing to common knowledge or its relatives. I will start by giving a theory of coordination in these situations which should be acceptable to proponents of common knowledge and more generally to those who believe that examples of public information are united by relevant psychological commonalities. I will then show how the theory can be altered to eliminate the use of common knowledge and in fact even to eliminate the appeals to public information.

[15] See, for example: Fagin *et al.* (1995, p. 198, 454) (cf. Halpern & Moses (1990, pp. 2,19)), Chwe (2001, p. 96). (For more references, and discussion of some famous formal arguments, see Lederman (forthcomingb, esp. n. 24). For discussion of more conceptual arguments, see Lederman (forthcominga).) In philosophy, the most notable proponent of a similar position has been Dan Greco, who has recently argued that if an epistemological theory is inconsistent with the existence of common knowledge, it would force us to reject explanations from the social sciences, which counts as "strong theoretical motivation" for rejecting the epistemological theory in question (Greco (2014a,b)). I don't see the strong theoretical motivation. First, models which use common knowledge in the social sciences are not well confirmed empirically; if anything they are disconfirmed (see again Lederman (forthcomingb)). Second, there are good explanations of why such models have become prominent which don't have to do with their accuracy: elegant and famous results which use common knowledge have led to a proliferation of models which explore these results using similar assumptions. Third, common knowledge isn't required to give an explanation of the phenomena, as I will now show.

The rationality of coordination depends in the first instance on knowing what others will do.[16] If Roman knows – by whatever means – that Columba will press her button, then since he knows that the mast is taller than the test, he knows that if he presses he will win $1000. If he does not press, he will gain nothing. So, in these circumstances, he should press the button.

The example of SAILBOAT is simple enough that, barring general skepticism about the future, it is clear that Roman knows that Columba will press her button. To explain the rationality of coordination, then, we must simply explain how he knows this. There are many theories one might give which do not feature common knowledge or public information; deciding between them would require substantial empirical work. My aim here is just to describe one such theory.

People can know what others will do on the basis of past experience of others' behavior in similar situations. We can provisionally describe some of these cases using the notion of public information. Most people are often in situations where it is public information that a certain pattern of actions will be best for everyone involved. When two people are walking towards each other on the street, it is public information that it would be best for each of them not to run into the other. When one's neighbor at dinner passes the soup, it is public information that it is best for all involved that the soup not spill. People tend not to run into each other on the street, and tend not to drop the soup without warning. More generally: if it is public information that a particular pattern of actions would be the best of the relevant outcomes for each party involved, people tend to do their part in that pattern of actions. In SAILBOAT it is public information that a particular pattern of actions will be best for each party involved. So people can reasonably infer that in this case, as in the past, others will do their part in the relevant pattern, that is, that others will press their buttons.[17]

Proponents and opponents of common knowledge alike will agree that expe-

[16] As usual the reader who is so inclined is invited to substitute "justifiedly believe", "believe", or "have sufficiently high confidence" where I use "know" in what follows.

[17] This explanation can be subsumed under a more general one if pressing the buttons is deemed "salient". The classic treatment of the importance of salience to coordination is Schelling (1960). Lewis (1969, 36-42) builds on Schelling to offer an account of coordination which is largely in the spirit of the present discussion. Lewis does not invoke common knowledge to explain behavior (nor does he invoke what he calls "common knowledge", which is not common knowledge, see Cubitt & Sugden (2003)); he is clear that higher-order expectations are unnecessary for explaining what people do. On my suggested picture, rational people would be responsive to salience because inductively supported generalizations can be stated using salience: people have tended to do their part to achieve the most salient good outcomes in the past, so they will do so in the present case as well. Perhaps an outcome is most salient of a set of outcomes just in case it most attracts normal people's attention.

rience plays a role in how we know what others will do.[18] All sides should hope to go beyond the simple description just given, to state exactly which features of new situations people respond to in assessing what others will do. Providing a detailed account of this kind is a difficult empirical problem. But the difficulty of giving such a detailed account does not cast doubt on the general form of the theory. When people encounter a doorknob which is different in shape and position from any doorknob they've encountered before, they are fairly reliable in knowing how to turn it, and whether to push or pull. When people encounter a new stove in a new house, with dials unlike any they've seen before, they usually know which dial does what. It is clear that people form beliefs about new doorknobs and stove dials in part on the basis of past experience, even if it is hard to say exactly how they do this. Similarly, it is clear that people make predictions about what others will do in new situations in part on the basis of past experience, even if it is not always clear exactly how they do this. Certainly, they have ample data to drawn on; it is just a question of which similarities they deem relevant for drawing inferences about the present case.

This story can be completed by filling in the "black box" of public information. But crucially, the story itself is not particularly sensitive to what substitution we make. For example, we could substitute mutual knowledge^4. In the past, whenever it has been mutual knowledge^4 that some pattern of actions gives everyone their best payoff, people have played their part in producing that pattern of actions. Since plausibly it could be mutual knowledge^4 in SAILBOAT that it would be best for all involved to press their button, this generalization would make sense of the belief that in this situation, others will press. This rational reconstruction does not use common knowledge, nor does it use public information.

The version of the story which uses mutual knowledge^4, however, may not be particularly realistic. In sizing up a situation, people do not often attempt to discern what others know about what they themselves know (never mind what others know about what they themselves know about what others know). Instead, people typically respond directly to heterogeneous, observable features of the situation.[19] Here is one way this could work in the present case. Two people

[18]There is no deductive argument from common knowledge of any of the background features of the case to the claim that all will press their buttons. For example, it is consistent with common knowledge of the payoffs, common knowledge of the height of the mast, and common knowledge that all are subjective expected utility maximizers that no one press their buttons.

[19]Some recent work in psychology suggests that people's first reaction is to treat information as appropriate to use for coordination; if they are distracted or given a load while performing a task, they err by coordinating when they should not, as opposed to failing to coordinate

are *jointly attending* to an object just in case they are engaged in a pattern of eye-movements between the object and one another's eyes, first looking at the object, and then checking that the other is also looking at the object.[20] It is *obvious* that an object has a feature just in case any normal person who looked at the object would know that the object has the feature.[21] If others learn from conversation that an outcome will be best for all involved provided an object has a particular feature, if it is obvious that the object has the feature, and if all participants' alertness is confirmed by an episode of joint attention, people tend to do their part in bringing about the mutually beneficial outcome. Roman and Columba's experience of this regularity makes it reasonable for them to infer that the other will do his or her part in SAILBOAT. To predict what others will do using this generalization, neither we nor they need to rely on claims which are well-described using the notion of "public information". In fact, Roman and Columba do not need to draw any sophisticated inferences about one another's minds.[22]

---

when they should (Rand *et al.* (2014), Keysar *et al.* (2000), Keysar *et al.* (2003), Epley *et al.* (2004), Lin *et al.* (2010) and Keysar (2007)). If in their first reactions people are not so different from automata who will act to coordinate in the right external conditions, this may explain why there is little need to think about others' minds to infer that they will coordinate.

[20]Some (e.g. Peacocke (2005), Campbell (2002)) have suggested that common knowledge or one of its relatives is needed in the characterization of episodes of joint attention. Their arguments are not compelling. The cases they adduce are consistent with the hypothesis that subjects jointly attend just in case they engage in a distinctive pattern of eye-movements. This analysis of joint attention does not use any iterated knowledge or belief; it is thus not subject to the main arguments of the paper.

[21]Thanks to Jeremy Goodman for discussion here.

[22]More complex behavior, for example conversations, may require more thinking about what others think. An influential model of conversations holds that participants in a conversation coordinate on a resolution of context-sensitive expressions by coordinating on a shared body of information, the "common ground", generally supposed to be determined (roughly) by what the participants commonly know (Stalnaker (2002, 2014), cf. e.g. Pinker (2007), Pinker *et al.* (2008)). Typically, little explicit motivation is given for the claim that common knowledge should be used here; perhaps the following two ideas have been in the background. First, since the common ground is typically public information, proponents of the idea may have supposed that it would coincide with common knowledge. Second, it may have been assumed that coordination in general requires common knowledge, so that linguistic coordination in particular would also require common knowledge. Neither motivation survives the arguments of the present paper. One can coordinate on resolutions of context-sensitive expressions by coordinating on what participants mutually know^n for low n (perhaps as low as 2). This "theory" of the common ground would be consistent with the formal treatment of the common ground in linguistic semantics, which employs only the assumption that the common ground corresponds to a body of information (which in the simplest models is represented by a set of possible worlds). Various laws of the epistemic logic of contexts which Stalnaker himself deems important would not come out valid on this model of common ground (see Stalnaker

Our beliefs, more or less accurately, reflect the world around us. The proponents of common knowledge and its relatives suppose that when two people meet their beliefs create a hall of mirrors: each person's beliefs reflect the world, but also the other's beliefs, which in turn reflect the first person's beliefs and so on, in an infinite sequence of reflections of reflections. They imagine that certain objects and events are so positioned in this hall of mirrors that these objects and events are recognizably represented in each of the reflections in this infinite sequence. They propose that people will coordinate successfully only if they manage to position their actions in this exact way in the marvelous hall of mirrors created by their minds. But our minds often reflect the world imperfectly. I have argued that the presence of slight aberrations on the surface of each mirror has the consequence that as we move further and further along this sequence of reflections, the original image becomes blurred and ultimately unrecognizable. I have sketched an explanation for why these slight errors in the way our minds reflect our actions and one another's minds are no serious impediment to what we are able to do, and even to what we are able to do together. People don't need to gaze deep into a hall of mirrors before they decide what to do. They know what has happened in similar situations before, so they can take one quick look, and they are ready to act.

A Appendix

A.1 Closure Principles and Logical Omniscience

In this appendix I state formally closure principles needed to run the main argument. I also show how this version of the argument does not require that Roman and Columba be "logically omniscient".

The argument based on SAILBOAT is easily formalized using a modal propositional logic with interpreted proposition letters. In the formal presentation, I will not assume the "factivity" of the modal operators; this will make the formal argument applicable both to knowledge and belief. In the argument in the main text the reasoners are assumed to be ideal. It is of considerable interest to see how "ideal" they must be. In giving the argument I will separate out the assumptions I use, to demonstrate that the required assumptions about ideal agents are in fact considerably weaker than assumptions implicit in standard models from epistemic logic.

Let the language $\mathcal{L}$ be given as follows:

$$L(r,i) \mid S(r) \mid \neg\phi \mid \phi \wedge \psi \mid K_i\phi \mid CK\phi$$

(2014, Appendix)), but these laws are not important to any empirically confirmed theories I am aware of.

where r is a rational number $\geqslant 0$, and i is either R or C. (We use rational numbers here instead of real numbers as a technical convenience to keep the language countable. Nothing essential hinges on this.) $L(r,i)$ is to be read as "the mast looks to be $r$cm tall to i"; $S(r)$ as "the mast is $r$cm tall"; K_i as "i knows that", and CK is to be read as "the agents commonly know that". The metalinguistic abbreviations $\vee$, $\rightarrow$ and $\Diamond_i$ are used in the standard way. Moreover, we use the following metalinguistic abbreviations for mutual knowledge: $M^1A = K_RA \wedge K_CA$; and $M^nA = M^{n-1}A \wedge M^1M^{n-1}A$. For simplicity, we assume the sole atoms of the language are sentences of the form $L(r,i)$ and $S(r)$.

The logic L^- is given by the following laws and rule:

(PL)	All theorem schemas of propositional logic.
(CKout)	$CK\varphi \rightarrow M^n_\psi$
(K)	$K_i(\varphi \rightarrow \psi) \rightarrow (K_i\varphi \rightarrow K_i\psi)$
(CONJUNCTION-OUT)	$K_i(\varphi \wedge \psi) \rightarrow (K_i\varphi \wedge K_i\psi)$
(CONJUNCTION-IN)	$(K_i\varphi \wedge K_i\psi) \rightarrow K_i(\varphi \wedge \psi)$
(CK K)	$CK(K_i(\varphi \rightarrow \psi) \rightarrow (K_i\varphi \rightarrow K_i\psi))$
(CK CONJUNCTION-OUT)	$CK(K_i(\varphi \wedge \psi) \rightarrow (K_i\varphi \wedge K_i\psi))$
(CK CONJUNCTION-IN)	$CK((K_i\varphi \wedge K_i\psi) \rightarrow K(\varphi \wedge \psi))$
(MP)	$\varphi, \varphi \rightarrow \psi \vdash_{L^-} \psi$

Note that we assume (K), (CONJUNCTION-OUT), (CONJUNCTION-IN) separately, since we have not assumed that any of the operators K_R, K_C or CK is factive. As mentioned above, this makes the proof sufficiently general that it applies to the case of belief as well as that of knowledge.

Lemma A.1. *For any n:*

1. $\vdash_{L^-} M^n(\varphi \rightarrow \psi) \rightarrow (M^n\varphi \rightarrow M^n\psi)$
2. $\vdash_{L^-} M^n(\varphi \wedge \psi) \rightarrow (M^n\varphi \wedge M^n\psi)$
3. $\vdash_{L^-} (M^n\varphi \wedge M^n\psi) \rightarrow M^n(\varphi \wedge \psi)$

The proof is by induction.

Let the logic L be given by adding the following axiom schemas to L^- and closing under the rules.

(Contra)	$K_i(\varphi \rightarrow \psi) \rightarrow K_i(\neg\psi \rightarrow \neg\varphi)$
(CKContra)	$CK(K_i(\varphi \rightarrow \psi) \rightarrow K_i(\neg\psi \rightarrow \neg\varphi))$

Lemma A.2 (Conditional Possibility)**.** *For any n:* $\vdash_L M^n(\varphi \rightarrow \psi) \rightarrow (\neg M^n\neg\varphi \rightarrow \neg M^n\neg\psi)$

Once again the proof is by induction.

For a fixed rational $0 < \epsilon < 1$, let the logic L^ϵ be formed by adding the following axiom schemas to L, and closing under modus ponens:

(INTERPERSONAL IGNORANCE)	$L(r,i) \to \Diamond_i L(\epsilon \cdot r, j)$
(CK INTERPERSONAL IGNORANCE)	$CK(L(r,i) \to \Diamond_i L(\epsilon \cdot r, j))$
(NO KNOWN ILLUSION)	$L(r,i) \to \Diamond_i S(r)$
(CK NO KNOWN ILLUSION)	$CK(L(r,i) \to \Diamond_i S(r))$

An argument exactly analogous to the one in the main text establishes the following proposition:

Proposition A.3. *For any $r > 100$, there is some $x \leqslant 100$ such that:*

$$\vdash_{L^\epsilon} L(r,i) \to \neg CK \neg S(x)$$

Every theorem of propositional logic is a theorem of this logic, but the logic does not contain the rule of necessitation (RN-K$_i$) (if $\vdash \varphi$ then $\vdash K_i\varphi$). If it did, it would be provable that the agents *know* the theorems of propositional logic no matter how complex (and indeed, given a plausible rule for the introduction of CK (if for all $n \vdash M^n\varphi$, then $\vdash CK\varphi$), that they commonly know these theorems). In other words, it would be provable that the agents are logically omniscient. But the rules of our logic are in fact quite weak: we do not have even the rule of equivalence (RE-K$_i$) (if $\vdash A \leftrightarrow B$ then $\vdash K_iA \leftrightarrow K_iB$). To see this informally, one can simply note that, on the assumption $K\varphi$, there are no laws which would allow the proof of $K\neg\neg\varphi$. More formally, it is elementary to provide a countermodel to (RN-K$_i$) and (RE-K$_i$), using a model with designated worlds at which classical logic holds, and undesignated worlds where it fails. The argument based on SAILBOAT does not require common knowledge of all tautologies of propositional logic – an assumption which was maintained in model-theoretic versions of related arguments, for example in Halpern & Moses (1990).

A.2 Consistency with Introspection Assumptions

In this section, I give a model illustrating the consistency of the premises of the argument with the strong logic **S4.2** (for standard presentations of this logic, see Chellas (1980) or Fagin *et al.* (1995); for the use of the latter in philosophy, see Stalnaker (2006, 2009, 2015), one of the main philosophical proponent of common knowledge). (I note in passing that a consistency proof of this kind can also be given for the very strong logic **S5**; I've focused on **S4.2** here because it represents a more plausible theory of belief and knowledge.)

In the model, the set of worlds is a subset of $\mathbb{Q}^3$, where the x-coordinate represents the height of the mast, the y-coordinate represents how the mast looks to Roman, and the z-coordinate represents how the mast looks to Columba. The model is a simple Kripke model, with special clauses for interpreted proposition letters. The truth clauses for complex sentences, including the modal operators in the language $\mathcal{L}$, will be given by the usual inductive clauses (for details, including the semantics of CK, see, e.g. Fagin *et al.* (1995, Ch. 3.3)); I will state only the clauses for interpreted proposition letters.

The model is for a language which adds two unary belief operators, one for each agent, to the language $\mathcal{L}$. (This is inessential to the presentation, but it will give us a richer picture of the epistemology of the case.) The model is a structure $\langle W, R_{B^R}, R_{B^C}, R_{K^R}, R_{K^C}, v\rangle$, where $W = \mathbb{Q}^3_{>0}$, and v is defined by:

- $v(S(r)) = \{\langle x, y, z\rangle \in W : r = x\}$;
- $v(L(r, R)) = \{\langle x, y, z\rangle \in W : r = y\}$;
- $v(L(r, C)) = \{\langle x, y, z\rangle \in W : r = z\}$.

Letting $k < 100$ be a positive rational constant, which parametrizes the agents' uncertainty about one another's beliefs, and $\sigma < 100$ a positive rational constant which parametrizes their uncertainty about the mast, their beliefs can be represented by the following accessibility relations:

- $R^R_B(\langle x, y, z\rangle) = \{\langle x', y', z'\rangle \in W : |x' - y| \leqslant \sigma, y' = y, |z' - y| \leqslant k\}$
- $R^C_B(\langle x, y, z\rangle) = \{\langle x', y', z'\rangle \in W : |x' - z| \leqslant \sigma, z' = z, |y' - z| \leqslant k\}$

By standard soundness results, this gives us a model of **KD45** for each individual's belief. We then use the accessibility relations for belief to define the ones for knowledge. Only the case of Roman will be presented here; the clauses for Columba's accessibility relation are obvious transformations of them:

Case 1 if $\langle x, y, z\rangle \in R^R_B(\langle x, y, z\rangle)$ then $R^R_K(\langle x, y, z\rangle) = R^R_B(x, y, z)$;

Case 2 if $|y-z| > k$ and $|y-x| \leqslant \sigma$ then $R^R_K(\langle x, y, z\rangle) = \{\langle x', y', z'\rangle \in W : |x'-y| \leqslant \sigma, y' = y, |z' - z| \leqslant |y - z|\}$;

Case 3 if $|y-x| > \sigma$ and $|y-z| \leqslant k$ then $R^R_K(\langle x, y, z\rangle) = \{\langle x', y', z'\rangle \in W : |x'-x| \leqslant |y - x|, y' = y, |z' - y| \leqslant k\}$; and finally

Case 4 if both $|y - x| > \sigma$ and $|y - z| > k$ then $R^R_K(\langle x, y, z\rangle) = \{\langle x', y', z'\rangle \in W : |x' - x| \leqslant |y - x|, y' = y, |z' - z| \leqslant |y - z|\}$.

It's easily checked that the model is a model of **S4.2** for each individual's knowledge, using the usual soundness theorems. The model is also a model of L^ϵ.

A.3 Common p-belief

According to a simple "Lockean" theory of belief, belief corresponds to confidence above some threshold p. This theory is typically motivated by examples where rational belief appears not to satisfy "multi-premise" closure, such as the preface paradox. Since my own arguments rely on a form of multi-premise, one might wonder, then, whether Lockeans can escape it. In this section I argue that, whatever its prospects for responding to the main argument of the paper, a notion of common belief based on an underlying Lockean notion of belief faces grave challenges.

An agent p-believes a proposition just in case she assigns it probability $r > p$; some agents have mutual p-belief^1 that q iff they all p-believe q; they have mutual p-belief^n that q iff they mutually p-believe^1 that they mutually p-believe^{n-1} that q. They commonly p-believe that q iff for all n, they mutually p-believe^n that q. These notions were first studied systematically by Monderer & Samet (1989).[23]

There is an important challenge to the use of common p-belief in the theory of public information, which is quite different from the challenge to common knowledge presented in the main text: the most promising arguments motivating the importance of common belief *require* multi-premise closure properties, so these formal motivations can't be carried over to common p-belief. There are some formal results in the literature which have been claimed to show that common p-belief is important in a variety of situations, but those results have what I believe is an important shortcoming: they assume a great deal of common 1-belief, that is, common certainty. To take just one example, the recent work of Dalkiran *et al.* (2012) shows that common p-belief has an intimate relationship to Bayes-Nash equilibrium. But their results use the following assumptions: that there is (a) common certainty of the payoff structure, (b) common certainty of the agents' priors, (c) common certainty that the agents are rational, and (d) common certainty that if one of the agents assigns a proposition probability p, she is certain that she assigns it probability p (which entails both negative and positive introspection for certainty).[24]

If we allow that particular claims can become not just common p-belief for $p < 1$, but also common certainty, an analogue of the argument in the main text can be run against common p-belief. In particular, the following assumptions suffice.

PROBABILISM: Roman and Columba are probabilistically coherent;

CC PROBABILISM: Roman and Columba are commonly certain that each of them is probabilistically coherent;

PROBABILISTIC INTERPERSONAL IGNORANCE: there's an ϵ such that: if the mast looks to be $r$cm to Roman, then for all Roman p-believes, it looks to be less than or equal to ϵrcm to Columba (and similarly, switching Roman and Columba);

CC PROBABILISTIC INTERPERSONAL IGNORANCE: for the same ϵ Roman and Columba are commonly certain that if the mast looks to be $r$cm to Roman, then for Roman p-believes, it looks to be less than or equal to ϵr to Columba (and similarly, switching Roman and Columba);

[23]Note that there are subtleties in the definition of common p-belief, which are not visible in the definitions used in the main text; they also won't matter to the following argument. For discussion, see: Morris (1999), Lismont & Mongin (2003), Heifetz (1999).

[24]Arguments which similarly rely on assumptions of common certainty can be found in Morris & Shin (1997), Morris (1999) and Morris (2014).

CC NO PROBABILISTIC ILLUSION: Roman and Columba are commonly certain that if the mast looks to be $r$cm to Roman, then for all Roman p-believes, it is at most $r$cm (and similarly for Columba).

The conclusion of the argument is that the agents do not have common p-belief that the mast is taller than 100cm.

If proponents of common p-belief accept extensive assumptions of common certainty in arguing *for* common p-belief, it is not open to them to reject the comparatively modest assumptions of common certainty in this argument.[25] Either public information entails common certainty, or it entails only common p-belief for some p which is less than 1. The arguments based on SAILBOAT show that subjects may have public information that φ, but fail to have common certainty that φ. But if common certainty is as rare as this argument suggests, we can't rely on it in arguing for the importance or prevalence of common p-belief. We need an argument for common p-belief, *which is based only on assumptions of common p-belief.* I am not aware of any compelling arguments of this form.

References

Barwise, Jon. 1988. Three views of common knowledge. *Pages 365–379 of: Proceedings of the 2nd Conference on Theoretical Aspects of Reasoning about Knowledge.* Morgan Kaufmann Publishers Inc.

Berker, A Selim. 2008. Luminosity regained. *Philosophers' Imprint*, **8**(2), 1–22.

Breckenridge, Wylie. 2007. *The Meaning of 'look'.* Ph.D. thesis, University of Oxford.

Campbell, John. 2002. *Reference and consciousness.* Oxford University Press.

Chellas, Brian F. 1980. *Modal logic: an introduction.* Vol. 316. Cambridge Univ Press.

Chwe, Michael Suk-Young. 2001. *Rational Ritual: Culture, Coordination, and Common Knowledge.* Princeton University Press.

Clark, H. H., & Marshall, C. R. 1981. Definite reference and mutual knowledge. *Pages 10–63 of:* Joshi, A. K., Webber, B., & Sag, I. (eds), *Elements of discourse understanding.* Cambridge: Cambridge University Press.

Clark, Herbert H. 1996. *Using language.* Cambridge University Press.

Cubitt, Robin P, & Sugden, Robert. 2003. Common knowledge, salience and convention: A reconstruction of David Lewis' Game Theory. *Economics and Philosophy*, **19**(2), 175–210.

Dalkiran, Nuh Aygun, Hoffman, Moshe, Paturi, Ramamohan, Ricketts, Daniel, & Vattani, Andrea. 2012. Common knowledge and state-dependent equilibria. *Pages 84–95 of: Algorithmic Game Theory.* Springer.

[25] A similar point applies to the results of Shin & Williamson (1996), who show that common certainty is required for convention. To reject the upshot of this result the proponent of common p belief will presumably want to deny their assumptions of common certainty. But that is in conflict with endorsing the assumptions of common certainty in the results which motivate the use of common p-belief.

Epley, Nicholas, Morewedge, Carey K, & Keysar, Boaz. 2004. Perspective taking in children and adults: Equivalent egocentrism but differential correction. *Journal of Experimental Social Psychology*, **40**(6), 760–768.

Ernst, Marc O, & Banks, Martin S. 2002. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, **415**(6870), 429–433.

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

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

Gescheider, George A. 1997. *Psychophysics: the fundamentals*. Third edition edn. Malwah, NJ: Lawrence Ehrlbaum Associates.

Gilbert, Margaret. 1989. *On social facts*. Princeton University Press.

Greco, Daniel. 2014a. Could KK be OK? *Journal of Philosophy*, **111**(4), 169–97.

Greco, Daniel. 2014b. Iteration and Fragmentation. *Philosophy and Phenomenological Research*, **88**(1), 656–673.

Halpern, Joseph Y, & Moses, Yoram. 1990. Knowledge and common knowledge in a distributed environment. *Journal of the ACM (JACM)*, **37**(3), 549–587.

Harman, Gilbert. 1974. Review of *Meaning* by S. Schiffer. *Journal of Philosophy*, **71**, 224–9.

Harman, Gilbert. 1977. Review of Jonathan Bennett's Linguistic Behaviour. *Language*, **53**(2), 417–424.

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

Heifetz, Aviad. 1999. Iterative and fixed point common belief. *Journal of Philosophical Logic*, **28**(1), 61–79.

Keysar, Boaz. 2007. *Communication and miscommunication: The role of egocentric processes*.

Keysar, Boaz, Barr, Dale J, Balin, Jennifer A, & Brauner, Jason S. 2000. Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science*, **11**(1), 32–38.

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

Lederman, Harvey. forthcominga. Common Knowledge. *In:* Jankovic, Marija, & Ludwig, Kirk (eds), *Handbook of Collective Intentionality*. Routledge.

Lederman, Harvey. forthcomingb. Two Paradoxes of Common Knowledge: Coordinated Attack and the Electronic Mail Game. *Noûs*.

Lewis, David K. 1969. *Convention: A philosophical study*. Harvard University Press.

Lin, Shuhong, Keysar, Boaz, & Epley, Nicholas. 2010. Reflexively mindblind: Using theory of mind to interpret behavior requires effortful attention. *Journal of Experimental Social Psychology*, **46**(3), 551–556.

Lismont, Luc, & Mongin, Philippe. 2003. Strong completeness theorems for weak logics of common belief. *Journal of Philosophical Logic*, **32**(2), 115–137.

Lu, Chia-Hao, Aman, Joshua, & Konczak, Jüergen. 2009. Integration of visuo-haptic information for judging object size. *Page 62 of: Studies in Perception and Action X: Fifteenth International Conference on Perception and Action*. Psychology Press.

Mertens, Jean-Francois, & Zamir, Shmuel. 1985. Formulation of Bayesian analysis for games with incomplete information. *International Journal of Game Theory*, **14**(1), 1–29.

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

Monderer, Dov, & Samet, Dov. 1989. Approximating Common Knowledge with Common Beliefs. *Games and Economic Behavior*, **1**(2), 170–190.

Morris, Stephen. 1999. Approximate common knowledge revisited. *International Journal of Game Theory*, **28**(3), 385–408.

Morris, Stephen. 2014. Coordination, timing and common knowledge. *Research in Economics*, **68**(4), 306–314.

Morris, Stephen, & Shin, Hyun Song. 1997. Approximate common knowledge and coordination: Recent lessons from game theory. *Journal of Logic, Language and Information*, **6**(2), 171–190.

Moses, Yoram O. 1986. *Knowledge in a distributed environment.* Ph.D. thesis, Stanford University.

Peacocke, Christopher. 2005. Joint attention: Its nature, reflexivity, and relation to common knowledge. *In:* Eilan, Naomi, Hoerl, Christoph, McCormack, Teresa, & Roessler, Johannes (eds), *Joint Attention: Communication and Other Minds.* Oxford: Oxford University Press.

Pinker, Steven. 2007. The evolutionary social psychology of off-record indirect speech acts. *Intercultural pragmatics*, **4**(4), 437–461.

Pinker, Steven, Nowak, Martin A., & Lee, James J. 2008. The logic of indirect speech. *Proceedings of the National Academy of Sciences*, **105**(3), 833–838.

Rand, David G, Peysakhovich, Alexander, Kraft-Todd, Gordon T, Newman, George E, Wurzbacher, Owen, Nowak, Martin A, & Greene, Joshua D. 2014. Social heuristics shape intuitive cooperation. *Nature communications*, **5**.

Rubinstein, Ariel. 1989. The Electronic Mail Game: Strategic Behavior Under "Almost Common Knowledge". *The American Economic Review*, **79**(3), 385–391.

Schelling, Thomas C. 1960. *The strategy of conflict.* Harvard University Press.

Schiffer, Stephen R. 1972. *Meaning.* Oxford: Oxford University Press.

Shin, Hyun Song, & Williamson, Timothy. 1996. How much common belief is necessary for a convention? *Games and Economic Behavior*, **13**(2), 252–268.

Stalnaker, Robert. 2002. Common ground. *Linguistics and Philosophy*, **25**(5), 701–721.

Stalnaker, Robert. 2006. On Logics of Knowledge and Belief. *Philosophical Studies*, **128**, 169–199.

Stalnaker, Robert. 2015. Luminosity and the KK thesis. *Pages 19–40 of:* Goldberg, Sanford C. (ed), *Externalism, Self-Knowledge, and Skepticism.* Cambridge University Press.

Stalnaker, Robert C. 2009. On Hawthorne and Magidor on Assertion, Context, and Epistemic Accessibility. *Mind*, **118**(470), 399–409.

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

Weatherson, Brian. 2004. Luminous Margins. *Australasian Journal of Philosophy*, **82**(3), 373–383.

Williamson, Timothy. 1992. Inexact knowledge. *Mind*, **101**(402), 217–242.

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