New Horizons for a Theory of Epistemic Modals

Justin Khoo & Jonathan Phillips

Australasian
Association of
Philosophy

Routledge
Taylor & Francis Group

Check for updates

New Horizons for a Theory of Epistemic Modals

Justin Khoo[a] and Jonathan Phillips[b]

[a]Massachusetts Institute of Technology; [b]Harvard University

ABSTRACT
Recent debate over the semantics and pragmatics of epistemic modals has focused on intuitions about cross-contextual truth-value assessments. In this paper, we advocate a different approach to evaluating theories of epistemic modals. Our strategy focuses on judgments of the incompatibility of two different epistemic possibility claims, or two different truth value assessments of a single epistemic possibility claim. We subject the predictions of existing theories to empirical scrutiny, and argue that existing contextualist and relativist theories are unable to account for the full pattern of observed judgments. As a way of illustrating the theoretical upshot of these results, we conclude by developing a novel theory of epistemic modals that is able to predict the results.

1. Introduction

At the core of the recent debate over the semantics of modals is a tension between modal language and its subject matter. Consider the following modal sentence:

(1) The keys might be in the drawer.

On the epistemic interpretation of 'might', this sentence seems to state that it is epistemically possible that the keys are in the drawer. However, it also seems that there are no such absolute epistemic possibility facts; rather, propositions are only epistemically possible relative to a body of evidence or information. Suppose that Jones doesn't know whether the keys are in the drawer, but Smith has seen that they aren't. Then, relative to Jones's evidence, it is epistemically possible that the keys are in the drawer, but, relative to Smith's evidence, it is *not* epistemically possible that the keys are in the drawer. Thus, we face a mismatch between the subject matter of epistemic modality and the language that we use to talk about it: propositions are only epistemically possible relative to a body of evidence, but such relativization is not explicitly marked in sentences like (1).

Existing semantic theories of epistemic modals resolve this tension in different ways. For instance, modal contextualism proposes that epistemic modal sentences like (1) do in fact express propositions about particular bodies of evidence. Contextualism holds that *which* body of evidence a particular epistemic modal claim is about depends on the context in which it is uttered [Hacking 1967; Kratzer 1981, 1991; DeRose 1991, Stanley 2005; Dowell 2011].

By contrast, modal relativism resolves the tension by proposing that epistemic modal sentences express propositions that are not about particular bodies of evidence, but are instead true or false relative to bodies of evidence (see Egan [2005, 2007] and Macfarlane [2011, 2014]). An important commitment of relativism is that the same epistemic modal proposition may be true as assessed in one context and false as assessed in another.

The debate between modal contextualism and relativism has focused on cross-contextual assessments of speech acts involving bare epistemic modal sentences like (1). These are cases where a better-informed eavesdropper evaluates the truth value of an epistemic modal claim made by a lesser-informed speaker in a different context. Our aim in this paper is to shift the perspective on this debate and to advocate a different way of evaluating contextualist and relativist theories of epistemic modals. In section 2, we argue for testing the predictions of contextualism and relativism by exploring intuitions about the (in)compatibility of epistemic modal claims across contexts. In section 3, we report the results of a new empirical study, which reveals new data that neither existing contextualist nor relativist theories predict. This motivates the search for an alternative to orthodox contextualist and relativist theories; we sketch one such theory in section 4. We conclude by considering how our approach might generalize to a broader range of linguistic expressions.

2. Distinguishing Contextualism from Relativism

Throughout, we focus on epistemic possibility sentences like (1). We make the standard assumption that such sentences are composed of a modal, 'might', scoping over a sentence p. We use corner quotes, as in '⌜might p⌝', to denote the sentence schema whose substitution instances are epistemic possibility sentences like (1): in such a sentence, 'might' is the modal and p is its prejacent. Lowercase italic letters $p, q, r, \ldots$ are used as variables over propositions. Finally, we assume that the truth of such sentences depends both on a modal base f (a function from a world to a set of worlds: cf. Kratzer [1981]) and a world $\mathbf{w}$, as follows:

> *Simple Semantics*
> ⌜Might p⌝ is true at $f,\mathbf{w}$ iff $\exists \mathbf{w}'$ in $f(\mathbf{w})$: p is true at $f,\mathbf{w}'$.

Think of an epistemic modal base as a characterization of a body of information or evidence: what Jones knows, the information compatible with the security camera footage, and so on. Thus, we can read Simple Semantics as stating that ⌜might p⌝ is true relative to a world and body of information iff p is compatible with that body of information (at that world). Contextualist and relativist theories differ as to how modal bases enter into the semantics and pragmatics of epistemic modals.

Orthodox contextualist theories of epistemic modals are distinguished by two main claims:

> **Context Dependence**. Epistemic modal sentences are context dependent: in different contexts, the same epistemic modal sentence may express different propositions.

> **Truth Absolutism**. The proposition expressed by an epistemic modal sentence has its truth value fixed by the world. In other words, it determines a set of possible worlds.

Thus, contextualist theories hold that an epistemic modal sentence expresses a proposition that is about a particular modal base, which is determined by the context in which it is uttered. For instance, in context **c1**, the sentence $\ulcorner$might p$\urcorner$ says that p is compatible with evidence $f_{\mathbf{c1}}$ (that is, it expresses the proposition $\diamondsuit_{\mathbf{c1}}\, p$); in context **c2**, it says that p is compatible with evidence $f_{\mathbf{c2}}$ (that is, it expresses the proposition $\diamondsuit_{\mathbf{c2}}\, p$); and so on.

By contrast, standard relativist theories of epistemic modals deny both **Context Dependence** and **Truth Absolutism**, and instead endorse:

> **Context Invariance**. Epistemic modal sentences are context invariant: in different contexts, the same epistemic modal sentence expresses the same proposition (so long as its prejacent is not context-dependent).

> **Truth Relativism**. The truth value of the proposition expressed by an epistemic modal sentence can vary across dimensions of evaluation at the same world.

Thus, according to the relativist, in every context $\ulcorner$might p$\urcorner$ says that p is epistemically possible. In other words, $\ulcorner$might p$\urcorner$ expresses the proposition $\diamondsuit p$, which is true relative to evidence f iff p is compatible with evidence f.[1]

The differences between these theories matter when it comes to norms governing the acceptance of propositions. Because of their commitment to **Truth Absolutism**, contextualists can appeal to a norm involving truth, knowledge, or justified belief that is not relativized to contexts. Here is one possible such norm:[2]

> *Truth Norm*
> Accept p only if p is true.

However, the relativist's norm of acceptance makes important use of her commitment to epistemic modal propositions varying in truth value across contexts of assessment:

> *Relativist Truth Norm*
> Accept p at **c** only if p is true at **c**.

Relativist Truth Norm entails that whether person A ought to accept $\diamondsuit p$ depends on whether the evidence in A's context is compatible with p. By contrast, Truth Norm only applies to propositions whose truth value is fixed by the world: it says that whether A ought to accept $\diamondsuit_{\mathbf{c}}\, p$ depends on whether $\diamondsuit_{\mathbf{c}}\, p$ is true—that is, on whether the evidence in **c** is compatible with p. Say that an orthodox contextualist theory is one that endorses **Context Dependence**, **Truth Absolutism**, and Truth Norm, and that an orthodox relativist is one that endorses **Context Invariance**, **Truth Relativism**, and Relativist Truth Norm. Thus understood, orthodox contextualist and relativist theories make different predictions about assessments of epistemic modal claims across contexts.

---

[1]Formally, this difference amounts to a difference in the formal object that the theories use to model the propositions expressed by epistemic modal sentences. For contextualists, the proposition expressed by $\ulcorner$might p$\urcorner$ at **c** just is an ordinary possible-worlds proposition: $\{\mathbf{w}\colon \exists \mathbf{w'} \in f_c(\mathbf{w})\colon p$ is true at $f,\mathbf{w'}\}$. But, for relativists, the proposition expressed by $\ulcorner$might p$\urcorner$ (in any context) is not reducible to a possible-worlds proposition. Instead, it can be modelled as a world/evidence-proposition: $\{<\mathbf{w},f>\colon \exists \mathbf{w'} \in f(\mathbf{w})\colon p$ is true at $f,\mathbf{w'}\}$.

[2]We will focus on truth norms here, but this is just for the sake of concreteness. We do not want to take a stand on whether, instead, knowledge or justified belief or high credence of truth should be the norm of acceptance.

Before we get to the cases that have been used to distinguish between contextualism and relativism, we articulate two more standard background assumptions that are necessary for generating predictions in such cases. The first is:

> *Says-Talk*
> ⌜X's claim⌝ and ⌜What X said⌝ refer to the content of X's assertion on some particular occasion of utterance.

This is a very orthodox assumption, shared by many contextualists and relativists, that goes back to Kaplan [1989]. One motivation for this hypothesis is the behaviour of indexicals in discourses like this:

(2) A: I am hungry.
    B: What A said is true.

Here, A utters the sentence, 'I am hungry' and thereby asserts that A is hungry. What does B assert? Intuitively, B asserts something equivalent to this: A is hungry. This is predicted by Says-Talk: according to it, B refers to the proposition A asserted, and asserts of it that it is true—as such, B asserts that the proposition that A is hungry is true, and this seems equivalent to asserting that A is hungry.

The second assumption is a deflationary semantics for 'is true' and 'is false':

> *Deflationism*
> The proposition expressed by ⌜What X said is true⌝ (⌜What X said is false⌝) is equivalent to the proposition (the negation of the proposition) referred to by ⌜What X said⌝ on that occasion of use.

This principle is also widely accepted by both contextualists and relativists. We appealed to Deflationism in generating the prediction that B says something equivalent to what A said. Furthermore, this principle does seem intuitively plausible: endorsing what someone says by calling it 'true' does not seem to add anything to what they said. We return to both of these assumptions in more detail below.

2.1 Eavesdropper Intuitions

Given these assumptions, we can use intuitions about cross-contextual eavesdropper cases to test the predictions by orthodox contextualism and relativism. The reason to focus on cross-contextual eavesdropping is that in such cases the Truth Norm and Relativist Truth Norm can come apart, since the assessment takes place in a context that is distinct from the context in which the claim is made. Suppose that A utters ⌜might p⌝ and that B is not part of A's context, but is instead secretly listening in on A's conversation (perhaps via a wiretap). Given Says-Talk and Deflationism, contextualism predicts that whether B ought to think that what A said is true depends on whether the evidence relevant in A's context is compatible with p. By contrast, relativism predicts that whether B ought to think that what A said is true depends on whether the evidence relevant in B's context is compatible with p.

A common intuition alleged in the literature is that it would be correct for B to think that what A said is false, if B has evidence ruling out p, even if A's evidence was

compatible with p (compare Egan et. al [2005], Egan [2007], and Macfarlane [2011, 2014]). This alleged intuition is easily explained by relativism, since, according to relativism, it is B's evidence that matters to whether it would be correct for B to think that what A said is true. However, this intuition is not easily explained by contextualism, since, according to contextualism, it is A's evidence that matters to whether it would be correct for B to think that what A said is true.

Recently, however, this argument in favour of relativism over contextualism has been contested on the ground that this intuition is not widely shared (see Dowell [2011] and Knobe and Yalcin [2014]). In what follows, we discuss the results of Knobe and Yalcin, who empirically tested intuitions in such a case. In their first study, they asked native English speakers to read a vignette about two Experts who viewed skilfully faked evidence that suggested that a mobster named Fat Tony was murdered at the docks. After viewing the evidence, Expert A makes a non-modal claim: 'Fat Tony is dead.' Expert B instead makes a modal claim: 'Fat Tony might be dead.' Participants were asked to rate their agreement with one of the following statements:

- (nonmodal-true) What Expert A said is true.
- (nonmodal-false) What Expert A said is false.
- (modal-true) What Expert B said is true.
- (modal-false) What Expert B said is false.

The results from this experiment are summarized in the following figure:
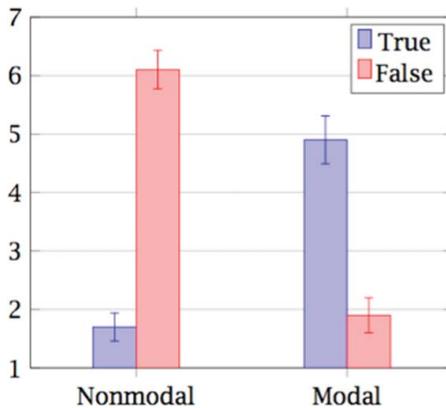

Figure 1. Knobe & Yalcin data: higher scores indicate more agreement

Participants tended to think that what Expert A said was false, and that what Expert B said was true. This is exactly the opposite of what orthodox relativism predicts—namely, that in such cases assessors will use their own evidence (which includes the information that Fat Tony is alive) in assessing what Expert B said, and thus will conclude that what Expert B said is false.

Importantly, Knobe and Yalcin do not claim that these data are evidence that relativism is false; instead, the data are merely presented as evidence that ordinary speakers do not share the eavesdropper intuitions alleged by relativists. This is an important point, and one that bears reemphasizing here.

To see how the results of the experiment are compatible with relativism, notice that it is not a commitment of relativism that the evidence relevant in the assessor's context is *always* the strongest evidence available to her at that time. It is open for the relativist to adopt a flexible position that allows for the relevant evidence in the assessor's context to sometimes be only the evidence that was available to the person uttering the epistemic modal sentence. Along these lines, a relativist may account for the results of Knobe and Yalcin's first experiment by arguing that participants tended to judge that the relevant information (in their context) was the evidence available to Expert B, which suggested that Fat Tony had died. Which information participants took to be relevant when they were assessing these claims was simply not tested, and thus these studies do not provide direct evidence against relativism. At most, data failing to produce the alleged eavesdropper experimentally may count as inductive evidence against relativism.

There might still be other ways of empirically exploring eavesdropping intuitions to test between relativism and contextualism. For example, experimental evidence that a single modal utterance is simultaneously assessed as true and as false by two eavesdroppers with distinct evidence would make for strong evidence against contextualism. Yet the current state of empirical evidence about intuitions in eavesdropping cases does not yet cleanly decide between the two theories. Rather than pursuing alternative variations in testing eavesdropper intuitions, we want to explore what we think is a more promising alternative approach that focuses on judgments of the (in)compatibility of different epistemic modal claims.[3]

2.2 Structural Predictions

Our strategy is to replace truth value judgments with truth-incompatibility judgments. Whereas the truth value predictions of contextualism and relativism depend on some particular version of those theories (in particular, what body of evidence they predict is relevant in the context of utterance or assessment), truth-incompatibility predictions do not—thus, we call these structural predictions. In the remainder of this section, we spell out the structural predictions by contextualism and relativism, respectively.

2.2.1 Compatibility of Modal Utterances

Suppose that X and Y are in different contexts (that is, they are not talking to one another). X has evidence that is compatible with *p*, whereas Y has evidence that rules out *p*. X utters ⌜might p⌝ in her context, while Y utters ⌜not-might p⌝ in her context. Call this a **Modal Utterances** case. Contextualism and relativism make different predictions about the following statement about X and Y's claims:

(Q) At least one of X or Y's claims must be false.

Here is why. According to contextualism, X asserts $\diamondsuit_{cX} \, p$ (that *p* is compatible with the evidence in X's context), while Y asserts $\diamondsuit_{cY} \, p$ (that *p* is compatible with the evidence in Y's context). And it is possible that *p* is compatible with the evidence in X's context and not compatible with the evidence in Y's context—in that case, (Q) would be false.

But it could also be that X's evidence is sufficiently similar to Y's evidence, so that it cannot be that p is compatible with X's evidence and not compatible with Y's—in that case, (Q) would be true. Thus, contextualism can predict either that (Q) is true or that (Q) is false, depending on how we understand the evidence in X and Y's contexts.

By contrast, relativism entails that X and Y's claims are incompatible. According to relativism, what Y asserts (call it $\Diamond p$) just is the negation of what X asserts ($\sim\Diamond p$). But then, relative to any context of assessment, at least one of their claims must be false. This is because no world and context of assessment $\mathbf{c}_A$ is such that p is compatible with the evidence of $\mathbf{c}_A$ and not compatible with the evidence of $\mathbf{c}_A$. Therefore, relativism always predicts that (Q) is true.

2.2.2 Compatibility of Modal Assessments

Suppose that X and Y are again in different contexts, listening in on a conversation taking place at a third context in which Z utters ⌜might p⌝. As before, X has evidence that is compatible with p, whereas Y has evidence that rules out p. X utters 'What Z said is true', while Y utters 'What Z said is false.' Call this a **Modal Assessments** case. In this kind of case, both contextualism and relativism predict that (Q) is true.

According to contextualism, when Z says ⌜might p⌝, he asserts that p is compatible with the evidence in his (Z's) context. Given Says-Talk and Deflationism, when X says 'What Z said is true', she asserts a proposition equivalent to the one that Z asserted; and when Y says 'What Z said is false', she asserts the negation of this proposition. Since at least one of these propositions must be false, (Q) follows.

According to relativism, when Z says ⌜might p⌝, he asserts $\Diamond p$. By Says-Talk and Deflationism, when X says 'What Z said is true', she asserts a proposition equivalent to $\Diamond p$, and when Y says 'What Z said is false', she asserts a proposition equivalent to $\sim\Diamond p$. Since, relative to any context of assessment, at least one of $\Diamond p$ or $\sim\Diamond p$ must be false, (Q) follows.

We summarize these predictions of contextualist and relativist theories as follows:

	Contextualism	Relativism
Modal Utterances	✓/✗	✓
Modal Assessments	✓	✓

It is an empirical question which, if any, of these predictions of contextualism and relativism are accurate. We explore this question in the next section.

3. Testing the Structural Predictions

3.1 Participants

To test the pattern of predictions discussed above, we conducted an experiment. Two hundred and forty participants (M_{age} = 37.93, SD_{age} = 12.11, 49% female) were recruited through Amazon Mechanical Turk (www.mturk.com). All of the materials used in this survey, along with the, data, analysis code, and complete demographic information can be retrieved from: https://osf.io/mhsfe/.

3.2 Design

All participants were told about a case in which the police are investigating a rumour that a local mobster, Fat Tony, died at the docks. The Chief of Police assigns Inspector A to examine the evidence at the docks, and assigns Inspector B to examine the footage from the security camera. The evidence that Inspector A finds at the docks suggests, but does not prove, that Fat Tony died at the docks; the evidence that Inspector B finds from the security footage proves that Fat Tony did not actually die at the docks.

Within this setup, participants were randomly assigned to read one of six different vignettes that concerned either two **Utterances** or two **Assessments** of one of three different kinds of claims—an epistemic **Modal** claim, a **Non-Modal** claim, or an **Indexical** claim.

In the **Modal Utterances** vignette, Inspector A reviews the evidence at the docks and says, 'Fat Tony could have died at the docks.' and Inspector B reviews the evidence from the security camera and says, 'Fat Tony couldn't have died at the docks.'[4]

By contrast, in the **Modal Assessments** vignette, the Chief of Police says, 'Fat Tony could have died at the docks' during a news interview that both Inspectors are watching in their respective homes. Inspector A, who has seen the evidence from the docks, says, 'What the Chief said is true', and Inspector B, who has seen the evidence from the security footage, says 'What the Chief said false.'

In comparison to these **Modal** cases, other participants read vignettes that concerned **Non-Modal** claims. These cases were identical to the preceding ones except that the Inspectors'/Chief's claim(s) did not include the epistemic modal, and thus read 'Fat Tony died [did not die] at the docks.'

Finally, we compare both kinds of cases with an **Indexical** version, in which an indexical claim is made. In the **Utterances** case, Inspector A says, 'I have served on the police force for twenty years', and Inspector B says, 'I have not served on the police force for twenty years.' In the **Assessments** case, the Chief of Police says, 'I have served on the police force for twenty years', while Inspector A says, 'What the Chief said is true', and Inspector B says, 'What the Chief said is false.'

After participants finished reading one of these six vignettes, they were reminded that the inspectors had made two different claims and were then asked whether they agreed or disagreed that 'At least one of the inspector's claims must be false.' Participants rated their agreement on a scale from 1 ('Completely Disagree') to 7 ('Completely Agree').[5]

---

[4]We changed the target sentence from a modal claim about the present ('Fat Tony could be dead') to one about the past ('Fat Tony could have died at the docks') to remove the possibility of thinking that Tony didn't die at the docks but instead died sometime thereafter. However, doing so opens up the possibility that some participants interpret 'could have' metaphysically, rather than epistemically (as we intend). However, we do not think that this interpretive possibility interferes with our results, since the metaphysical interpretation is clearly deviant in the context in which the sentence is uttered (the Inspector is asked to report what he has found in the course of investigating whether Fat Tony died at the docks).

[5]Note that our design is both similar to and differs from that of Katz and Salerno [2017]. Katz and Salerno tested (in)compatibility judgments of two distinct modal utterances (by asking participants whether they agree or disagree that 'X's claim and Y's claim can both be true (at the same time)'). We also ask about the compatibility of modal utterances, but compare (more continuous) rates of agreement with (Q) for both **Utterances** and, additionally, **Assessments** across **Modal**, **Non-Modal**, and **Indexical** variants.

3.3 Results and Discussion

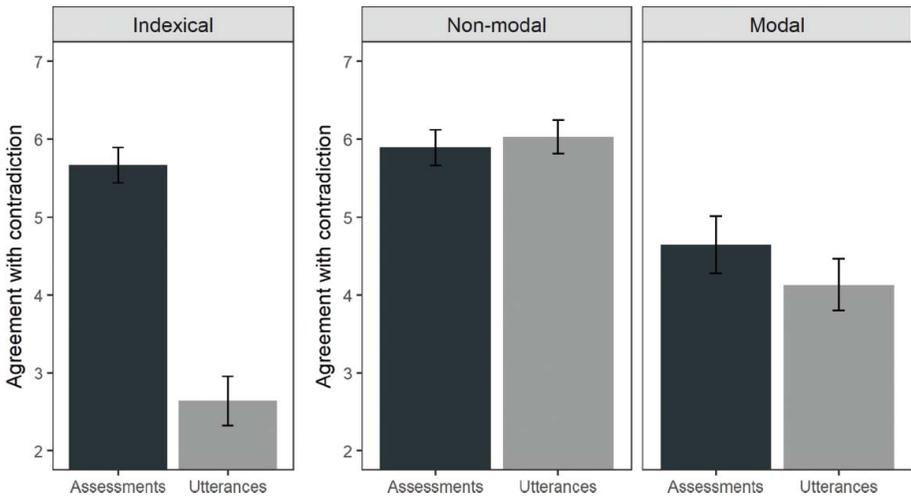Here is the overall pattern of results that we found:

Figure 2. Participants' mean level of agreement that at least one of the inspectors' claims must be false. Errors bars indicate +/- 1 *SE*.

Focus first on the **Modal/Non-Modal** conditions. We analysed participants' compatibility judgments in the Modal and Non-modal conditions with a 2 (Claim: **Modal vs Non-modal**) × 2 (Case: **Utterances vs Assessments**) *ANOVA*. We found that participants' ratings were significantly affected by whether or not the claims involved an epistemic modal, $F(1,140) = 29.204$, $p < .001$, $\eta^2_p = .171$, such that they strongly agreed that one of the inspector's claims must be false when they uttered/assessed a non-modal claim ($M = 5.96$, $SD = 1.33$), but not when they uttered/assessed a modal claim ($M = 4.38$, $SD = 2.10$), $t(119.89) = 5.11$, $p < .001$, $d = .902$. By contrast, we did not observe a significant effect of whether the Inspectors made conflicting utterances or conflicting assessments, $F < 0.5$, and most importantly did not find an interaction effect between these two variables, $F < 1.5$, meaning that the *difference* between the modal and non-modal claims did not significantly differ between the **Assessments** and **Utterances** conditions.

Thus, in line with contextualism, but not with relativism, participants were more inclined to judge that the two modal (vs. non-modal) claims were compatible in the **Utterances** condition.[6] At the same time, in conflict with *both* relativism and contextualism, participants were also more inclined to judge that the two assessments were compatible when they concerned a modal (vs. non-modal) claim in the **Assessments**

---

[6]This result builds on the recent findings of Katz and Salerno [2017], who explored a different variant of compatibility judgments about two utterances across a variety of cases. In the case most similar to ours (the FBI case), they found that about 50% of participants agreed that the two modal utterances could both be true (at the same time). In addition, they also included two 'screening' questions which involved the compatibility of two conflicting utterances of a non-modal or indexical claim. Further complementing the patterns that we observed, only 24% of participants judged that both non-modal utterances could be true, while 97% judged that both indexical utterances could be true. The convergence of these two data sets paints a clear picture of the pattern of judgments about conflicting utterances.

condition. Thus, the above results present a challenge both to contextualist and to relativist theories. Before we turn to presenting our alternative to orthodox contextualism and relativism, we pause here to discuss a move that may be made on behalf of each theory to account for our data.

3.3.1 Defending Orthodox Contextualism?

A contextualist might be tempted to deny Says-Talk, in order to account for the data in the **Assessments** case. In particular, a contextualist might instead hold that ⌜What X said⌝ has two interpretations: on one, it refers to the proposition X asserted, while on the other it refers (in context **c**) to the proposition that the sentence X uttered would have expressed had that sentence been uttered in **c**.[7] Given this assumption, in the **Modal Assessments** case there would be an interpretation of A's utterance of 'What the Chief said is true' in which it expressed the proposition that the Chief's sentence would have expressed, had it been uttered in A's context—namely, that it is possible, given A's evidence that Fat Tony died at the docks. And, likewise, there would be an interpretation of B's utterance of 'What the Chief said is false' in which he says that it is not possible, given B's evidence that Fat Tony died at the docks. These two propositions are not (necessarily) incompatible. Thus, given this assumption about ⌜What X said⌝, the contextualist would be able to predict our results for the **Assessments** conditions.

The trouble with this assumption is that it does not sit well with the pattern of results from the **Indexical** conditions, which are correctly predicted by Says-Talk. Consider the results from the **Indexical Assessments** case, where the two Inspectors made conflicting assessments about the Chief's utterance of 'I have served on the police force for twenty years.' In this case, participants strongly agreed that at least one of the Inspectors' claims must be false ($M = 5.67$, $SD = 1.40$), and agreed much more strongly than in the **Indexical Utterances** case, ($M = 2.64$, $SD = 1.96$), $t(68.69) = -7.82$, $p < .001$, $d = 1.77$ (see Fig. 2). In this **Indexical Assessments** case, Says-Talk predicts that 'What the Chief' said refers to the proposition that the Chief has served on the police force for twenty years. A says of this proposition that it is true, while B says of this proposition that it is false. Since at least one of these claims must be false, Says-Talk correctly predicts strong agreement with (Q). By contrast, if the contextualist response above is correct, we expect there to be an interpretation of A's/B's utterances of 'What the Chief said is true'/'What the Chief said is false' in which A says that A has served on the police force for more than twenty years and B says that B has not served on the police force for more than twenty years. But these two claims are clearly compatible.[8] We take these data both to confirm Says-Talk and to cast doubt on this kind of response on behalf of the contextualist.

3.3.2 Defending Orthodox Relativism?

A relativist might be tempted to reject both Says-Talk and Deflationism to account for our data. On this kind of approach, expressions like ⌜What X said⌝/⌜X's claim⌝ refer to something that encodes both the proposition X asserted and the context in which it was

---

[7]This idea is proposed by Björnsson [2010: 21], in the course of defending metaethical contextualism.
[8]This interpretation of our data is bolstered by participants' responses to a check question that asked them who the Inspectors were talking about when speaking. In the **Indexical Utterances** condition, participants overwhelmingly thought that Inspector A and Inspector B had themselves in mind (79.49% and 82.05%, respectively). In the **Indexical Assessment**s condition, participants instead overwhelmingly thought that Inspector A and Inspector B had the Chief in mind (84.62% in both cases).

asserted—perhaps the act of X's assertion. We'll model this as a pair of a proposition (the content of the assertion) and a context (the context of the assertion): $<p, \mathbf{c}>$. This approach then adopts an ambiguity view of 'is true'/'is false'. For example, focus on falsity. On one interpretation, 'is false' is deflationary: it maps an assertion to the negation of its content. On that interpretation, ⌜What X said is false⌝ is equivalent to the negation of the proposition X asserted. On the other interpretation, 'is false' means 'false-as-uttered': an assertion $<p, \mathbf{c}>$ is true-/false-as-uttered iff p is true/false at $\mathbf{c}$. Importantly, then, an assertion $<\Diamond p, \mathbf{c}>$ is true-/false-as-uttered iff $\Diamond p$ is true/false relative to evidence $f_{\mathbf{c}}$.

To see how this can predict the data from section 2.3, first consider the **Modal Utterances** case, where A says, 'Fat Tony could have died at the docks' and B says, 'Fat Tony couldn't have died at the docks' (in different contexts). If participants understand 'is false' in (Q) as deflationary, they will agree with it: this is because they should think that either Fat Tony could have died at the docks or Fat Tony couldn't have died at the docks. But if participants understand 'is false' in (Q) as false-as-uttered, they should disagree with this statement, as long as they think that (i) the evidence in A's context is compatible with Fat Tony having died at the docks and (ii) the evidence in B's context is not compatible with Fat Tony having died at the docks. In this case, both A's and B's assertions would be true-as-uttered.

Next, turn to the **Modal Assessments** case. If participants understand Inspector A's utterance of 'What the Chief said is true', Inspector B's utterance of 'What the Chief said is false', and the statement 'At least one of the Inspectors' claims must be false' all in the deflationary way, then they should agree with the latter statement. This is because, on this interpretation, A asserts that Fat Tony could have died at the docks, B asserts the negation of this proposition, and one of these two propositions must be false. However, if instead they understand 'At least one of the Inspectors' claims must be false' as false-as-uttered, they should disagree with this statement, at least if they think that (i) A's evidence is compatible with Fat Tony having died at the docks and (ii) B's evidence is not compatible with Fat Tony having died at the docks. For then both of A's and B's assertions would be true-as-uttered.

One challenge to this strategy is that, if 'true' and 'false' are genuinely ambiguous in the proposed way, it seems that it should be possible to disambiguate them in such a way to make sense of C's response in the following dialogue:

A: Fat Tony might be dead.
B: What A said is false.
C: #I agree with you, B—what A said is false; but also A's claim is true, since A didn't have the evidence proving that Fat Tony is alive.

Yet, to our ears, such dialogues sound irreparably incoherent (cf. Macfarlane [2009: 248]). Compare this dialogue with one involving a polysemous term, such as 'book', which may mean either 'a bound collection of pages' or 'a literary work':

D: (pointing to a bound volume with blank pages) This is a book.
E: I agree with you, D – that is a book; but it also isn't a book, since it's not a literary work.

In contrast with the exchange between B and C, E's response to D here seems coherent. Thus, we think that the relativist strategy of appealing to an ambiguity in 'is true'/'is false' will face some steep challenges.

We conclude that our data raise a challenge to both orthodox contextualist and orthodox relativist theories. We turn next to a different approach to thinking about

modal semantics, one that we think is independently plausible and can account for the data discussed in this section.

4. Contextualist Situation Semantics

In this section, we offer a new theory of epistemic modals, and show how it can predict the complete pattern of judgments that we have found. As a brief preview, our theory will accept context-dependence **Context Dependence**, and truth relativism **Truth Relativism**, and deny Deflationism in favour of a contextualist theory of truth.

We state our theory within a situation semantic framework, on which propositions are sets of situations rather than sets of worlds (compare Kratzer [1989, 2012]).[9] A situation is a possibly partial world—that is, each situation determines a unique world, but there may be distinct situations that determine the same possible world. A proposition p is true at situation $\mathbf{s}$ iff $\mathbf{s} \in p$. Since it is possible for some proposition p to be true at some situations at some world but not others at that same world, adopting a situation semantic framework allows us to predict truth-relativism, **Truth Relativism**.

We capture context-dependence, **Context Dependence**, by allowing that the same epistemic modal sentence may vary in what proposition it expresses at different contexts of utterance. Just as with orthodox contextualist theories, we hold that epistemic modal sentences are about particular modal bases. However, unlike orthodox contextualist theories, we hold some epistemic modal claims are about modal bases that may vary in output across situations; hence, on our theory, a modal base is now a function from a situation (rather than a world) to a set of worlds. This will allow us to define modal bases that are **situation-variant**: such modal bases vary in output across situations that share the same world. We can also define **situation-invariant** modal bases: these vary in output only across situations at different worlds. To see the difference here, consider the following two epistemic modal bases:

> X's evidence at t.
> The best available evidence.

X's evidence at t is **situation-invariant**: it varies across worlds (in so far as X could have had different information than X in fact had) but not across situations at the same world. By contrast, the best available evidence is arguably **situation-variant**, since different evidence is available in different situations.[10]

---

[9]A centred worlds framework would have worked just as well for our purposes (cf. Egan [2007]).

[10]Some independent evidence that epistemic modal claims can be interpreted as about situation-variant modal bases or situation-invariant modal bases comes from data about retraction. Whether you should retract a claim depends on whether it is true in your situation. If you utter ⌜might p⌝ in a context in which it is assigned a situation-variant modal base, and you later learn $\sim p$, then you should retract your utterance in light of this new evidence, since your claim would now be false at your new situation. Suppose instead that you utter ⌜might p⌝ in a context in which it is assigned a situation-invariant modal base (about, say, your current evidence), and your current evidence is compatible with p. Then if you later learn $\sim p$, you would not be required to retract your previous utterance, since it would still be true at your new situation. The observation (from von Fintel and Gillies [2011]) that there are two possible responses (retract, or stick to one's guns) after uttering ⌜might p⌝ and later learning $\sim p$ is thus *prima facie* evidence that we sometimes use epistemic modal sentences to express situation-variant, and sometimes situation-invariant, claims:

A: The keys might be in the drawer.
B: No, I have them with me. Why did you say that?
A1: Oh, I guess that I was wrong.
A2: I didn't say that they were in the drawer, only that they might be, and they might have been!

We continue to accept Says-Talk. Finally, we reject Deflationism in favor of Contextualist Truth/Falsity:

> *Contextualist Truth/Falsity*
> The proposition expressed by ⌜What X said is true⌝ (⌜What X said is false⌝) at **c** is equivalent to the proposition that the referent of ⌜What X said⌝ is true (false) at the situation of **c**.

Contextualist Truth/Falsity predicts that a truth/falsity-assessment of p is a claim that p is true at the situation in which the claim is being made. Suppose that p is the proposition expressed by Z's utterance, and that p is in fact true at s_1 and false at s_2. Suppose that X says, 'What Z said is true' in situation s_1, and that Y says, 'What Z said is false' in situation s_2. Then we predict that both X and Y's claims are true. As we will see in a moment, this will be the key to predicting the data that we found in the **Modal Assessments** case.[11]

4.1 Predicting the Data

Recall the **Modal Utterances** case, where Inspector A looks at the evidence at the docks and utters,

(3) Fat Tony could have died at the docks.

while Inspector B sees Fat Tony plant the evidence and utters,

(4) Fat Tony could not have died at the docks.

We aim to predict the possibility of semantically competent people disagreeing about (Q) in this scenario:

(Q) At least one of Inspector A's and Inspector B's claims must be false.

We predict this by allowing that there are two interpretations of A's and B's claims. On one interpretation, Inspector A made a claim about what was possible, given what he knows from the evidence at the docks (let d be the proposition that Fat Tony died at the docks):

$$\{s: \exists w \in f_{\text{docks}}(s): w \in d\}$$

Similarly, perhaps Inspector B made a claim about what was possible, given what he knows from the security camera footage:

$$\{s: \sim\exists w \in f_{\text{camera}}(s): w \in d\}$$

---

[11]A potential worry is that Contextualist Truth/Falsity predicts that truth-/falsity-ascriptions are either necessarily true or necessarily false. One way to avoid this problem, and still to hold that truth-/falsity-ascriptions convey information about particular situations, is to appeal to counterpart relations among situations. Thus, instead of indexing 'true' and 'false' to a particular situation, we index them to a function from situations to situations, g_s, that is anchored to a particular situation, s. The idea is that g_s maps a situation s' to the counterpart of s that is world-mates with s'. Then the view is that ⌜What X said is true⌝ expresses the proposition (at **c**) that the proposition referred to by ⌜What X said⌝ at **c** is true at the counterpart of s_c—that is, the proposition $\{s: p_X$ is true at $g_{sc}(s)\}$ (where p_X is the proposition that X asserted).

It's not hard to imagine a continuation of the story that (tacitly) supports this interpretation of the inspectors' claims. For example, the Chief might commend both Inspectors on their conclusions—Inspector A because he correctly assessed which possibilities had been left open by what he learned at the docks, and Inspector B because he correctly assessed which possibilities had been left open by what he learned from the security camera. On this intuitive picture, there is no reason why one of the Inspectors' claims must be false, so someone interpreting the case this way should think that (Q) is false.

But this is only one possible way of interpreting the two inspectors' claims. A different interpretation is that both Inspector A and Inspector B made claims about the best available evidence in the case. On this interpretation, Inspector A asserted the following proposition, while Inspector B asserted its negation (where f_{best} is situation-variant):

$$\{s: \exists w \in f_{best}(s): w \in d\}$$

On this interpretation, relative to any situation, at least one of A's and B's claims must be false. Thus, someone interpreting the Inspectors in this way should think that (Q) is true. Variability in whether the modal base is situation-variant or situation-invariant is how our contextualist situation semantics allows for agreement and also disagreement with (Q) in the **Modal Utterances** case, and thus allows us to predict the mid-point agreement observed on average.

Turn next to the **Modal Assessments** case. In this case, the Chief utters (3). Inspectors A and B (with the same evidence as above, and in different contexts from the Chief and each other) overhear the Chief's utterance and say, respectively:

(5) What the Chief said is true.
(6) What the Chief said is false.

In this case, there is a unique proposition expressed by the Chief; what our theory predicts about (Q) depends on what proposition this is. One possibility is that the Chief intended to make a claim just about what he knows at the time. Perhaps he wanted to state something that he could be very confident was true, and to which he could remain committed even after he gained further evidence. On this interpretation, the Chief asserted this content (where f_{Chief} is situation-invariant):

$$\{s: \exists w \in f_{Chief}(s): w \in d\}$$

Then, by Contextualist Truth/Falsity, Inspector A says that this proposition is true at A's situation, and Inspector B says that this proposition is false at B's situation. But, since A's situation and B's situation are at the same world, the Chief's proposition cannot be true at A's situation and false at B's situation. Therefore, on this interpretation, we predict that (Q) is true: someone interpreting the Chief in this way should thus agree with (Q).

However, there is another possibility: perhaps the Chief wanted to make a claim about the case overall, a claim which he may have to revise upon later learning more information. That is, on this interpretation Chief instead expressed situation-variant content about the best available evidence:

$$\{s: \exists w \in f_{best}(s): w \in d\}$$

Then, by Contextualist Truth/Falsity, Inspector A says that this proposition is true at A's situation, and Inspector B says that this proposition is false at B's situation. Given the plausible assumption that in A's situation the best available evidence *is* compatible with *d*, A's claim would be true. And, given the equally plausible assumption that at B's situation the best available evidence is *not* compatible with *d*, B's claim would also be true. Therefore, on this interpretation, we predict that (Q) is false: someone interpreting the Chief in this way should thus disagree with (Q).[12] Taken together, these two interpretive possibilities again allow our theory to predict the midpoint agreement that we observed in the **Modal Assessments** cases.

Finally, we note here why our contextualist situation semantic theory predicts that (Q) is true for the **Indexical Assessments** version of the case. We assume that the semantic value of an indexical like 'I' is not situation-variant. One strategy is that 'I' is directly referential and context-sensitive (à la Kaplan [1989]). Then, when the Chief says, 'I have served on the force for more than twenty years', he expresses the following proposition (let 'Chief' be a name that rigidly designates the Chief):

{**s**: Chief served on the force for more than twenty years at $\mathbf{w_s}$}

When A says, 'What the Chief says is true', he asserts that this proposition is true at his situation and when B says, 'What the Chief says is false', he asserts that the negation of this proposition is true at his situation. But, since A and B's situations share the same world, at least one of their claims must be false. Thus, we predict the high agreement with (Q) in the **Indexical Assessments** case.

5. Conclusion

We argued for evaluating theories of epistemic modals by exploring inter-contextual (in)compatibility judgments. Implementing this strategy, we found a pattern of ordinary judgments that is not predicted by orthodox contextualist or relativist theories. We ended with a novel contextualist situation semantics that *does* predict the pattern of judgments that we found.

Stepping back from the particulars of the contextualism/relativism debate over the semantics of epistemic modals, it bears emphasizing that the phenomenon we have uncovered is likely to extend beyond epistemic modal expressions. A much wider range of expressions are known to exhibit some kind of contextual variability (normative expressions, predicates of taste, gradable adjectives, quantifiers, conditionals, and so on), but don't obviously pattern with so-called 'automatic indexicals', like first person pronouns. A promising possibility is that the present proposal can be fruitfully

---

[12]An anonymous reviewer suggests that, even in this case, it would be natural for A to retract her claim upon accepting B's claim. If so, this is not immediately predicted by our view. Our theory predicts that A says that the best evidence available in A's situation does not rule out the possibility that Fat Tony died at the docks, which is true even at a situation in which the best available evidence rules out that Fat Tony died at the docks. Still, there may be other reasons why it would be appropriate for A to retract her claim in such a case. Claiming that *p* is compatible with the best available evidence in a particular situation may implicate that no nearby situation has better available evidence. If so, then learning that the evidence in some nearby situation rules out *p* would naturally introduce pressure to retract the original claim. We are not convinced that this is the best way to go, but it is worth pointing out that such a case would raise a quite general challenge. There are two intuitions in tension in this kind of case: one is that A's and B's claims are incompatible (in our technical sense) and the other is that it would be appropriate for A to retract her claim in light of accepting B's claim (see Knobe and Yalcin [2014]'s Experiment 4 for a similar case). We hope that further work will be done to explore these intuitions empirically.

extended to this wider class of expressions, which have been of central interest across philosophy, linguistics, and cognitive science. Critically, the methodology that we developed in this paper should be straightforwardly applicable to this wider class of expressions, allowing us to test empirically whether these expressions exhibit the same kind of contextual variation that we have observed in epistemic modals, and to account semantically for this variation where it arises. This extension of the present proposal would not only provide a clearer empirical picture of the context-sensitivity of these expressions, but would shed light on the extent and scope of the underlying phenomenon that gives rise to this kind of pattern in our use of epistemic modals.[13]

References

Björnsson, Gunnar and Stephen Finlay 2010. Metaethical Contextualism Defended, *Ethics* 121/1: 7–36.

DeRose, Keith 1991. Epistemic Possibilities, *The Philosophical Review* 100/4: 581–605.

Dowell, Janice 2011. A Flexible Contextualist Account of Epistemic Modals, *Philosophers' Imprint* 11/14: 1–25.

Egan, Andy 2007. Epistemic Modals, Relativism, and Assertion, *Philosophical Studies* 133/1: 1–22.

Egan, Andy, John Hawthorne, and Brian Weatherson 2005. Epistemic Modals in Context, in *Contextualism in Philosophy: Knowledge, Meaning, and Truth*, ed. Gerhard Preyer and Georg Peter, Oxford: Clarendon Press: 131–170.

Hacking, Ian 1967. Possibility, *The Philosophical Review* 76/2: 143–68.

Kaplan, David 1989. Demonstratives: *An Essay on the Semantics*, Logic, Metaphysics and Epistemology of *Demonstratives* and other Indexicals, in *Themes from Kaplan*, ed. Joseph Almog, John Perry, and Howard Wettstein, New York: Oxford University Press: 481–563.

Katz, Jonah and Joe Salerno 2017. Epistemic Modal Disagreement, *Topoi* 36/1: 141–53.

Knobe, Joshua and Seth Yalcin 2014. Context-Sensitivity of Epistemic Possibility Modals: Experimental Data, *Semantics & Pragmatics* 7/4: 1–21.

Kratzer, Angelika 1981. The Notional Category of Modality, in *Words, Worlds, and Contexts: New Approaches in Word Semantics*, ed. Hans-Jürgen Eikmeyer and Hannes Rieser, Berlin: Walter de Gruyter: 38–74.

Kratzer, Angelika 1989. An Investigation of the Lumps of Thought, *Linguistics and Philosophy* 12/5: 607–53.

Kratzer, Angelika 1991. Modality, in *Handbuch Semantik/Handbook of Semantics*, ed. Arnim von Stechow and Dieter Wunderlich, Berlin: Walter de Gruyter: 639–50.

Kratzer, Angelika 2012. *Modals and Conditionals: New and Revised Perspectives*, Oxford: Oxford University Press.

MacFarlane, John 2005. Making Sense of Relative Truth, *Proceedings of the Aristotelian Society* 105/1: 305–23.

MacFarlane, John 2008. Truth in the Garden of Forking Paths, in *Relative Truth*, ed. Manuel García-Carpintero and Max Kölbel, Oxford: Oxford University Press: 81–102.

MacFarlane, John 2009. Nonindexical Contextualism, *Synthese* 166/2: 231–50.

MacFarlane, John 2011. Epistemic Modals are Assessment-Sensitive, in *Epistemic Modality*, ed. Andy Egan and Brian Weatherson, Oxford: Oxford University Press: 144–78.

MacFarlane, John 2014. *Assessment Sensitivity: Relative Truth and its Applications*, Oxford: Oxford University Press.

Stanley, Jason 2005. Fallibilism and Concessive Knowledge Attributions, *Analysis* 65/286: 126–31.

von Fintel, Kai, and Gillies, Anthony 2011. Might Made Right, in *Epistemic Modality*, ed. Andy Egan and Brian Weatherson, Oxford: Oxford University Press: 108–30.