

# Defining Common Ground

Seth Yalcin

yalcin@berkeley.edu

December 7, 2023

## 1 Introduction

[Stalnaker \[2014\]](#) develops an array of ideas about common ground. Here I fix on two especially. The first is the idea that the common ground of a conversation is definable in terms of an iterated propositional attitude of *acceptance*, so that  $p$  is common ground iff  $p$  is commonly accepted. The second is the idea that the “default setting” of conversational acceptance is belief, so that as a default, what is accepted in conversation coincides with what is (commonly) believed.

In this paper I defend a pair of contrasting theses. First, I argue that we should identify the common ground with what is common knowledge about what is accepted, so that  $p$  is common ground iff it is common knowledge that  $p$  is accepted. Thus the attitude that is iterated in the definition of common ground is not acceptance but knowledge. Second, I argue that the “default setting” for conversational acceptance is not belief, but knowledge.

What I offer here is meant to be read as an intramural critique of the two Stalnakerian theses listed. I take on board much of Stalnaker’s conceptual background without argument, and aim to preserve the basic structure of Stalnaker’s context set model of common ground. The central claim is comparative: by adjusting Stalnaker’s account in the two basic ways I suggest, we gain some advantages. Thus my ambition is not to defend the resulting story from first principles, or from the myriad rival models of conversation in the literature.<sup>1</sup>

Here is the plan. In §2, I review the basics of Stalnaker’s picture. In §3, I set out the basics of an alternative definition of common ground, one founding

---

<sup>1</sup>Thus, for instance, I don’t engage objections from the alleged impossibility of nontrivial iterated attitudes (as raised for instance by [Fagin et al. \[1999\]](#), [Lederman \[2018a,b\]](#)), or the sort of worries that have motivated the models of the informational context we find in dynamic semantics (e.g., [Heim \[1982\]](#), [Kamp \[1988\]](#); see [Stalnaker \[1998\]](#) for discussion).

common ground on common knowledge. I argue for the alternative account in §4. In §5, I argue for the view that the “default setting” of the common ground is common knowledge. The paper then concludes.

## 2 Common ground as founded on acceptance

In a series of classic papers (Stalnaker [1970, 1974, 1975, 1978]), Stalnaker brought into focus a notion of *common ground* that has exerted considerable influence in both philosophy and linguistics. Stalnaker made a reductive proposal about what a common ground is in Stalnaker [2002], updating it in Stalnaker [2014]; page references are to the latter unless otherwise noted. The general perspective is to “... think of the common ground as a propositional attitude, or as a complex state definable in terms of the propositional attitudes of the participants of the conversation” (37). The concrete suggestion is to build the common ground out of a propositional attitude Stalnaker calls *acceptance*. Stalnaker explains acceptance like this:

...“accept” is the attitude that is iterated in the definition of common ground... The basic propositional attitude that is the basis for the iterative definitions of common ground and speaker presupposition is not given a formal analysis, but is to be understood as something like “acceptance for the purposes of the conversation.” (4)

The notion of acceptance is modeled on the notion of belief, and I will assume that the relevant kind of acceptance has the same formal properties as belief, and so that the formal properties of common ground will be the same as those of common belief. (45)

What are these formal properties? Following an idealization he favors for belief (see, e.g., Stalnaker [2006]) Stalnaker assumes that acceptance satisfies certain “introspection” axioms, those corresponding to the modal axioms 4 and 5: if one accepts something, one accepts that one accepts it; and if one does not accept something, one accepts that one does not accept it:

Positive introspection.  $A\phi \rightarrow AA\phi$

Negative introspection.  $\neg A\phi \rightarrow A\neg A\phi$

Say that if an attitude satisfies both of the corresponding principles, it is *transparent*. Acceptance (like belief) is taken to be transparent. Transparency helps to secure the result that if world  $y$  is compatible with what an agent accepts in

world  $w$ , then what the agent accepts in  $y$  is the same as what she accepts in  $w$ .

One basic motivation for positive introspection is that it would let us extend to common ground some ideas from modal logic used for modeling common knowledge and belief. In the textbook approach to interpreting common belief operators via Kripke-style models (see, e.g., [Fagin et al. \[1995\]](#)), to arrive an accessibility relation capturing what is common belief in a group, we take the transitive closure of the union of the doxastic accessibility relations of the individual members of the group. Such a formalization of common belief would have little interest in the absence of the assumption of positive introspection for the belief states of the individual agents. (And similarly common knowledge.) Stalnaker ultimately wants to formalize common ground along the same pattern, by defining an accessibility relation for the common ground that corresponds to the transitive closure of the union of the acceptance accessibility relations of the members of the group. This formalization too would not have interest without the assumption of positive introspection for acceptance.

It is worth underlining that (in [Stalnaker \[2014\]](#)) ‘acceptance’ is best understood as labeling, not a *category* of attitudes, but rather a *sui generis* propositional attitude of its own—it is assumed to be an attitude *alongside* belief, supposition, etc.<sup>2</sup> But given that acceptance is modeled on belief, it is natural to interpret Stalnaker as holding that belief and acceptance are species of a common genus of attitudes. This would fit with what Stalnaker says elsewhere (e.g., in [Stalnaker \[1984\]](#)).

With this basic grip on the notion of acceptance, we can define *common acceptance* along the familiar pattern of the iterated or hierarchical definition of common knowledge:

$p$  is *commonly accepted* in group  $g$  just in case everyone accepts  $p$ ,  
everyone accepts that everyone accepts  $p$ , etc.

---

<sup>2</sup>This contrasts with the categorical way the word is used in [Stalnaker \[1984\]](#), where acceptance is described as a “generic propositional attitude concept with such notions as presupposing, presuming, postulating, positing, assuming and supposing as well as believing falling under it” (79). There the word picks out a type or category of attitudes. One basic reason this usage does not fit that of [Stalnaker \[2014\]](#) is that, as discussed further below, most of these attitudes are obviously not even ideally transparent. (If one fails to postulate  $p$ , it doesn’t follow that one postulates that one isn’t postulating  $p$ ; etc.) [Stalnaker \[2014\]](#) also regularly describes acceptance as “acceptance for some purpose or other” (39). This would make little sense on the categorical understanding of ‘accepts’, where belief would be understood as a variety of acceptance; for we don’t believe things for purposes.

One could use ‘accepts’ both ways, so that acceptance (the *sui generis* attitude) is a variety of acceptance (the putative category of attitudes identified in [Stalnaker \[1984\]](#)); but I will avoid punning on the term, and will set the categorical use aside.

—and now we can state Stalnaker’s thesis that the common ground of a conversation is to be identified with what is commonly accepted among the interlocutors:

CG=CA.  $p$  is common ground in  $g$  iff  $p$  is commonly accepted in  $g$ .

We can call this a “definition” of common ground if we want, but I will read this as a substantive thesis about the metaphysics of common ground. It is a claim about how the common ground is grounded in the states of the conversing agents—about what the common ground is in more fundamental terms.

We could put it also in terms of role and realizer. Stalnaker puts the notion of common ground to work in modeling the most basic dimensions of communication and assertion, and also puts it to work in clarifying a range of specific linguistic phenomena—for instance, presupposition triggers, conditionals, and attitude verbs. (And of course, the literature is rich with myriad extensions of the framework, too numerous to cite.) These applications frame an explanatory role for common ground—a job description for the notion. CG=CA is a proposal about what realizes that role, about what performs this work. So one can agree with Stalnaker (as we will here) that something does do all this work, while disagreeing him about whether it is common acceptance that does it.

By design, on Stalnaker’s account, common ground has an iterative structure exactly like that of common belief and common knowledge. Stalnaker traces the social and public character of common ground to this iterative structure:

... common ground, and speaker presupposition, are social and public attitudes, but I want to explain the social, public character of the attitude in terms of the iterative structure, as in common knowledge and belief. (38)

Defining a common ground operator in Stalnaker’s setting in the standard way,<sup>3</sup> the operator satisfies positive introspection: if it is common ground that  $p$ , it is common ground that it is common ground that  $p$ .

Why not sidestep talk of acceptance, and just identify the common ground with what is common belief in the group? A central motivation is that it “may serve the purposes of the conversation to engage in some mutually recognized pretense, or to carry on a conversation within the scope of some mutually recognized supposition” (46), and in such cases one is conversationally accepting what one doesn’t believe. Still, these sort of cases do involve a certain departure

---

<sup>3</sup>I.e., as a normal necessity modal whose accessibility relation corresponds to the transitive closure of the union of the acceptance accessibility relations of the members of the group.

from the norm on Stalnaker’s view. He proposes that normally, as a default, the things that we (commonly) accept are indeed things that we (commonly) believe:

In simple straightforward serious conversations, what is accepted will coincide with what is believed... (4)

I will assume that belief (and common belief) are the default settings for acceptance for the purpose of the conversation (and for common ground). (45)

To be clear, the proposal is not that states of acceptance are sometimes literally *constituted* by states of belief. Rather, it is that in normal conditions, the propositions that the agents (commonly) accept for the purposes of conversation align with the propositions they (commonly) believe. I will call this the idea that there is a *belief default*:

BD. As a default,  $p$  is common ground in  $g$  iff  $p$  is commonly believed in  $g$ .

I will take the idea to be that this is a default norm governing cooperative communicative exchanges. (It could also be described as a default concerning the “conversational tone” in the sense of Yalcin [2007].) This default may be overridden with some suitable signal that the norm is being suspended (e.g., a preface like “Suppose hypothetically the following...” or “Let me tell you a story...”, etc.). It also may be overridden by other indications that the context isn’t normal, such as indications that one or more speakers are not being cooperative in roughly the sense of Grice [1975]. Meanwhile BD captures a perfectly natural idea about what *doesn’t* take any special setup in conversation. Normally, if you walk up to a person and assert a proposition with no preliminary qualifications or hedges, you’ll be understood to be saying what you believe, and presumed to be attempting to transmit that belief. BD captures this simple point.<sup>4</sup>

CG=CA and BD are the main two Stalnakerian theses I want to scrutinize. But we should mention some further important notions that Stalnaker defines in terms of common ground. First we have (speaker) presupposition:

---

<sup>4</sup>Exegetically speaking, I claim only that one could reasonably pull BD of the quotes recently mentioned. The main issue will be whether BD is true, not whether it is Stalnaker’s view. But on the exegetical point, it is worth noting that Stalnaker makes a number of remarks that would appear to cohere better with a knowledge default rather than a belief default. I will discuss this further in §5.

An agent in  $g$  *presupposes*  $p$  iff she accepts that  $p$  is common ground in  $g$ .

Stalnaker holds that there is a *norm of agreement* governing presupposition. The norm enjoins interlocutors to accept that  $p$  is common ground only when  $p$  is in fact common ground; equivalently, it enjoins everyone to presuppose just the same things.<sup>5</sup> If the norm is violated, it means one or more speakers takes something to be common ground that is not. In that case, the context is said to be *defective*.

Is the attitude of presupposition transparent? Given how presupposition is defined, this is equivalent to the question whether the complex attitude expressed by “accepts that it is common ground that” ( $A^{CG}$ ) satisfies these principles:

$$\begin{aligned} A^{CG}\phi &\rightarrow A^{CG}A^{CG}\phi \\ \neg A^{CG}\phi &\rightarrow A^{CG}\neg A^{CG}\phi \end{aligned}$$

Presupposition is indeed positively introspective, but it is not in general negatively introspective. Consider a case where you think your interlocutor mistakenly takes you to be presupposing  $p$ . You are not in fact presupposing  $p$  ( $\neg A^{CG}\phi$ ), but you think this fact is not common ground, so you do not accept that it’s common ground that you are not presupposing  $p$  ( $\neg A^{CG}\neg A^{CG}\phi$ ). Thus failing to presuppose that  $p$  does not imply that one presupposes one isn’t presupposing  $p$ . This is a case where you take the context to be defective. But if we restrict attention to cases where the agent presupposes that their context is non-defective, conceiving of that as the normal case, Stalnaker claims that the negative introspection principle does hold (232). That is: restricting to nondefective contexts, presupposition is transparent.

Given an individual agent in conversation, we can speak of the set of possible worlds compatible with what they presuppose. This is that agent’s *context set*. If an agent presumes their context set is nondefective, then if world  $x$  is compatible with what an agent presupposes in world  $w$ , then what the agent presupposes in  $x$  is the same as what she presupposes in  $w$ . In other words, an agent’s context set does not vary across the worlds it leaves open. If a context actually is nondefective, then the context sets of all of the conversing agents are the same, and we can speak of *the* context set of the conversation.

---

<sup>5</sup>Are there things that agents accept for the purposes of conversation, yet don’t presuppose? Space for this possibility exists given the definitions Stalnaker provides, apparently. (If we wanted to eliminate this possibility, we might restate the relevant norm as one directly governing acceptance: accept  $p$  only when  $p$  is in fact common ground.)

Stalnaker’s famous model of assertion as a proposal to add the content asserted to the common ground (Stalnaker [1978]) relies on the idea that in the nondefective ideal case, it is mutually apparent to the interlocutors what the common ground is. (And thus the model relies on the transparency of acceptance, and on the idea that presupposition is normally transparent). For an utterance  $u$  will only be mutually understood to be the assertion of a proposition  $p$  if it is common ground that  $u$  expresses  $p$ ; and an assertion will only result in its intended effect if the interlocutors arrive at a new common ground that is a function the previous one together with the content asserted. If my context set worlds were to vary in respect of what was common ground at each of them, then the mutually understood content of my assertion might vary across my context set worlds (since that content is generally sensitive to what is common ground), and in a way I could recognize in advance of my assertion. But then the rational reconstruction of my speech act as a proposal to add *the* content of my assertion to *the* common ground would break down. Stalnaker’s transparency assumptions thus play an important role in legitimating the context set model.<sup>6</sup>

### 3 Common ground as founded on knowledge

Contrast CG=CA with:

CG=CKA.  $p$  is common ground in  $g$  iff it is common knowledge in  $g$  that  $p$  is accepted by everyone in  $g$ .

In the next section, I argue that we should prefer CG=CKA to CG=CA. In this section, I clarify what CG=CKA says, and its package of background assumptions.

CG=CKA defines common ground in terms of two attitudes: acceptance and common knowledge. We assume the standard iterative definition of common knowledge. We assume that knowledge is positively introspective (i.e., the  $KK$  principle), but not negatively introspective. Turning to ‘accepts’, here I mean something close enough to what I take Stalnaker to be targeting that I have opted to use the same word. I take acceptance to be a *voluntary* attitude, roughly of a kind with attitudes like taking for granted, entertaining, supposing, imagining, and postulating (call these *entertainment* attitudes). Like Stalnaker, I don’t try to explain acceptance in more fundamental terms, but take it as basic. But there is one very important difference in the way we will understand

---

<sup>6</sup>See Hawthorne and Magidor [2009] and Stalnaker [2009] for additional relevant discussion of the connection between transparency principles and Stalnaker’s model of assertion.

acceptance in the context of CG=CKA: here we will not assume that acceptance is transparent. What CG=CKA assumes instead is that acceptance is *epistemically* transparent. By this we will mean that acceptance satisfies both of these principles:

$$\begin{aligned} A\phi &\rightarrow KA\phi \\ \neg A\phi &\rightarrow K\neg A\phi \end{aligned}$$

That is: when you accept  $p$ , you know that you accept it; and when you don't accept  $p$ , you know that you don't.

In the context of CG=CKA, 'presupposition' is just another label for 'acceptance': an agent *presupposes*  $p$  iff she accepts  $p$ . Under CG=CKA's assumptions that knowledge is positively introspective and that acceptance is epistemically transparent, a limited form of negative introspection will hold for knowledge:  $\neg K\phi \rightarrow K\neg K\phi$  will hold whenever  $\phi$  is of the form  $A\psi$  or  $\neg A\psi$ . This legitimates the representation of an agent's presuppositions as a set of possibilities: for the set doesn't vary across the agent's epistemic alternatives. An agent's *context set* is the set of worlds compatible with what they accept.

We assume that there is a *norm of agreement* enjoining interlocutors to accept alike—to be in acceptance states representable with the same context set. A context of interlocutors  $g$  is *nondefective* just in case it is common knowledge in  $g$  that for any  $p$  and all  $a$  in  $g$ ,  $a$  accepts  $p$  iff  $p$  is common ground in  $g$ ; otherwise it is *defective*.

It is worth distinguishing weak and strong varieties of defectiveness. Say a context is *strongly defective* iff it is not the case that the interlocutors accept exactly the same things. In a strongly defective context, not all of the context sets of the agents match. These will be cases where the norm of agreement is violated.

But a weaker sort of defect in the context is also possible. Suppose that everyone is accepting just the same things, but that this fact—that everyone is indeed accepting alike—is not itself common knowledge. For instance, maybe neither of us is presupposing  $p$ , but I incorrectly think you might be. Then there is a  $p$  such that I do not know that: you presuppose  $p$  iff  $p$  is common ground between us. Call this sort of context *weakly defective*. This is a defective context insofar as your context set is not invariant across my epistemic alternatives. The norm of agreement is not violated, but it isn't common knowledge that it isn't violated. A fully nondefective context is one where our context sets are not only the same, but are commonly known to be the same: they are the same across our respective epistemic alternatives. In the ideal context we want not



just *de facto* matching of acceptance states, and not just common knowledge of what is being accepted for the purposes of the conversation, but also common knowledge that the same things are being accepted.

As I hope is plain, CG=CKA doesn't simply identify the common ground with common knowledge. Falsehoods may be common ground on this view, because acceptance is not a factive (veridical) attitude. When  $p$  is common ground, what follows is (not that  $p$  is true, but) that  $p$  is accepted, and that it is commonly known to be accepted. If a false  $p$  is common ground and wrongly thought by the interlocutors to be known, then it might be common ground that they know that  $p$ , and it may even be common ground that they commonly know  $p$ ; but neither of these facts (about what is common ground) implies that  $p$  is true.

The notion of common ground provided by CG=CKA is not itself iterable.<sup>7</sup> That is, given CG=CKA, if  $p$  is common ground, it does not necessarily follow that it is common ground that  $p$  is common ground. However, if  $p$  is common ground, it does follow that it is common knowledge that  $p$  is common ground. (This owes to the iterability of the common knowledge operator.)

This suffices to get the basic structure of CG=CKA on the table. The best way to bring its content further into focus is to argue for the view. In the next section, I will bring out three interconnected ways that CG=CKA seems to improve upon CG=CA.

## 4 Acceptance versus knowledge as a foundation

### 4.1 Transparency troubles

CG=CA hypothesizes an acceptance attitude that is transparent. In this section I argue that the hypothesis is problematic. I further argue that even if granted, the hypothesis would not serve to clarify the public character of the common ground in the way that it is intended to. Meanwhile CG=CKA does better on both counts: it avoids the assumption that acceptance is transparent, and provides the better path for explaining the common ground's public character.

What is the problem with the idea that acceptance is transparent? The problem is that it is in tension with the requirement that acceptance is also voluntary—that it is under the control of the agent. Acceptance must be voluntary for the obvious reason that one can freely opt in or not to accepting something for the purposes of conversation. In this important way, acceptance differs from belief, and classifies more naturally with what we above labelled *en-*

---

<sup>7</sup>Thanks here to Harvey Lederman.

*tertainment* attitudes: entertaining, imagining, taking for granted, and the like. But on the face of it, voluntariness and transparency pull in conflicting directions. Ordinary entertainment attitudes are not transparent (not even ideally).<sup>8</sup> On the other hand, the attitudes that have the best shot at satisfying one or more of the introspection axioms—belief and knowledge—are clearly not voluntary. Stalnaker’s hypothesized notion of acceptance is thus more unusual than it may at first appear, combining features that are, on reflection, characteristic of very different classes of attitudes.

To be clear, I don’t claim that voluntariness and transparency are logically incompatible.<sup>9</sup> The main point is just that the content of normal entertainment attitudes are under executive control at least to the extent that we can simply deliberately entertain combinations of contents that would violate transparency: one can, for instance, imagine that one is asleep, or that one doesn’t exist. So it is not enough to explain acceptance by pointing to the familiar class of entertainment attitudes and saying: “It’s like those attitudes—except, it’s transparent”. More needs saying, for it is not obvious that there are any natural examples of voluntary, transparent mental states. One could stipulate a term for a subclass of entertainment states that *de facto* satisfy transparency, but the stipulation wouldn’t bring a new joint in psychological nature into existence, or render the proposed notion of acceptance less artificial.

That is the first worry about assuming acceptance is transparent. I mean it only as a warm-up to a second and deeper worry, which is this. When we compare acceptance to ordinary entertainment attitudes, a problem about the putative transparency of acceptance comes into focus. The entertainment attitudes are not transparent. But even if they were, it is important to see that they would *not thereby be* in any sense *more epistemically accessible* states—more transparent in an epistemic sense.<sup>10</sup> The attitude of imagination, for

---

<sup>8</sup>That is: one is just mistaken about the basic nature of (e.g.) imagination if one thinks that to imagine that there are unicorns, one must additionally imagine *that one is imagining* there are unicorns. Likewise, if you *fail* to suppose that *p*, or fail to entertain the truth of *p*, it does not follow that you must be supposing that you are not supposing *p*, or entertaining the proposition that you are not entertaining *p*. And so on, for all other entertainment attitudes.

<sup>9</sup>As a reviewer notes, one place we may see voluntariness and iterability come together is in *trying*. Trying seems voluntary; and one might argue, following [Holguín and Lederman \[2023\]](#), that the logic of trying is iterative: if one tries to do something, one tries to try to do it. Note however that in the main text we are talking about transparent *attitudes*—and trying is not a kind of mental state, but a kind of action. (One might think any action of trying implies a corresponding mental state of intending; but Holguín and Lederman argue against this.) Also, even if *trying* implies *trying to try*, presumably *not trying* does not imply *trying not to try*—that is, trying doesn’t satisfy the analogue of negative introspection.

<sup>10</sup>This brings out the point that calling instances of the modal axioms 4 and 5 “introspection” or “transparency” principles makes questionable sense except in connection with belief or knowledge.

instance, is not iterative, but even if it had that logical structure, that would be orthogonal to the question of epistemic access—to the question of whether or under what conditions I know what I am imagining or not imagining. Similarly, without additional assumptions, the claim that acceptance is transparent is orthogonal to the question of whether or under what conditions we know what we are or are not accepting.

The public character of the common ground, however, seems tied up exactly with that question. A proposition we each accept is not public between us if we do not even know or believe that we are accepting it. (This is why the standard view is that publicity is grounded in common knowledge or common belief.<sup>11</sup>) But this basic fact about the public character of the common ground wouldn't follow from the thesis that acceptance is transparent, as that thesis says nothing at all about belief or knowledge. Once we see clearly that the putative transparency of acceptance does not imply epistemic transparency, it is hard to see what the transparency of acceptance *per se* could have to do with the public character of the common ground.

The proponent of  $CG=CA$  could at this stage wheel in BD, and note that on their view, common acceptance normally coincides with common belief; and they might add the claim that it is really the latter that ultimately grounds the publicity of the common ground. But this would be to concede the point I am pressing, which is that publicity is not explained until belief or knowledge enter the story, and therefore it is not the iterative structure of acceptance specifically that explains publicity.

This position would also make for a needlessly indirect account of the public character of common ground, given that we can embrace  $CG=CKA$  instead. On that view, no one attitude is alleged to be both voluntary and iterative. Acceptance is voluntary, but not iterative. Knowledge is iterative, but not voluntary. The publicity of the common ground owes to the simple fact that it is common knowledge in the group what their common ground is.  $CG=CKA$  fits the standard view that common knowledge is what makes for publicity, while avoiding the postulation of an attitude with an apparently unprecedented combination of properties. If there is a transparency thesis in the vicinity of acceptance that is plausible, it is epistemic transparency.

---

<sup>11</sup>In a survey article on common knowledge, [Lederman \[2017\]](#) characterizes this as the “default position” in that literature.

## 4.2 The necessary ties to belief and knowledge

That common acceptance couldn't suffice to explain publicity connects with a broader problem for  $CG=CA$ : it carries a burden about explaining what look to be obvious necessary connections between the common ground and what the agents in the conversation believe or know. This burden is not one  $CG=CKA$  carries.

Here is a platitude: communication doesn't happen between agents who don't even believe that they are communicating, no matter how many layers of acceptance they undertake. Where we have need to speak of a common ground between agents, we are dealing with agents that have some interesting amount of mutual belief—belief concerning, at least, the facts about their communication. Maybe the conversation in question is one where many of the propositions being accepted are not themselves believed-true (they include temporary suppositions, etc.); maybe each interlocutor is deeply skeptical of what the other has to say. Still, in any nondefective case, the facts concerning *what propositions are being accepted in the conversation* are themselves going to be a matter of mutual belief, if not knowledge—if not common knowledge.

Given  $CG=CKA$ , none of that is surprising. On that view, mutual belief about what is being taken for granted is just constitutive of common ground to the extent that it is constitutive of common knowledge, so this connection between common ground and belief secured quite directly. Given  $CG=CA$ , however, the tie between common ground and belief must be indirect, for acceptance does not imply belief, and neither does common acceptance imply common belief.<sup>12</sup> Does common acceptance nevertheless somehow imply common belief about what is accepted? I think this is what the defender of  $CG=CA$  should want to hold—after all, if it were metaphysically possible for two agents to be in a state of common acceptance while failing to have any common belief about what they accept, or even about whether they were in conversation at all, those would seem to be cases where  $CG=CA$  predicts a common ground where there just isn't one.

But it's not obvious how  $CG=CA$  hopes to forge a suitable connection between common ground and common belief or knowledge. One might try to say that common acceptance can only be metaphysically brought about by way of common belief, though it is not itself constituted by common belief. This way of developing  $CG=CA$  would imply that wherever there is common ground, there are really *two* infinitely iterated attitudes involved: there is what is commonly

---

<sup>12</sup>We have noted Stalnaker's thesis that there is a belief default (BD), but this is just a defeasible constraint governing what content is normally treated as common ground—it does not describe a necessary connection between acceptance and belief.

accepted (the common ground), and then there is common belief about what is commonly accepted (which causally sustains the common ground, though it doesn't partly constitute the common ground). That would reveal  $CG=CA$  to be a more complex story than it seemed to be—and more complex than  $CG=CKA$ , which doesn't invoke more than one iterated attitude in the full story about what it takes to bring a common ground into existence.

Considering questions like these—Is common acceptance without common belief or knowledge metaphysically possible? Is common ground caused by, but not constituted by, common belief? Etc.—it is natural to have the feeling that we have somehow gone off the rails. I think that feeling is exactly right. One may assume  $CG=CA$  while ignoring these questions, but they are questions that the proposal raises, and the fact that it raises these questions is a concern. One sign that a theory is carving at the wrong joints is that it raises spurious questions.  $CG=CKA$  does not raise questions like these.

To summarize: if we do not found the common ground on common belief or common knowledge in the first place, we are saddled with the burden of restoring what are by all appearances necessary connections between these notions.  $CG=CKA$  founds the common ground on common knowledge, so this is a burden it does not carry.

### 4.3 The need for an epistemic foundation

If it is conceded that the above considerations motivate an approach that would found common ground directly on common belief or common knowledge, the question arises which of those two to go with. In fact, Stalnaker's original reductive thesis (Stalnaker [2002]) invoked common belief. The proposal was:

$CG=CBA$ .  $p$  is common ground in  $g$  iff it is common belief in  $g$  that  $p$  is accepted by everyone in  $g$ .

$CG=CBA$  fares better than  $CG=CA$  with respect to the challenges so far raised. There are reasons, however, to favor  $CG=CKA$  instead. And these reasons to favor  $CG=CKA$  over  $CG=CBA$ , we'll see, are further reasons for favoring it over  $CG=CA$ .

The reasons connect with a much broader debate, about how common belief and common knowledge compare in their power to support robust explanations of coordinated behavior in general. Yalcin [2023] argues that the comparison is not close. There I argue common belief in the absence of relevant common knowledge is a much rarer phenomenon than is often assumed, and that uncontroversial examples of it do not in fact enable something recognizable as

coordination at all. I offer a general case that common knowledge is the central notion in the explanation of coordinated action, that common belief makes for robust explanations of coordination only insofar as it approaches common knowledge. The case is relevant here, since I assume, with Stalnaker, that common ground essentially emerges out of coordinated action: there is no common ground without successful coordination on states of acceptance. While I cannot review the full argument of Yalcin [2023], let me draw out the key point as it bears on common ground.

Given that common knowledge implies common belief,  $CG=CKA$  is the strictly more restrictive view: anything that will count as a common ground according to  $CG=CKA$  will count as a common ground according to  $CG=CBA$ , but not the reverse. Thus  $CG=CBA$  will recognize the existence of a common ground where  $CG=CKA$  does not. To adjudicate between the views, therefore, we need to bring into focus the putative sort of common ground that  $CG=CBA$  but not  $CG=CKA$  would recognize, and use it to assess which proposal draws the more natural boundary.

So we can ask: what would it be for agents to have common belief about what they accepting that is as close as possible to, without being, common knowledge? Remember, acceptance is not factive. The mere fact that a common ground can contain false propositions is no problem for  $CG=CKA$ . To find cases where agents have common belief about what they accepting that is not common knowledge, the closest it appears we can come is to consider what are effectively “thoroughly Gettiered” cases of common belief. Here is an example from Yalcin [2023]:

In a noisy and darkly lit bar, Alice sees from behind a person she mistakes for Carl. (He’s sitting in Carl’s usual seat, looks just like him from behind, and so on.) She leans in whispers the following into the stranger’s ear: “It’s me, Alice. I’ve gotta run now, but let’s meet tomorrow at the cafe at noon.” The stranger is deaf, so he doesn’t respond, but he gives a thumbs up without looking at Alice. Alice interprets this as agreement about the plan, but in fact the stranger is just signaling to the bartender that he wants another shot. Meanwhile, Carl is at another bar and coincidentally undergoes a completely analogous confusion: he mistakes a stranger for Alice, says “Let’s meet tomorrow at the cafe at noon”, and experiences the illusion of assent from the stranger.

Here Alice and Carl have common belief, but not common knowledge, that they are talking to each other. We might even say that they have common belief that they both accept that they are talking. Is the proposition that they are talking *common ground* between them? There seems no reason to say so. Alice

and Carl are not even connected, and not actually communicating. If this is the kind of case that  $CG=CBA$  but not  $CG=CKA$  permits, then what it helps to reveal is that  $CG=CKA$  carves the more natural joint.

Where we have need to speak of common ground, we have agents in some state of coordination. I will not offer a conceptual analysis of ‘coordination’, but it is obvious that the relevant notion is a modally robust. There is no “lucky” or “accidental” coordination; coordinated action involves an interdependence in action. (As Yalcin [2023] argues, if Alice and Carl meet at the cafe at the next day, that is not because they were actually coordinated; it is because of a bizarre coincidence.) Belief is compatible with lucky correctness, and common belief is compatible with bizarrely coincidental origins, as in the story above. But one cannot luck into knowledge, or into common knowledge. Knowledge requires some kind of appropriately robust connection to the truth known,<sup>13</sup> and thus common knowledge requires the agents to bear some kind of appropriately robust connection to each other. This modally robust connection is what is needed to ground coordinated action, and more specifically, the sort of coordination we are interested in when we speak of common ground. Common belief, on the other hand, seems of little value for grounding coordination where it does not constitute common knowledge.

It can be tempting to think that basically all rational action, whether individual or collective, is fully explicable in terms of belief and desire. From that starting point, it can feel like something must have gone wrong in the point I am making here. But as Greco [2016] notes, very often what we are interested in explaining about an action is why it *succeeded*—why it resulted the securing of the preferred outcome, whatever that was. And as Greco brings out, knowledge and mere true belief differ in their power to undergird a robust explanation of success in action. Success that is explained by way of true belief that isn’t knowledge is generally success by luck—which is to say, it is success that doesn’t really admit of much explanation at all. We can agree that Alice and Carl’s individual belief and desire states make predictions about what they’ll do—both concerning what they will accept, and how they will act. But it is important to be clear what can and cannot be robustly explained by appeal to these states. Their individual belief states do help to predict that they will each show up at the cafe at noon the next day. But if we are interested in explaining how they got what they wanted—how they managed to show up at the same time and

---

<sup>13</sup>Of course, the target sense of robustness has been cashed out by philosophers in different ways, for instance via the idea that knowledge is *sensitive* (Nozick [1981]) or *safe* (Sosa [1999], Williamson [2000] Pritchard [2005]). See especially Greco [2016], who considers the matter from the perspective of the role of robustness plays in explanation in general.

place, in the way they both wanted—we have to grant that there’s not much to explain there: this was fundamentally luck. There was no actual coordination, no stable sensitivity of the actions of one to that of the other. The situation would be different if they had actually talked: in that case, their success in meeting becomes explicable. For their states of mind would have a robust sensitivity to each other, in a way that makes the outcome finally observed nonaccidental. This robust sensitivity is what common knowledge makes for.

What I claim is that this at least suffices to place a burden on CG=CBA and on CG=CA. Common ground requires non-accidental coordination in states of acceptance, and there is a reasonable *prima facie* case that a foundation in common knowledge is required where there is non-accidental coordinated action. CG=CBA and CG=CA deny that common ground must be founded on common knowledge, and imply that it can exist without it. The burden on these views is to provide the examples that show that the demand of knowledge is too strong—that there exists a common ground where CG=CKA would recognize none. Looking for these examples ourselves, what we seem to find are bizarrely coincidental cases, like those of Alice and Carl. But CG=CKA’s verdict that there is no common ground in these cases is the right one.

Let me consider one reply to the above argument on behalf of CG=CBA and on CG=CA, a reply that has been suggested to me. The reply is to double-down on the idea that the notion of acceptance these theories assume is acceptance *for the purposes of the conversation*. One might have initially assumed that describing acceptance in this way was just meant to call attention to the usual sort of *reason* speakers have for accepting as they do; but the idea now is think of acceptance attitudes as essentially *relativized to* conversations. One does not accept *p simpliciter*; rather, one accepts *p* only relative to conversation *c*. How is this supposed to help? In the case of Alice and Carl, the idea would be to stress a point already made, viz., that they are not in conversation in the first place. So although they do in some sense each individually accept that they are talking to each other, they don’t accept this for the purposes of “their conversation”. So these theories don’t after all predict that it is common ground between Alice and Carl that they are talking; so the case isn’t, after all, a problem for those theories.

A striking feature of this reply is that it brings in precisely that feature of real conversations that grounds the possibility of common knowledge about what is accepted, viz., a genuine causal-informational connection between the agents. (At least assuming that whatever a “conversation” is, it requires at least that.) But it far from clear that we can make sense of two agents’ being in



conversation—their having the suitably robust causal-informational connection that anything worth calling “conversation” requires—in the absence of any common knowledge at all between the agents. This way of developing  $CG=CBA$  and  $CG=CA$  thus brings them very close to  $CG=CKA$ —so close that it now seems hard to tease them apart. What sort of example is supposed to show that  $CG=CKA$  makes predictions that are too strong? What is a case where agents are in fact in conversation, and yet somehow only have common belief or common acceptance about what they accept, while lacking any relevant common knowledge? We can all agree that actually being in conversation requires actual connection, a connection that allows information to flow between the agents. But information flow of this sort between agents implies that they are in contentful states of mind that exhibit a modally robust mutual sensitivity to each other, one sustained by their causal connection. But that is precisely the kind of thing common knowledge is constituted by.

If a conversation-relativized interpretation of  $CG=CBA/CG=CA$  does just turn it into  $CG=CKA$  in another guise, there is nothing to argue about, except perhaps which guise is most perspicuous. On that, a notable difference is that the version of  $CG=CBA/CG=CA$  currently under discussion invokes a new primitive notion: that of a group’s being in conversation. (Or it makes explicit a primitive that was there all along.) This notion is now used to explain what common ground is. But one might have thought that what we were up to, in introducing the idea of common ground, was the reverse: we were hoping, *inter alia*, to explain what a conversation is in terms of common ground.

Let me close this section by responding to one last inchoate (and very non-Stalnakerian) reason for hesitating about appealing to common knowledge in explaining common ground: the skeptical impulse. Philosophers sometimes have a tendency to exalt the state of knowledge, so that laying claim to it can take on the appearance of extravagance. Such a conception of knowledge hardly has presumptive status. Another view, the view assumed here, is that knowledge, and common knowledge, are perfectly ordinary and commonplace. If anything, it is common belief *without* any relevant common knowledge that is strange and unusual—as the case of Alice and Carl brings out.

## 5 On the default of common ground

Let me turn now to the question of what the “default” of the common ground is. Again, the view I read into Stalnaker is:

BD. As a default,  $p$  is common ground in  $g$  iff  $p$  is commonly believed in  $g$ .

In contrast, in the section, I defend:

KD. As a default,  $p$  is common ground in  $g$  iff  $p$  is common knowledge in  $g$ .

In certain ways I will explain, the debate that could be had between these views connects with issues that go beyond the boundaries of this paper. Accordingly my defense of KD will have a somewhat tentative character. But not all of the intended value of this section of the paper resides in its defense of KD. The basic question here—of whether there is a “default setting” of the common ground, and if so, what it is—seems hardly to have been recognized or discussed, so part of what I aim to do is to simply bring the issue into focus.

Indeed, though I read BD into Stalnaker on the basis of the earlier cited quotes, there is reason to doubt his commitment to it runs deep: he does not single the idea out in the way I have, and in various places he can seem to have KD in mind instead: for he describes the common ground in passing in many places as *presumed* common knowledge (see [Stalnaker \[2014\]](#) 25; 29; 36; 45). But in many contexts he also seems to follow the common practice in the literature of using “common knowledge” indiscriminately to mean common knowledge or common belief—thus making it hard to gauge whether he means to emphasize knowledge specifically in those places. Suffice to say that the question of BD versus KD is not clearly on the radar of [Stalnaker \[2014\]](#). Setting exegesis aside, my main aim is to bring out the question, and to highlight the *prima facie* advantages KD has.

## 5.1 The norm of assertion

There is a considerable literature on the question what the norm of assertion is, much of it originating in epistemology (in classic work by Unger, DeRose, and Williamson) and concerned with the connection between assertion and knowledge.<sup>14</sup> It has often been unclear how if at all this debate about the norm of assertion connects with the modeling proposal about assertion advanced in [Stalnaker \[1978\]](#). But once we have in focus the question of what the default setting of common ground is, we can see a rather direct point of connection. BD and KD are default norms governing the common ground. Given that assertion is, as

---

<sup>14</sup>See [Unger \[1975\]](#), [DeRose \[1991, 2002\]](#), [Williamson \[1996, 2000\]](#).

Stalnaker [1978] suggested, a kind of proposal to add the proposition asserted to the common ground, we expect that BD and KD will generate corresponding default norms for assertion. That is: if speakers as a default view the common ground as tracking common belief (knowledge), then unless the default is overridden, we expect them to assert only what they believe (know), to the extent they are cooperative.

There may be norms governing assertion that are not parasitic in this way on the norms governing common ground. But it is worth framing the minimal position, which is that there aren’t any such norms (cf. Pagin [2016]): that *any norms governing assertion are derivative from the norms regulating common ground*. There’s a debate to be had—not here—about whether this attractively simple hypothesis is correct. I will assume it moving forward; let us see where it takes us.

Given this background, KD but not BD suggests that *as a default*—when the default norm governing common ground is in force—one should not assert  $p$  unless one knows it. That is a weaker version of the knowledge norm of assertion famously defended by Williamson [1996, 2000], among many others.<sup>15</sup> But it is worth stressing some important differences. Williamson understands the rule that one must assert  $p$  only if one knows  $p$  as “constitutive” and “essential” to assertion. This view of assertion would not follow from KD, and we do not assume it. What is essential to assertion, we assume, is only its characteristic dynamic effect on the common ground—roughly along the lines advanced in Stalnaker [1978]. KD is a norm governing common ground—and merely a *default* norm at that. If this default is in force and KD is correct, that makes the prediction that one should assert only what one knows. But if the default is overridden, assertion is still, for all we have said and will assume, perfectly possible. And it may be that the default is overridden, or temporarily suspended, in systematic ways in ordinary discourse. In this way there is no essential connection assumed here between assertion and knowledge (or, for that matter, between assertion and belief). More specifically, we don’t assume that it is part of the identity of assertion that it is regulated by norms at all: we don’t assume that norms need to be mentioned in order to explain what assertion is.

The weakened version of the knowledge norm predicted by KD is motivated by all the facts typically wheeled in to support its stronger counterpart. It immediately predicts the default infelicity of Moore-paradoxical sentences like:

- (1) ??It’s raining and I don’t know it.

---

<sup>15</sup>See Turri [2016] and citations therein.

BD, on the other hand, fails to predict this without supplementation. Suffice to say that there is a range of familiar data here. For instance, if one asserts that it’s raining (or whatever), the question:

(2) How do you know?

is generally fine, at least where the answer isn’t obvious. But the question presupposes that the speaker purported to know in asserting, and that fits naturally with KD, not BD.

The weakened norm is also arguably as strong as we should want. For it seems obvious one can override the default setting of the common ground with some suitable preface or signal, and thereby render it permissible to assert things that one doesn’t know. For instance, if the context is right and I say, “Once upon a time...”, I may launch into a barrage of assertions of things I don’t know. One could describe this, as Frege [1918] did, as “sham assertion”, but to say that seems to miss the extent of overlap between the case where the default is in force and the case where it isn’t. Whether I am reporting history (i.e., with KD in force) or telling a fairy tale (with KD overridden), I will use declarative sentences to incrementally update some kind of shared body of information. The key difference between the two cases resides in our shared attitude about what this body of information corresponds with, a fact reflected in the setting of common ground. When I say “It was a stormy night” in either case, the sentence will presumably update the relevant shared body of information in exactly the same characteristic way—and to this extent we seem to have something with title to be called “assertion” in both cases.<sup>16</sup>

Meanwhile BD unsupplemented predicts the assertibility of unassertable propositions. Here is an example. I don’t know who won the game last night, but I am confident that the Warriors beat the Knicks, since the Warriors have a much

---

<sup>16</sup>That said, the view I have described is compatible with alternative accounts of such storytelling (or mixed accounts). Another possible view is that when I say “It was a stormy night” in the context of relating a fiction, my sentence isn’t understood as proposal to update the “matrix” context set, or common ground, of the conversation, but is instead taken as a proposal to update what Stalnaker [1988, 2014] calls a *derived* or *subordinate* context, a body of information determined somehow as a function of the common ground. On this view, my utterance does eliminate possibilities from the matrix context set, but it only has this result mediate, by way of its impact on a corresponding derived context. The details will depend on how exactly the target derived context is fixed as a function of the matrix context set. If one says this, then the utterance in question is *not* a proposal to add the content asserted to the *matrix* context set—and so isn’t an assertion, at least in one clear sense. On such a view, the storyteller might be understood as (indirectly) communicating what she knows about what happens in the fiction—and to that extent, such a view might say that storytelling doesn’t always require an override of KD. The questions of how derived contexts are triggered, how they are grounded in matrix context sets, and how they are interact with update and compositional semantics are all intricate, so I don’t engage them here.

stronger record, had the home advantage, etc. This state of mind puts me in position to assert things like:

(3) I think/believe the Warriors won.

(4) The Warriors probably won.

It does not, however, put me in position to assert:

(5) The Warriors won.

For example, if you ask me “Who won the game?” the mere fact that (3) is true does not put me in position to assert (16), as BD would suggest. This basic point is made by [Hawthorne, Rothschild, and Spectre \[2016\]](#) and [Rothschild \[2020\]](#), who put it as the claim that belief is *weak*. (More on what “weak” means below.)

## 5.2 The insufficiency of common belief for common ground

BD predicts that as a default, if  $p$  is commonly believed, then  $p$  is common ground. This appears not to be the case. Consider the following exchange:

Alice: I believe that the movie starts at 7, but it might start later.

Bert: I also believe that it starts at 7, but I’m not sure either.

Alice: Well, I guess we should call and check.

If common belief in the proposition that the movie starts at 7 is achievable at all, the first two steps in this exchange seem as apt as anything to produce it. After all, each agent directly reports their belief in that proposition to the other, and each conversationally accepts what the other says. But it is clear that this proposition does not become taken for granted in this conversation in the way that it would were it common ground. For it is obvious that possibilities where the movie starts later are being treated as conversationally open. If such possibilities weren’t still open, Alice’s suggestion that they call and check wouldn’t make sense or sound right, and the second conjuncts of each of what Alice and Bert say would sound weird. The surest way to add the proposition that the movie starts at 7 to the common ground is (again following [Stalnaker \[1978\]](#)) to just assert that proposition. But as is familiar, and in contrast to what we see above, these are not felicitous:

(6) ??The movie starts at 7, but I’m not sure it does.

(7) ??The movie starts at 7, but it might start later.

The upshot is that it can be common ground that  $p$  is believed by all the interlocutors, and yet not common ground that  $p$ —even in a conversation where no move made (“Suppose the following...”, etc.) to override the default about what is being accepted for the purposes of the conversation. This suggests that whatever the default is, common belief in a proposition is not a sufficient condition for its entry into the common ground.<sup>17</sup>

There is no comparable objection to KD. If we replace ‘believe’ with ‘know’ throughout Alice and Bert’s exchange, the result makes little sense.

What all this suggests is that taking  $p$  to be common ground is in a certain sense a stronger posture to strike towards  $p$  than just (to some mutual degree) believing it. Taking  $p$  to be common ground involves excluding  $\neg p$  worlds from relevant consideration,<sup>18</sup> whereas belief—and with it, mutual and common belief—apparently does not require that. This is as expected, given the observations of Hawthorne et al. [2016] and Rothschild [2020]. (See also Beddor and Goldstein [2018], Bledin and Lando [2018].)

A related example to consider in this connection is this exchange:

- (8) Alice: It’s raining.
- (9) Bert: I very strongly suspect you are right about that!
  - (a) So let’s assume that it is.
  - (b) So let me just double check the weather app on my phone.

Bert’s initial reply affirms his confidence in what Alice said, and yet is compatible with his not being willing to add the proposition Alice asserted to their common ground just yet, as in continuation (b). But even if his continuation were (a), we have the sense that although Bert is willing to allow the proposition that it’s raining into the common ground, he has intervened on the default setting, marking this content as merely assumed, not common knowledge. “I agree” and “Let’s assume you’re right” are both ways of letting an asserted proposition into the common ground, but the second appears to be a way of excepting the proposition from being subject to KD.

---

<sup>17</sup>A reviewer observes that the following discourse seems fine:

A: Are you sure it’s raining?  
 B: Yes I’m sure. Aren’t you?  
 A: Oh yes, I’m sure too. Should we look and see?

As the reviewer notes, this would suggest that “common surety” also does not suffice for common ground status.

<sup>18</sup>Though see Ciardelli [2020] for some further subtleties.

### 5.3 Two senses of weak/strong

To avoid confusion, let me pause here to separate out two senses of “strong” (“weak”). [Hawthorne, Rothschild, and Spectre \[2016\]](#) use these terms in talking about “evidential standards”: their claim is that “the evidential standards that are required for belief are very low” and that “the evidential standards for assertion, whatever they are, are much higher than those for belief” (1394). A presupposition of their discussion is that belief and assertion are subject to evidential standards in (close enough to) the same sense, such that comparison of the strength of the standard makes sense. But this presupposition is not obviously correct. For whereas it is presumably some kind of conceptual truth that rationally believing  $p$  is believing in accord with the evidential norms or standards (whatever they are) governing belief, the same is not true for assertion: rationally asserting  $p$  is *not* just asserting in accord with the epistemic norms/standards (whatever they are) governing assertion. Like any other action, the rationality of asserting something is tied up with its practical utility. Relatedly, belief seems constitutively tied up with evidential standards; but on the view carved out here, assertion is not. It is subject only to a default norm, one derivative from KD (a norm that directly regulates, not assertion, but the common ground). The sense in which belief is “subject to evidential standards” thus shakes out as quite different from the sense in which assertion is. This makes meaningful comparison far from straightforward.

Fortunately there is another weak/strong distinction in the vicinity. Consider this definition, restricting to attitude operators that embed epistemic modals, and where  $\Diamond$  corresponds with epistemic possibility:

An attitude operator  $A$  is *strong* just in case for any (nonmodal)  $\phi$ ,  $\{A\phi, A \Diamond \neg\phi\}$  is inconsistent; otherwise it is *weak*.

(We can extend this distinction in the obvious way from attitude verbs to the corresponding attitude states.) This *modal* weak/strong distinction is a logical distinction between attitude ascriptions in interaction with epistemic possibility modals. It is not, at least in any direct sense, one about the (nonlogical) epistemic norms or standards governing the attitudes. I see the claim that belief operators are weak in this modal sense as the central empirical claim of [Hawthorne, Rothschild, and Spectre \[2016\]](#) and of [Beddor and Goldstein \[2018\]](#). It is fundamentally the modal claim I take to be at work in the relevant data cited above.

In contrast to belief operators, knowledge operators seem modally strong. This ties in with the decades-old point ([Unger \[1975\]](#), [Dretske \[1981\]](#)) that “To

speak of... knowledge despite uneliminated possibility of error, just *sounds* contradictory” [Lewis, 1996, 549]. For instance:

(10) ??I know it’s heads, but it might be tails.

The entertainment attitudes also appear to be modally strong. Invitations like this seem difficult to make sense of (cf. Yalcin [2007]):

(11) Imagine that it’s not raining. ??Also, imagine that it might be.

Given that conversational acceptance is a variety of entertainment attitude, data like this lead us to expect acceptance to be modally strong. By all appearances, it is: this is at the root of the familiar point that epistemically contradictory discourses are infelicitous (Veltman [1996]):

(12) ??It’s not raining. It might be raining.

—and as we observed already with (7) in the movie example above. Observe here that although acceptance is modally strong, it is not plausibly strong in the sense of Hawthorne et al. [2016]—it is not essentially subject to high evidential standards. (Indeed, the state seems not essentially subject to *any* evidential standards: one can accept as one pleases, just as one can imagine as one pleases.)

Note that this modal weak/strong distinction applies to attitudes/attitude operators only. It is not defined for activities like assertion. So it does not make sense to ask: “Is assertion modally weak?” The nearest sensical question is just the question whether conversational acceptance is strong—and as just noted, it plausibly is.

If belief is modally weak but acceptance is modally strong, that is a significant logical difference between the two. This mismatch presents difficulties for understanding how belief could ever be the tone or setting for the common ground (let alone the default setting, as BD has it). If belief is weak, then (to use Beddor and Goldstein [2018]’s example) one can believe both of these:

(13) The house is empty.

(14) The house might not be empty.

Suppose I do. Then in response to the question “Is the house empty?” I am in position to assert *either* (13) or (14) compatible with BD. But given acceptance is modally strong, I can’t assert both, since these two things I believe are not jointly acceptable. How then should I answer? Shall I pick one at random? This isn’t what people do, though it would appear fine by the lights of BD.<sup>19</sup>

---

<sup>19</sup>People also don’t say:



The point is that the difference in modal strength between belief and acceptance means that the latter is essentially limited in its power to accurately reflect the former. One cannot conversationally accept what one believes, when believes both (13) and (14). The problem does not arise for KD, because knowledge is strong.

## 5.4 Asserting what it makes sense to assume

Let me turn now to an objection concerning one of the data points cited above against BD. Recall the case where I am confident that the Warriors beat the Knicks based on some general background facts, though I don't take myself to know they won. I alleged that in such a case, contra BD, I am not in position to reply to the question "Who won the game?" with (16) (repeated):

(16) The Warriors won.

Is this claim correct? Mandelkern and Dorst [2022] bring into focus an interesting way that something like (16) can be improved.<sup>20</sup> Suppose I asserted (16) without taking myself to know it, but followed this up immediately with (17):

(17) After all, they have a much stronger record, and they had the home advantage.

We might have the intuition that despite my epistemic condition, the whole discourse (16)+(17) is assertable for me after all. But then that would seem to show (16) is indeed assertable without knowledge.

Before analysis, it helps to have a second example. Suppose Bill and Ted are waiting for a subway train. They are tired and eager to get home. It's way past midnight, and everyone knows the R line is unpredictable this late; sometimes service is even suspended without notice. They don't actually know whether the train is about to come (and it's common knowledge that they don't know). After around 30 minutes of waiting, Bill utters the following sentences (in order):

(18) It's not coming.

(19) It's been over half an hour.

(20) Let's get a taxi.

---

(15) The house is empty. ??But I think it might not be.

—though it is unclear why not, if BD could be true.

<sup>20</sup>See Mandelkern and Dorst [2022, 4-5], examples (8), (9), (10), (12), (15).

This whole discourse seems fine. But in saying (18), Bill says something he doesn't know, and doesn't take himself to know.

What is happening in these two examples? I suggest roughly this: first, something is said,  $p$ , that isn't known by the speaker, but which the speaker thinks would make sense to assume, given the interests and purposes of the conversation. The speaker then follows this up with an assertion of some  $q$  he does know, and that is intended to bring out his reason (evidence), or one of his reasons, for thinking it would make sense to assume  $p$ . The implicitly understood rhetorical structure is thus roughly:

[It would make sense to assume]  $p$  [because]  $q$

The saying of  $q$  additionally signals that  $p$  was *not* something the speaker necessarily took himself to know. This seems clear from these examples. If one says (16) and stops talking, one creates the impression one knows the result of the game; but if one elaborates with (17), this makes it transparent that one's evidence does not quite secure knowledge of (16). Similarly with (18) when followed by (19). With the general structure clear, further examples are easy to multiply.

The idea that a sequence of sentences in discourse is usually heard to be organized or implicitly structured by some sort of rhetorical or coherence relations is familiar, and independently plausible (see, e.g., Hobbs [1979, 1985], Asher and Lascarides [2003], Kehler [2006]). For just one example, if you say "Tell me about your brother", and I reply:

(21) Well, he just went to Istanbul. ??He really loves tacos.

My reply is weird, though both of my claims are perfectly relevant to your request. The weirdness owes to the sense we have that the second sentence is supposed to explain or otherwise cohere with the first sentence, and it doesn't.

Various theories of coherence relations in discourse exist. Without assuming any particular one, my suggestion is just that the data in question fit a recognizable and apparently conventionalized rhetorical schema.

KD is a default. It would be challenged by the observation of random exceptions, but not by the observation that there are systematic, predictable ways we make exceptions. Examples like (16) and (18) are of the second sort. The departure is highly constrained, and in a sense localized to (16) and (18). That is, although one can say (16) without knowing it if one follows it up with (17), this is intuitively permissible only one actually *knows* (17); the constraint that

one says only what one knows is already back by the time we get to (17).<sup>21</sup>

It generally makes sense to assume whatever is known, but it often makes sense to assume things that are not known. The question whether it makes sense to assume something is, unlike the question whether we know it, partly a practical question—it’s the question whether acting like we know is apt by our lights to achieve our ends. Whether it makes sense for Bill and Ted to assume (18) is tied in with their wanting to get home already, and with how they weigh the cost of waiting against the cost of the taxi. (If Bill and Ted were instead hired as outside auditors to measure train timeliness, (18) makes less sense to assume, though Bill’s evidence would not be relevantly different.)

Often the common ground of a conversation will involve a mix of propositions that are common knowledge and propositions that are commonly known only to make sense to assume. But judging by the examples recently discussed, we seem sensitive in discourse to the difference—we seem to take the trouble to flag whether the proposition we are proposing to add to the common ground falls short of knowledge.

## 5.5 Strong belief?

Consider now the question whether the proponent of BD can respond to these worries by explicitly framing their proposal in terms of some non-weak notion of belief. I will put the idea like this:

SBD. As a default,  $p$  is common ground in  $g$  iff  $p$  is commonly strongly believed in  $g$ .

Thus strong belief is the state iterated in the statement of the default. The task is to explain what “strong” belief is, given that natural language belief talk suggests a weak notion, and is therefore an infirm guide.

---

<sup>21</sup>Forced guessing seems constrained in much the same way. As [Mandelkern and Dorst \[2022\]](#) observe, one may assert  $p$  without knowing it if  $p$  is one’s guess in answer to question  $Q$ , and one is explicitly asked to guess the answer to  $Q$ . But the permission to assert merely what one guesses-true is localized to  $Q$ , as is obvious from the weirdness of discourses like:

Mark: What’ll John bring for dinner?  
Liam: Not sure.  
Mark: Well, what’s your guess?  
Liam: Indian. That’s his favorite.  
Mark: Oh really? How do you know that’s his favorite?  
Liam: ??I don’t. That’s just my guess. I mean, John loves spicy food.  
Mark: I didn’t know he likes spicy food.  
Liam: ??Well, I actually don’t know whether he does. Just my guess.

What must strong belief be like, to do the work needed? If we want it to help us to explain the problem with (6), strong belief must imply certainty (understood as a state expressible with ‘I’m sure that’). (Though see fn. 17.) If we want it to help us explain the problem with (7), the state of strongly believing  $p$  must be incompatible with, or at least in some kind of intelligible tension with, the state of strongly believing that it might not be that  $p$ . There are intuitively describable states that are candidates for this job that are not knowledge: for instance, *certainty* (cf. Stanley [2008]), or perhaps the state of *believing that you know* (cf. Williamson [2000], Stalnaker [2006]).

I doubt that either of these ways of cashing out SBD can do as well as KD explaining the basic problem with Moorean sentences like (1), or explaining the propriety of the question “How do you know?” in response to assertion (i.e., (2)). But the literature on the norm of assertion contains rather extensive investigation of the empirical adequacy of the knowledge norm and its rivals in respect of predicting a range of data of this sort, and there is not space here to properly enter that fray.<sup>22</sup>

So I will press instead a *prima facie* advantage KD has that is more conceptual in character. Whatever the default norm of the common ground is, there is an important question to answer about it: why that norm rather some other (or none at all)? A theory of the default norm that fits the data may yet not make sense of the data, if it remains unclear why we would have any rational interest in the norm theorized. Whatever the norm is, it shouldn’t be mysterious why it exists, or why we have it instead of another.

If the norm is KD, we can say something straightforward. KD is the norm because that is the norm that most directly facilitates our interest in using the common ground to transmit knowledge—and we have a strong interest in transmitting knowledge. We value knowledge in itself, its possession facilitates the achievement of our aims, and (as we observed above) common knowledge enables coordination in a way that mere common belief could not.

What sort of answer can we give instead, if BD or SBD are assumed? How would one of these norms tend to serve our interests better, enough anyway that it makes sense that we would adopt one of them rather than KD? From the perspective of these norms, the default role of common ground would be to facilitate the sharing of belief or of confidence. What is the value of coordinating on (strong) belief *per se*, in abstraction from whether the (strong) belief constitutes knowledge? And why think whatever value it would have exceeds the value KD would offer?

---

<sup>22</sup>For a sense of the fray, see Pagin [2016], Turri [2017], Kneer [2018], and citations therein.

I will have to be satisfied with framing the question. The pathways for responding here are diverse, and in a way that is incompatible with a comprehensive assessment. If knowledge were assumed to be a scarce commodity, for instance, a belief norm might make better sense than KD. But that’s a place where the debate becomes entangled with questions beyond the scope of the paper, about how ubiquitous or context-sensitive knowledge is, and about how to respond to the skeptic. Or again, one might argue that BD or SBD has better prospects than KD in respect of cohering with an independently attractive Bayesian and probabilistic representation of agents and of rational agency. But that also ties in with a complex debate—this time, about whether credence can itself constitute knowledge (Moss [2013, 2018]). Furthermore, the sort of value BD or SBD would offer as compared to KD may depend on what view we take of the norms governing (strong) belief—another complex issue beyond our scope.<sup>23</sup>

I will therefore avoid trying to play whack-a-mole against all the ways I can imagine a developer of BD or SBD rationalizing it over KD. Instead I settle for a more modest conclusion, which is merely that KD can give a comparatively straightforward answer to the assigned question. Combined with the point that KD makes comparatively short work of the data reviewed in sections §5.1-2, the initial appearances are that KD makes good sense of the facts. Whether SBD can be developed in ways that match or exceed the performance of KD is a question that we can leave open. We have framed the sort of work such an account would need to do.

## 6 Solidarity forever

I have spent the paper developing alternatives to two ideas in Stalnaker [2014], so let me anticipate and dispel a misapprehension this could encourage. If we have shown that  $CG=CKA$  has benefits over  $CG=CA$ , and/or that KD has advantages over BD, these ought to be read as contributions to the basic program Stalnaker initiated. The ambition here was not to undermine the program and replace it with something new. The view described in §3 and defended in §4 of course differs from the letter of Stalnaker [2014], and in ways that (I have argued) matter; but its main contours are obviously Stalnakerian.<sup>24</sup> Furthermore, if §5 was on track, it brought into focus the way in which Stalnaker’s model, by

---

<sup>23</sup>For instance, if belief were subject to the norm that one should believe  $p$  just in case one knows it (a view often associated with Williamson [2000]), then the predictions of BD and KD seem difficult to distinguish.

<sup>24</sup>It also fits naturally with the “knowledge first” approach advanced in Stalnaker [2015, 2019a] (see also the introduction to Stalnaker [2019b]).

enabling us to frame the question of the default norm(s) governing common ground, can shed light on the question of the norm of assertion. If the default norms governing common ground can explain most, or all, of the facts that have traditionally been explained by appeal to the idea that assertion is essentially governed by norms, this clears the way for the view that the account of assertion provided in [Stalnaker \[1978\]](#) does not in fact leave something essential about assertion out.<sup>25</sup>

## References

- Nicholas Asher and Alex Lascarides. *Logics of conversation*. Cambridge University Press, 2003.
- Bob Beddor and Simon Goldstein. Believing epistemic contradictions. *Review of Symbolic Logic*, 11(1):87–114, 2018.
- Justin Bledin and Tamar Lando. Epistemic modals and closure. *Philosophy and Phenomenological Research*, XCVII(1):3–22, 2018.
- Ivano Ciardelli. Indicative conditionals and graded information. *Journal of Philosophical Logic*, 49(3):509–549, 2020.
- Keith DeRose. Epistemic possibilities. *Philosophical Review*, 100(4):581–605, 1991.
- Keith DeRose. Assertion, knowledge, and context. *Philosophical Review*, 111(2):167–203, 2002.
- Fred Dretske. The pragmatic dimension of knowledge. *Philosophical Studies*, 40:363–78, 1981.
- Ronald Fagin, Joseph Y. Halpern, Yoram Moses, and Moshe Y. Vardi. *Reasoning about Knowledge*. MIT Press, Cambridge, MA, 1995.
- Ronald Fagin, Joseph Y. Halpern, Yoram Moses, and Moshe Y. Vardi. Common knowledge revisited. *Annals of Pure and Applied Logic*, 96(1-3):89–105, 1999.
- Gottlob Frege. Thought. In M. Beaney (with translation by P. Geach and R. H. Stoothoff), editor, *The Frege Reader*, pages 325–45. Blackwell 1997, 1918.

---

<sup>25</sup>For very helpful conversations on the topics of this paper, many thanks to participants in my graduate seminars at Berkeley in 2021 and 2023, and to Bob Beddor, Sam Berstler, Daniel Hoek, Anhui Huang, Cameron Johnson, Amanda Lopatin, John MacFarlane, Emily Massey, Milan Mosse, Robbie Williams, Stephen Yablo, and Snow Zhang. For detailed comments on earlier drafts, I am indebted to three anonymous reviewers, and especially to Harvey Lederman and Paolo Santorio. Thanks finally to Bob Stalnaker for enduring earlier drafts of this work, for generous engagement on these issues over the years, and of course, for the pioneering work this paper responds to.

- Daniel Greco. Safety, explanation, iteration. *Philosophical Issues*, 26(1):187–208, 2016.
- H Paul Grice. Logic and conversation. In Peter Cole and Jerry L. Morgan, editors, *Syntax and Semantics, Volume 3: Speech Acts*, pages 41–58. Academic Press, 1975.
- John Hawthorne and Ofra Magidor. Assertion, context, and epistemic accessibility. *Mind*, 118(470):377–397, 2009.
- John Hawthorne, Daniel Rothschild, and Levi Spectre. Belief is weak. *Philosophical Studies*, 173(5):1393–1404, 2016.
- Irene Heim. *The Semantics of Definite and Indefinite Noun Phrases*. PhD thesis, University of Massachusetts, 1982.
- Jerry Hobbs. Coherence and coreference. *Cognitive Science*, 3(1):67–90, 1979.
- Jerry Hobbs. On the coherence and structure of discourse. Technical Report Tech. Rep. CSLI-85-37, Stanford University, 1985.
- Ben Holguín and Harvey Lederman. Trying without fail. Unpublished manuscript, February 2023.
- H. Kamp. Comments. In R. Grim and D. Merrill, editors, *Contents of Thought*, pages 156–181. 1988.
- Andrew Kehler. Discourse coherence. *The Handbook of Pragmatics*, pages 241–265, 2006.
- Markus Kneer. The norm of assertion: Empirical data. *Cognition*, 177:165–171, 2018.
- Harvey Lederman. Common knowledge. In Marija Jankovic and Kirk Ludwig, editors, *The Routledge Handbook of Collective Intentionality*, pages 181–195. Routledge, 2017.
- Harvey Lederman. Two paradoxes of common knowledge: Coordinated attack and electronic mail. *Noûs*, 52(4):921–945, 2018a.
- Harvey Lederman. Uncommon knowledge. *Mind*, 127(508):1069–1105, 2018b.
- David Lewis. Elusive knowledge. *Australasian Journal of Philosophy*, 74(4):549–567, 1996.
- Matthew Mandelkern and Kevin Dorst. Assertion is weak. *Philosophers’ Imprint*, 22(19), 2022.
- Sarah Moss. Epistemology formalized. *Philosophical Review*, 122(1):1–43, 2013.
- Sarah Moss. *Probabilistic Knowledge*. Oxford University Press, 2018.

- Robert Nozick. *Philosophical Explanations*. Belknap Press, 1981.
- Peter Pagin. Problems with norms of assertion. *Philosophy and Phenomenological Research*, 93(1):178–207, 2016.
- Duncan Pritchard. *Epistemic Luck*. Clarendon Press, 2005.
- Daniel Rothschild. What it takes to believe. *Philosophical Studies*, 177(5):1345–1362, 2020.
- Ernest Sosa. How to defeat opposition to moore. *Philosophical perspectives*, 13:141–153, 1999.
- Robert Stalnaker. Pragmatics. *Synthese*, 22(1-2):272–89, 1970.
- Robert Stalnaker. Pragmatic presuppositions. In Milton K. Munitz and Peter Unger, editors, *Semantics and Philosophy*. New York University Press, 1974.
- Robert Stalnaker. Indicative conditionals. *Philosophia*, 5(3):269–286, 1975.
- Robert Stalnaker. Assertion. In Peter Cole, editor, *Syntax and Semantics 9: Pragmatics*, pages 315–332. Academic Press, 1978.
- Robert Stalnaker. *Inquiry*. MIT Press, Cambridge, MA, 1984.
- Robert Stalnaker. Belief attribution and context. In Robert Grimm and Daniel Merrill, editors, *Contents of Thought*. University of Arizona Press, 1988.
- Robert Stalnaker. On the representation of context. *Journal of Logic, Language, and Information*, 7:3–19, 1998.
- Robert Stalnaker. Common ground. *Linguistics and Philosophy*, 25(5-6):701–21, 2002.
- Robert Stalnaker. On logics of knowledge and belief. *Philosophical Studies*, 128(1):169–199, 2006.
- Robert Stalnaker. On Hawthorne and Magidor on assertion, context, and epistemic accessibility. *Mind*, 118(470):399–409, 2009.
- Robert Stalnaker. *Context*. Oxford University Press, Oxford, 2014.
- Robert Stalnaker. Luminosity and the KK thesis. In Sanford C. Goldberg, editor, *Externalism, Self-Knowledge and Skepticism*, pages 17–40. Cambridge University Press, 2015.
- Robert Stalnaker. Contextualism and the logic of knowledge. In *Knowledge and Conditionals*, pages 129–148. Oxford University Press, 2019a.



- Robert Stalnaker. *Knowledge and Conditionals: Essays on the Structure of Inquiry*. Oxford University Press, 2019b.
- Jason Stanley. Knowledge and certainty. *Philosophical Issues*, 18:35–57, 2008.
- John Turri. *Knowledge and the norm of assertion: An essay in philosophical science*. Open Book Publishers, 2016.
- John Turri. Experimental work on the norms of assertion. *Philosophy Compass*, 12(7):e12425, 2017.
- Peter Unger. *Ignorance: a case for skepticism*. Oxford University Press, Oxford, 1975.
- Frank Veltman. Defaults in update semantics. *Journal of Philosophical Logic*, 25(3): 221–61, 1996.
- Timothy Williamson. Knowing and asserting. *Philosophical Review*, 105(4):489–523, 1996.
- Timothy Williamson. *Knowledge and its Limits*. Oxford University Press, Oxford, 2000.
- Seth Yalcin. Epistemic modals. *Mind*, 116(464):983–1026, 2007.
- Seth Yalcin. Common knowledge first. Unpublished manuscript, 2023.