Eavesdropping: What is it good for?

Jonathan Phillips and Matthew Mandelkern*

Draft of April 2, 2020. Feedback welcome.

Abstract Eavesdropping judgments (judgments about truth, retraction, and consistency across contexts) about epistemic modals have been used in recent years to argue for a radical thesis: that truth is assessment-relative. We argue that judgments for 'I think that p' pattern in strikingly similar ways to judgments for 'Might p' and 'Probably p'. We argue for this by replicating three major experiments involving the latter and adding a condition with the form 'I think that p', showing that subjects respond in the same way to 'thinks' as to modals. This poses a serious challenge to relativist treatments of the modal judgments, since a relativist treatment of the corresponding 'thinks' judgments is totally implausible, so if a unified account of the phenomena is to be found, it cannot be a relativist one. We briefly sketch how a unified account might look.

1 Introduction

Consider the following scenario, from MacFarlane 2011:

> **Might Boston**: You overhear George and Sally talking in the coffee line. Sally says, 'Joe might be in Boston right now.' You think to yourself: Joe can't be in Boston; I just saw him an hour ago here in Berkeley.

Here are some natural questions to ask about what Sally said. First, did she speak falsely? Second, would it be appropriate for her to take back what she said? Third, was your thought (that Joe can't be in Boston) inconsistent with what Sally said? In the recent literature, relativists have argued that the answers to questions like these—questions that involve eavesdropping, or more generally, evaluation across contexts—can be 'yes', and that this reveals something striking about truth: the truth of claims like Sally's depends on the *assessor's* evidence, not the speaker's evidence. Such a position requires a radical readjustment in our thinking about truth: on this way of thinking, the truth of sentences is not fixed by the context in which it is asserted and the world where it is asserted; it also depends on the context where it is

* Thanks to Bob Beddor, Andy Egan, Justin Khoo, Josh Knobe, and Rose Lenehan for very helpful and extensive discussion and feedback.

evaluated.[1] Expressivists have used similar considerations to argue for a different, equally radical conclusion—namely, that sentences like Sally's are neither true nor false at all, but rather serve only to express a certain state of mind.[2]

Contextualists, who hold that the context of assertion, together with the world of evaluation, suffice to fix the truth-value of claims like Sally's, have pushed back against these revisionary claims, aiming to account for the judgments that have motivated relativism and expressivism in the framework of classical contextualist theories.[3]

In this paper we will argue that eavesdropping judgments—the patterns of judgments about truth-value, retraction, and joint consistency that have motivated this debate—in fact do not help us decide between these views at all. Our argument is simple: cross-contextual judgments about constructions of the form ⌜I think that p⌝ pattern in essentially the same ways as judgments about corresponding epistemic modal claims of the form ⌜Might p⌝ or ⌜Probably p⌝. Thus, whatever explanation one gives of eavesdropping judgments about constructions involving ⌜I think that p⌝ will most likely also account for the parallel patterns involving ⌜Might p⌝ or ⌜Probably p⌝. But, for reasons we will explain, it looks unlikely that the explanation of the former will essentially involve relativist, expressivist, or, indeed, contextualist resources; and so it looks unlikely that the judgments in the latter cases will involve these resources either. Insofar as contextualism is the default view, this could be seen as an argument for contextualism. But our central claim is not that contextualism is correct, but rather that these particular judgments do not provide support for relativism, expressivism, or contextualism.

To sketch our argument in a bit more detail, consider the following variant on MacFarlane's case, which substitutes 'I think Joe is in Boston right now' for 'Joe might be in Boston right now':

> **Think Boston**: You overhear George and Sally talking in the coffee line. Sally says, 'I think Joe is in Boston right now.' You think to yourself: Joe can't be in Boston; I just saw him an hour ago here in Berkeley.

Now let's ask the key eavesdropping questions about falsity, retraction, and consistency as regards this variant. Did Sally speak falsely? Would it be appropriate for her to take back what she said? Is what you thought to yourself inconsistent with what Sally said? Our intuition is that it can be reasonable to answer 'yes' to these questions, and, indeed, that this can be reasonable to just the same degree that it is reasonable to answer 'yes' to the parallel questions in the 'might' variant.

1 E.g. Egan et al. 2005, Lasersohn 2009, Stephenson 2007a, Egan 2007, Kölbel 2009, Egan 2011, MacFarlane 2014, Beddor & Egan 2018, Khoo & Phillips 2019.

2 E.g. Yalcin 2007, Swanson 2015, Moss 2015.

3 E.g. von Fintel & Gillies 2008, 2011, Dowell 2011, Khoo 2015.

Suppose we are right about this; what would that show? One thing we might take it to show is that 'I think that Joe is in Boston' is sensitive to the evidence the assessor has in the context of assessment (on a relativist line), or is non-truth valued (an expressivist line), or is sensitive to the salient information in the context of assertion (a contextualist line), in an exactly parallel manner to the way in which 'might' or 'probably' is. But as far as we know, none of these views has been proposed in the literature, and for good reason: none of these is a plausible theory of attitude reports. While there may be features of context (of utterance or assessment) which influence how we interpret attitude ascriptions, it does not seem plausible that the truth of a report about what S believes will generally depend on what the *assessor* of the report believes.

Another option would be to give two different explanations of our two phenomena: one for the pattern of judgments about attitude ascriptions, and another for the pattern of judgments about epistemic modal claims. But, as we will show, these track each other so closely that this option looks ad hoc at best. It would be much more theoretically parsimonious to offer a single explanation for the general pattern. Such an explanation is unlikely to have much to do with epistemic modals specifically, and if that's right, then it isn't going to turn in any interesting way on relativist/contextualist/expressivist features of epistemic modals. A more plausible explanation will instead account for both sets of patterns by way of general considerations concerning the way we think about truth, retraction, disagreement, and consistency.

Our main goal in this paper is simply to argue that the kind of evidence that has played a central role in the literature probably is not helpful in distinguishing between relativism, contextualism, and expressivism. Such a claim obviously does not commit us to any one of these views being correct: it is compatible with our main claims that any one of these is correct. However, insofar as these data have been more critical for motivating relativism and expressivism, and insofar as contextualism is often taken as a the default view, our points here may (but need not) be taken as indirect support for a contextualist position.

Our plan is as follows. In the first three sections, we advance our claim that, with regards judgments about truth/falsity, retraction, and consistency, epistemic modal claims and 'think'-reports pattern in similar ways. To argue for this, we look at the three most significant empirical explorations of these judgments, namely those in Knobe & Yalcin 2014, Beddor & Egan 2018, and Khoo & Phillips 2019. We choose the most critical experiment from each of those papers and describe results of a variant of that experiment which simply replaces the relevant epistemic modal claim with a 'think'-claim. In each case, we find that subjects' judgments pattern the same way in the 'think'-variant as in the original epistemic modal variant. (Relativists adduce similar evidence involving deontic modals, predicates of personal taste,

knowledge, and so on; in this paper we will focus on epistemic modals, largely because those judgments have been studied the most systematically.)

In the final section we do more to explain why we think that these judgments should be explained in a uniform way; why we think that relativist/contextualist/expressivist considerations probably play no part in a unified explanation; and how a unified explanation might go.

2 Knobe & Yalcin 2014: Falsity and Retraction

The first, and simplest, experiment we discuss is from Knobe & Yalcin 2014. That experiment aims to directly test judgments about falsity and retraction in a case like the one with which we started (MacFarlane 2011). We selected the final experiment in Knobe & Yalcin 2014 (Experiment 4) because this experiment tested judgments of both retraction and falsity within a single experiment and involved a case which has been prominent in the literature. Our only change to the original study was to include new conditions in which the relevant epistemic modal claim is replaced with an attitude report.

The original study found that participants were reluctant to judge an epistemic modal claim as false in MacFarlane-inspired cases, though they did agree that the claim should be retracted. We seek to both directly replicate this finding and investigate whether it extends to otherwise similar attitude reports.

2.1 Methods

We collected a sample of 242 participants ($M_{age} = 37.09$; $SD_{age} = 11.3$; 104 females) from Amazon Mechanical Turk (www.mturk.com). Participants were randomly assigned to one of four conditions, two of which were exact reproductions of the conditions in Experiment 4 in Knobe & Yalcin 2014. As in Knobe & Yalcin 2014, all participants read a version of the following scenario, which builds on MacFarlane's original case. The different versions varied only in what exactly Sally said. For example, in the non-modal variant, Sally makes a non-modal claim about Joe's location:

> **Non-modal**: Sally and George are talking about whether Joe is in Boston. Sally carefully considers all the information she has available and concludes that there is no way to know for sure. Sally says: "Joe is in Boston."
>
> Just then, George gets an email from Joe. The email says that Joe is in Berkeley. So George says: "No, he isn't in Boston. He is in Berkeley."

In the **Modal** variant, Sally instead says "Joe might be in Boston".

Our addition was to add an **Attitude** variant in which Sally says "I think Joe is in Boston." No other changes were made to the original materials.

After reading one of these three variants, participants either answered a question about whether it would be appropriate for Sally to retract what she said or answered a question about whether what Sally said was false:

> **Retraction question:** We want to know whether it would be appropriate for Sally to take back what she said (for example, by saying "Ok, scratch that"). So please tell us whether you agree or disagree with the following statement:
>
> - It would be appropriate for Sally to take back what she said.
>
> **Falsity question:** We want to know whether what Sally said is false. So please tell us whether you agree or disagree with the following statement:
>
> - What Sally said was false.

In both cases, participants indicated their agreement on a scale from 1 ('Completely disagree') to 7 ('Completely agree'). Finally, all participants completed a brief demographic questionnaire.

2.2 Results

Data and code for all of our experiments are available at [redacted for anonymous review].

2.2.1 Replication

To statistically characterize the pattern of responses (see Fig.1), we first asked whether we replicated the original finding in Knobe & Yalcin 2014.[4] Similar to Knobe & Yalcin (2014), we observed the critical interaction effect between question-type (Retraction vs. Falsity) and statement-type (Non-modal vs. Modal), $F(1,156) =$

4 An overall analysis of variance revealed that participants' agreement ratings were significantly affected by whether the agent uttered a non-modal assertion, an epistemic modal claim, or an attitude report, $F = 68.2$, $p < 0.001$, $\eta_p^2 = 0.17$. We also observed a significant effect of whether participants were asked about whether it would be appropriate for the agent to retract her claim or whether the agent's claim was false, $F = 24.12$, $p < 0.001$, $\eta_p^2 = 0.224$. More importantly, we also observed an interaction between these two variables, $F = 14.87$, $p < 0.001$, $\eta_p^2 = 0.112$, meaning that the pattern of participants' judgments about the different claims differed depending on whether they were asked the retraction or falsity question.

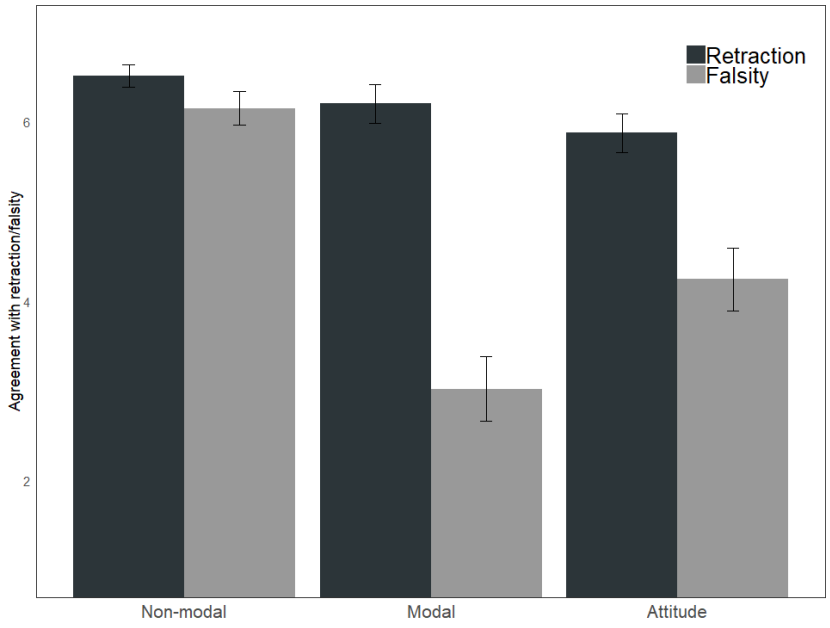

Figure 1 Graph of participants' agreement ratings that retraction would be appropriate (dark bars) or that the claim was false (light bars) as a function of whether the utterance involved a bare assertion, an epistemic modal claim, or an attitude report.

34.64, $p < .001$, $\eta_p^2 = 0.182$. This interaction was driven by the fact that we observed no significant difference between judgments of falsity and retraction for non-modal claims, $t(66.38) = -1.62$, $p = 0.109$, $d = 0.363$, but did observe a large difference between falsity and retraction judgments for epistemic modal claims, $t(64.99) = -7.65$, $p < 0.001$, $d = 1.689$.

2.2.2 Extension

We next asked whether the observed pattern for epistemic modal claims extended to attitude reports. Indeed, we found a similar pattern: when the agent's claim involved an attitude report rather than an epistemic modal claim, participants were more likely to judge that the agent should retract the claim than that the claim was false $t(66.91) = -4.00$, $p < 0.001$, $d = 0.883$. Moreover, we found no significant difference in participants' agreement that retraction would be appropriate when the the claim involved an epistemic modal or an attitude report, $t(79) = 1.06$, $p = 0.291$, $d = 0.236$, and found that, if anything, participants more agreed with the falsity of the claim when it involved an attitude report than when it involved an epistemic modal, $t(80) = -2.45$, $p = 0.016$, $d = 0.541$. We return to this notable finding in the general discussion.

In short, Knobe & Yalcin (2014) found that when the prejacent of an epistemic modal claim turns out to be false, participants judge that the modal claim should be retracted, but that participants are reluctant to say that it is clearly false (giving only midpoint-level agreement). We found a similar pattern of judgments when the epistemic modal claims were replaced with attitude reports.

3 Beddor & Egan 2018: Falsity and QUDs

The second experimental paradigm we explored probes judgments about truth and falsity. In particular, Beddor & Egan (2018) argue that judgments about the truth-value of epistemic modal claims depend on the question under discussion (*QUD*) in the context of assessment. They take this as evidence for a particular kind of relativism: one on which truth is relative to the assessors' body of evidence and their QUD.

Beddor and Egan argue for their claim on the basis of several structurally similar experiments. We will focus on one particular experiment, Experiment 5, which we think controls for some independent confounds in a helpful way which Beddor and Egan discuss. (We come back to one these potential confounds, namely prejacent targeting, in §5.) We replicate their experiment and add two new conditions in which the relevant epistemic modal claims are replaced with attitude reports.

3.1 Methods

We collected a sample of 500 participants ($M_{age} = 38.33$; $SD_{age} = 13.21$; 240 females) from Amazon Mechanical Turk (www.mturk.com). Participants were randomly assigned to one of four conditions, two of which directly replicated the conditions of Experiment 5 in (Beddor & Egan 2018).

In Beddor and Egan's original set-up, participants first read the following:

> John is worried he might have strep throat. He goes to his primary care physician and she runs an initial test that indicates that there is a 75% chance that John does not have strep. Based on the initial test results, John's doctor says: "You probably don't have strep throat. However, we should do a throat culture in order to be safe. If it turns out that you have strep throat, we should put you on antibiotics."

The case went on as follows in two conditions that varied the QUD. In the QUD-PREJACENT condition, the continuation focused on whether the prejacent is true—i.e., on whether John has strep throat:

> QUD-PREJACENT CONDITION:
>
> John comes back two days later to find out the results of the throat culture, and sees a different doctor. The throat culture comes up positive, which indicates there is a 90% chance that John has strep throat. John has not yet seen the results of these tests, but his new doctor has. John asks the new doctor: "I'm trying to figure out whether I need to take antibiotics. My primary care physician told me, 'You probably don't have strep.' Is what she said true?"
>
> Which of the following responses would be correct?
>
> (a) "No, it's not"
>
> (b) "Yes, it is"

In the QUD-COMPETENCE condition, the continuation focused instead on the competence of John's primary care physician:

> QUD-COMPETENCE CONDITION:
>
> John comes back two days later to find out the results of the throat culture, and sees a different doctor. The throat culture comes up positive, which indicates there is a 90% chance that John has strep throat. But now John wants to know whether his primary care physician made a mistake administering the initial test, so he asks: "I'm trying to figure out whether I can rely on my primary care physician. She told me, 'You probably don't have strep'. Is what she said true?"

> The new doctor reviews the initial tests, and confirms that John's primary care physician had not made any mistakes interpreting the results. Given this, which of the following responses would be correct?
>
> (a) "No, it's not"
>
> (b) "Yes, it is"

In our variants, we simply replaced the doctor's utterance, 'You probably don't have strep throat' with the utterance 'I don't think you have strep throat'. In other words, participants in these two conditions first read the following preamble:

> John is worried he might have strep throat. He goes to his primary care physician and she runs an initial test that indicates that there is a 75% chance that John does not have strep. Based on the initial test results, John's doctor says: "I don't think you have strep throat. However, we should do a throat culture in order to be safe. If it turns out that you have strep throat, we should put you on antibiotics."

Subjects then were randomly assigned to one of the two following conditions. In the THINK / QUD-PREJACENT condition, the continuation again focused on whether the sentence uttered is true:

> THINK / QUD-PREJACENT CONDITION:
>
> John comes back two days later to find out the results of the throat culture, and sees a different doctor. The throat culture comes up positive, which indicates there is a 90% chance that John has strep throat. John has not yet seen the results of these tests, but his new doctor has. John asks the new doctor: "I'm trying to figure out whether I need to take antibiotics. My primary care physician told me, 'I don't think you have strep.' Is what she said true?"
>
> Which of the following responses would be correct?
>
> (a) "No, it's not"
>
> (b) "Yes, it is"

By contrast, in the THINK / QUD-COMPETENCE condition, the continuation focused on the competence of John's primary care physician:

> THINK / QUD-COMPETENCE CONDITION:
>
> John comes back two days later to find out the results of the throat culture, and sees a different doctor. The throat culture comes up positive, which indicates there is a 90% chance that John has strep throat. But now John

> wants to know whether his primary care physician made a mistake administering the initial test, so he asks: "I'm trying to figure out whether I can rely on my primary care physician. She told me, 'I don't think you have strep'. Is what she said true?"
>
> The new doctor reviews the initial tests, and confirms that John's primary care physician had not made any mistakes interpreting the results. Given this, which of the following responses would be correct?
>
> (a) "No, it's not"
>
> (b) "Yes, it is"

As in Beddor and Egan (2018), participants also completed a comprehension check question. For the first two conditions, we used the same check as in Beddor and Egan, namely: 'In the scenario you just read, which of the following did John's primary care physician say?' The options were '(a) You probably have strep. (b) You probably don't have strep. (c) You probably have pneumonia. (d) You probably don't have pneumonia'. In the other two conditions, we changed the comprehension check question such that the options were: '(a) I think you have strep. (b) I don't think you have strep. (c) I think you have pneumonia. (d) I don't think you have pneumonia.' Lastly, all participants were asked to complete a brief demographic questionnaire. No other changes were made to the original materials.

3.2 Results and analysis

We excluded the 27 participants who failed to answer the comprehension check question correctly, and analyzed the responses from the remaining 473 participants.

3.2.1 Replication

Replicating the general pattern in Beddor & Egan 2018, we found that 83.05% of participants in the QUD-PREJACENT condition selected 'No, it's not', while only 28.81% of the participants in the QUD-COMPETENCE condition did.

3.2.2 Extension

More importantly, we found a strikingly similar pattern of responses when the utterance was instead about the what the doctor believed. We found that 74.79% of participants in the THINK / QUD-PREJACENT condition selected 'No, it's not', while only 27.12% of the participants in the THINK / QUD-COMPETENCE condition did, a strikingly similar pattern to Beddor and Egan's finding (Figure 2).

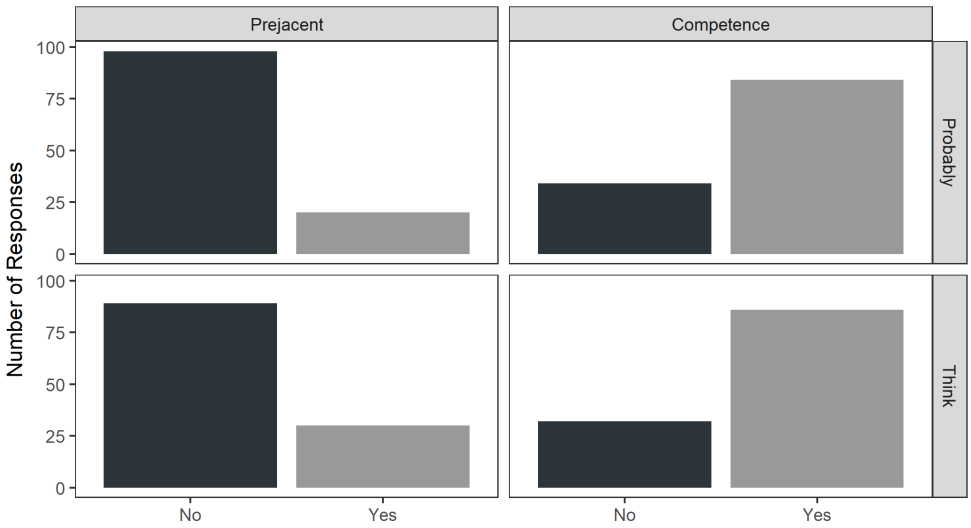

Figure 2 Graph of the number of times each response option was selected as a function of both whether or not the utterance involved an epistemic modal and whether or not the assessment QUD targeted the doctor's competence.

Moreover, analyzing all of the data together using a generalized linear model, we found only a main effect of the assessment QUD manipulation, $z = 7.826$, $p < 0.001$, and no effect of whether the utterance involved an epistemic modal or an interaction between them, $p \geq 0.121$.

In sum, Beddor and Egan (2018) found that the degree to which subjects agreed with the 'Yes' versus 'No' responses varied with the QUD condition. In particular, subjects were much more likely to judge 'Yes' in the COMPETENCE condition (61%) than in the PREJACENT condition (27%). Beddor and Egan argued that this tells in favor of a particular kind of QUD-sensitive relativism for epistemic modal claims (this experiment focuses on 'probably' claims, while other experiments focus instead on 'might' claims). We found the same pattern of QUD-relativity in judgments for attitude reports.

4 Khoo & Phillips 2019: Consistency

The final paradigm we explore concerns judgments about consistency. Khoo & Phillips (2019) investigate the degree to which subjects judge that at least one of apparently conflicting epistemic modal claims, or judgments about epistemic modal claims, must be false. Relativist and contextualist approaches differ at a structural

level on their predictions for these questions. While the pattern found by Khoo and Phillips is difficult to capture on either a standard relativist or standard contextualist approach, we will not focus on this aspect of the data. Instead, our aim will be simply to show that consistency judgments of apparently conflicting 'think'-claims (or assessments of a single 'think'-claim) exhibit a pattern similar to what was found by Khoo & Phillips (2019). We do this by replicating the experiment in (Khoo & Phillips 2019) while adding analogous conditions involving 'think'-claims.

4.1 Methods

In this experiment, 405 participants (M_{age} =37.99, SD_{age} =12.47; 192 females) were recruited through Amazon Mechanical Turk (http://www.mturk.com). Participants each completed a single trial, which involved reading a vignette about an ongoing police investigation. In all cases, participants first read the following background information:

> The police are on the trail of Fat Tony, a local mobster. This morning, they learn of a rumor that Fat Tony has died at the docks.
>
> The Chief of the Police assigns Inspector A to examine the evidence at the docks. Meanwhile, the District Attorney assigns Inspector B to review the footage from the security camera at the docks.

How this background continued depended on the condition to which participants were randomly assigned. Participants were assigned to either an **Utterances** or an **Assessments** condition, and additionally to make assessments of either epistemic **Modal** claim, a **Non-Modal** claim, or an **Indexical** claim.

In the **Modal Utterances** case, the background continued as follows:

> Inspector A takes a good look at the evidence down by the docks, and concludes that it suggests, but does not prove, that Fat Tony died at the docks. The Chief calls Inspector A at the docks and asks him, "What have you found?"
>
> **Inspector A replies, "Fat Tony could have died at the docks."**
>
> Meanwhile, Inspector B reviews the security camera footage and concludes that the footage proves that Fat Tony did die at the docks. The District Attorney calls Inspector B and asks him, "What have you found?"
>
> **Inspector B replies, "Fat Tony couldn't have died at the docks."**

By contrast, in the **Modal Assessments** case, the background instead continued:

> Inspector A takes a good look at the evidence down by the docks, and concludes that it suggests, but does not prove, that Fat Tony died at the docks. Afterwards, he goes home. That evening, Inspector A and his wife watch the Chief of Police talking with reporters on TV. The reporters on the news ask the Chief what his investigation had found.
>
> **The Chief tells the reporters: "Fat Tony could have died at the docks."**
>
> Inspector A's wife knows that Inspector A was examining the evidence at the docks and so she asks him, "Is that right?"
>
> **Inspector A replies, "What the Chief said is true."**
>
> Meanwhile, Inspector B reviews the security camera footage and concludes that the footage proves that Fat Tony did not die at the docks. That evening he watches the same TV broadcast with his wife, and they also hear the Chief tell the reporters, "Fat Tony could have died at the docks."
>
> Inspector B's wife knows that Inspector B was examining the evidence at the docks and so she asks him, "Is that right?"
>
> **Inspector B replies, "What the Chief said is false."**

Two other conditions differed slightly from these. In one, participants instead evaluated a **Non-Modal** claim. These cases were identical to the preceding ones except that the Inspectors'/Chief's claim(s) did not include the epistemic modal, and thus instead read: "Fat Tony [died / did not die] at the docks."

So far, this set-up exactly matches Khoo and Phillips' set-up. Our addition was an **Attitude** condition. These cases were identical to the preceding ones except that the Inspectors'/Chief's claim(s) took the form of an attitude report, and thus instead read: "I [think / don't think] Fat Tony died at the docks."

Finally, other participants instead assessed **Indexical** statements, which were also included in Khoo and Phillips' experiment. In the **Indexical Utterances** case, the background continued as follows:

> Inspector A takes a good look at the evidence down by the docks, and concludes that it suggests, but does not prove, that Fat Tony died at the docks. Later that evening, Inspector A gets a call from the Chief. The Chief knows that certificates of appreciation are being given to officers who have served on the police force for at least twenty years, so he asks Inspector A, "How long have you served on the police force?"
>
> **Inspector A replies, "I have served on the police force for twenty years."**
>
> Meanwhile, Inspector B reviews the security camera footage and concludes that the footage proves that Fat Tony did die at the docks. Later that evening, Inspector B gets a call from the District Attorney. The District Attorney also knows that certificates of appreciation are being given to officers who have

served on the police force for at least twenty years, so he asks Inspector B, "How long have you served on the police force?"

Inspector B replies, "I have not served on the police force for twenty years."

In the **Indexical Assessments** condition, the two inspectors instead made two different claims about the truth of the Chief's utterance:

> Inspector A takes a good look at the evidence down by the docks, and concludes that it suggests, but does not prove, that Fat Tony died at the docks. Afterwards, he goes home. That evening, Inspector A and his wife watch the Chief of Police talking with reporters on TV. The reporter on the news knows that certificates of appreciation are being given to officers who have served on the police force for at least twenty years, so she asks the Chief, "How long have you served on the police force?"
>
> **The Chief tells the reporters: "I have served on the police force for twenty years."**
>
> Inspector A's wife knows that Inspector A is on the police force, and so she asks him, "Is that right?"
>
> **Inspector A replies, "What the Chief said is true."**
>
> Meanwhile, Inspector B reviews the security camera footage and concludes that the footage proves that Fat Tony did not die at the docks. That evening he watches the same TV broadcast with his wife, and they also hear the Chief say to the reporter, "I have served on the police force for twenty years."
>
> Inspector B's wife knows that Inspector B was also on the police force, and so she asks him, "Is that right?"
>
> **Inspector B replies, "What the Chief said is false."**

After reading the entire vignette, participants were reminded that the inspectors had made two different claims and were asked whether they agreed or disagreed that "At least one of the inspectors' claims must be false." Participants rated their agreement on a scale from 1 ("Completely Disagree") to 7 ("Completely Agree").

After answering this question, participants also answered a manipulation check question. In the **Modal, Non-Modal,** and **Attitude** conditions, participants were asked to make a judgment about what was more relevant in Inspector A's conversation and, then separately, in Inspector B's conversation, which allowed us to test whether they tracked the differences across these two conversational contexts. In both cases, participants responded by selecting which of the following two options was more relevant in each conversation:

- What the evidence at the docks reveals about Fat Tony.
- What the security camera footage reveals about Fat Tony.

In the **Indexical** conditions, participants were instead separately asked who both Inspector A and Inspector B think has served on the police force for twenty years. They responded by selecting one of the following three options for each Inspector:

- Inspector A
- Inspector B
- The Chief

Finally, participants completed a brief and optional demographic questionnaire.

4.2 Results and analysis

No participants were excluded from the analyses. To ensure that participants correctly understood the relevant differences in Inspector A's and Inspector B's contexts, we first assessed participants' judgments of which evidence was most relevant in the two contexts. These judgments of relevance confirmed that participants clearly tracked the changes in the different contexts: participants found the evidence at the docks to be more relevant in Inspector A's context, and found the evidence from the security camera to be more relevant in Inspector B's context, $\chi^2(1) = 153.4$, $p < .001$, $V = 0.5$.

4.2.1 Replication

Following Khoo & Phillips (2019), we first analyzed participants' judgments of whether one of the Inspectors' claims must be false in the Indexical condition. In the **Indexical Utterances** condition, where Inspector A says, "I have served on the force for more than twenty years," and Inspector B says "I have not served on the force for more than twenty years," participants strongly disagreed that at least one of the Inspectors claims must be false (M =3.42, SD =1.94). However, in the **Indexical Assessments** condition, where the two Inspectors made conflicting assessments about the Chief's utterance of "I have served on the police force for twenty years," participants instead strongly agreed that at least one of the Inspectors claims must be false (M =5.65, SD =1.68, $t(94)$ =-6.02, p <.001, d =1.23).

Next, we analyzed participants' compatibility judgments in the **Modal** and **Non-Modal** conditions with a 2 (Statement: Bare vs. Modal) $\times$ 2 (Condition: Utterances vs. Assessments) ANOVA. Replicating Khoo & Phillips (2019), partic-

ipants' judgments were significantly affected by whether or not the claims involved a bare assertion or an epistemic modal claim, $F(1,199) = 17.95$, $p < .001$, $\eta_p^2 = 0.083$. More specifically, we found that participants more strongly agreed that one of the inspectors' claims must be false when they uttered/assessed a bare assertion ($M = 5.69, SD = 1.46$), than when they uttered/assessed a modal claim ($M = 4.62, SD = 2.06$), $t(184.22) = -4.27$, $p < .001$, $d = 0.6$. As in Khoo & Phillips (2019), we also did not observe a significant effect of whether the Inspectors made conflicting utterances or conflicting assessments, $F(1,199) = 0.295$, $p = .588$, $\eta_p^2 = .001$, and did not find an interaction effect between these two variables, $F(1,202) = 0.091$, $p = .764$, $\eta_p^2 < 0.001$, meaning that the difference between the different claims (**Bare** vs. **Modal**) did not significantly differ between the **Assessments** and **Utterances** conditions.

4.2.2 Extension

Finally, we asked whether the pattern we observed in the modal condition could similarly be found in the attitude report condition. Specifically, we did a similar analysis to that in Khoo & Phillips (2019), but replaced the modal condition with the attitude condition. Once again, we found that participants' judgments were significantly affected by whether or not the claims involved a bare assertion or an attitude report, $F(1,202) = 4.772$, $p = .030$, $\eta_p^2 = 0.023$. Specifically, we found that participants more strongly agreed that one of the inspectors' claims must be false when they uttered/assessed a bare assertion ($M = 5.69, SD = 1.46$), than when they uttered/assessed an attitude report ($M = 5.16, SD = 1.95$, $t(194.12) = -2.21$, $p = .028$, $d = 0.31$. We again did not observe a significant effect of whether the Inspectors made conflicting utterances or conflicting assessments, $F(1,202) = 0.052$, $p = .820$, $\eta_p^2 < .001$, and did not find an interaction effect between these two variables, $F(2,202) = 0.213$, $p = .644$, $\eta_p^2 = 0.001$, meaning that the difference between the different claims (**Bare** vs. **Attitude**) did not significantly differ between the **Assessments** and **Utterances** conditions.

Finally, the **Modal** and **Attitude** conditions did not differ significantly from one another, $t(205.63) = -1.94$, $p = .054$, $d = 0.27$.

In short, we replicated Khoo and Phillips' key finding that speakers are less likely to judge that at least one of the claims/judgments must be false in the **Modal** cases than in the **Bare** cases. And, critically, we found that in the **Attitude** cases, judgments patterned much the same way as in the **Modal** cases.

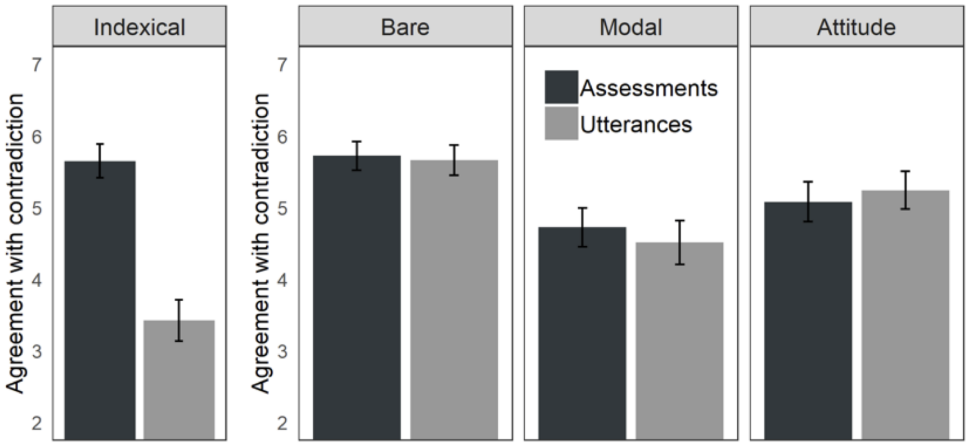

Figure 3 Participants' mean level of agreement that at least one of the inspectors' claims must be false. Errors bars indicate +/- 1 SEM.

5 Discussion

5.1 What is eavesdropping good for?

These results together support our hypothesis that epistemic modal claims and 'think' claims pattern together when it comes to the eavesdropping phenomena that have been at the heart of the debate between relativism, contextualism, and expressivism. We think that this, in turn, suggests that these phenomena cannot play a central role in deciding between these views. We lay out our argument for this conclusion in this subsection.

One possible response to these results is to hold that there are simply two different phenomena here: judgments about truth-value, retraction, and consistency for *epistemic modal* claims, and judgments about truth-value, retraction, and consistency for 'think' claims. If so, then we ought to pursue independent explanations for set of phenomena. We can't rule out a possibility like this, but from the point of view of theoretical parsimony, it is obviously unattractive. Given the strikingly parallel patterns we find across these domains, it strikes us as very unlikely that there are two completely unrelated explanations that simply happen to generate very similar patterns of results across this wide range of cases and questions.

So let us instead proceed under the assumption that the explanation of the judgments in the two domains will be closely related. Given this assumption, what kind of explanation would be plausible? Well, let's think about what a relativist might say here. Relativism about attitude ascriptions has in fact been defended with respect

to 'knows' in MacFarlane 2014: it's not unreasonable to think that sensitivity to skeptical scenario-raising varies with the context of assessment rather than assertion. And plausibly other kinds of context sensitivity that have been ascribed to attitude claims—for instance sensitivity to Fregean guises, or to questions under discussion—could be held to be provided by the context of assessment rather than the context of assertion.

But none of this, as far as we can see, is any help in explaining the present phenomena. Consider for example Knobe & Yalcin (2014)'s results and suppose we want a relativist explanation of why subjects think that Sally should retract her claim that 'I think Joe is in Boston', or why subjects are somewhat inclined to agree that the claim is false. A relativist explanation which parallels the kind of explanation given concerning epistemic modals would have to say that what matters in assessing the truth of 'I think Joe is in Boston', as asserted by Sally, is not Sally's attitude state in the context of assertion, but rather the evaluating subject's attitude state in the context of assessment: the fact that we know in the later context that Joe is in Boston would somehow have to suffice to make Sally's earlier claim 'I think Joe is in Boston' false, as assessed at the context of evaluation. But this idea has not been defended in the literature, and it strikes us as totally implausible.

Let us say a little bit more about why this is. Let c be the context in which Sally asserts 'I think Joe is in Boston' (for simplicity, think of c as a centered world). Suppose that, in c, Sally is very confident that Joe is in Boston; she has all the functional dispositions that are usually associated with a state of belief, including, of course, the disposition to sincerely assert that she believes Joe is in Boston. Now consider *our* context of assessment, where we know that Joe was not in Boston. Can *our* knowledge somehow retroactively remove Sally's belief that Joe was in Boston? It is hard to see how it could.

To take another example, think about Beddor and Egan's case, in which John's doctor says to John: 'I don't think that you have strep throat' in his context s. If we want to explain the QUD-relativity reported above by way of a relativist story, we will have to say that the truth of the doctor's claim in c depends on what information *we* have in our context and what question we are attending to; and in particular, that when *we* find out that John probably does have strep throat, this can on its own suffice to falsify the doctor's claim that *he* didn't think John had strep throat. Again, this seems plainly implausible. *We* can't change facts about what the doctor thought at some point in the past simply by finding out that it was false.

Our point here is not that relativism is *in general* implausible, nor even that relativism about 'think' is implausible. Our point is rather that the particular form of relativism that would be required to treat 'think' on an analogy with 'might' or 'probably' is grossly implausible. By contrast, the forms of relativism that have been put forward about epistemic modals, predicates of taste, and so on seem *prima facie*

much more plausible. To say that the truth of ⌜Might p⌝ or ⌜That's tasty⌝ depends on *our own* information, or standards of assessment, seems like a possibility worth exploration, in a way that the present hypothesis about 'think' just doesn't: the idea that the truth of ⌜I think that p⌝, as asserted by *S* in some particular context *c*, depends in general on whether *we* think that p is true strikes us as a hypothesis that cannot be taken seriously. There may, again, be more subtle forms of relativism about attitude claims that are more plausible—for instance, perhaps a relativist might hold that Fregean guises can be provided by the context of assessment, or that (as MacFarlane holds) what possibilities are made salient in assessing knowledge ascriptions depends on the context of assessment. But the radical form of relativism about 'think' that would be required to account for our data just does not seem plausible and we think that it is for good reason that relativism along those lines has never been defended.

The upshot of this is that if we want a unified explanation of cross-contextual judgments which covers both epistemic modals and 'think' claims, it will not use relativist resources to make sense of these judgments.

Similar considerations extend to expressivist approaches. Perhaps there are subtle forms of expressivism that can be defended about 'I think'. But what would be needed to give a unified expressivist account of cross-contextual judgments about 'think' and 'might' or 'probably' doesn't seem plausible. Such an account would have to argue that there is no fact of the matter about whether Sally spoke truly when she said 'I think Joe is in Boston'. Rather, the idea would go, we should set aside questions about truth and falsity, and should simply accept what she said if *we* think Joe is in Boston, and reject it if *we* think he's not. But this seems absurd. There *is* a fact of the matter about whether what Sally said was true, and this fact depends on whether Sally thought that Joe was in Boston. If we want to know whether Sally spoke truly, we should inquire about *her* mental states, not *ours*.

What about contextualism? Well, contextualist approaches to epistemic modals have generally tried to explain cross-contextual judgments by way of independently motivated, general considerations about how people think about disagreement, retraction, and so on (as in e.g. Khoo 2015). So contextualism is not usually proposed as itself being an explanation of these phenomena. So while contextualism about 'think', as about epistemic modals, is reasonable enough, in our view, contextualism does not advertise itself as on its own accounting for the phenomena in question; and thus it is doubtful that contextualist resources on their own can account for the relevant phenomena concerning 'think' and 'might'.

In short, then, it looks like, if we want a unified account of cross-contextual judgments involving 'think' and epistemic modals, that account will not essentially involve the resources of relativism, expressivism, or contextualism, but rather will come from independent considerations about cross-contextual judgments more gen-

erally. To be clear, these general considerations are likely to be consistent with an underlying theory of epistemic modals which is relativist, expressivist, or contextualist; our claim is that none of these theories will on its own explain cross-contextual judgments about the relevant phenomenon.

The dialectical situation here is somewhat delicate, so let us rehearse it briefly. Our argument has been that cross-contextual judgments do not motivate relativism, expressivism, or contextualism. While this claim, on its own, is neutral between these theories, the dialectical situation concerning them differs in an important way: namely, relativism and expressivism have been *largely motivated* by considerations about cross-contextual judgments. If we are right that this motivation is undermined by the data we have presented, then there will be correspondingly less motivation for these theories in a general sense. Moreover, insofar as contextualism is the default view that was complicated primarily by this kind of cross-contextual judgment, the data we have presented could be taken as an argument for contextualism. However, we don't want to commit to this point: there could be independent sources of motivation for relativism or expressivism. Our point is the more limited dialectical one that cross-contextual judgments don't on their own provide any evidence for or against relativism, expressivism, or contextualism.[5]

5.2 A positive account

This is our central point, and it is a negative one. Still, a natural question to raise at this point is how one *should* give a unified account of the cross-contextual judgments concerning epistemic modals and attitude reports. Though our main goal is not to answer this question, we want to say enough to suggest that an answer can be given. This is important dialectically because if it turned out to be impossible to find a satisfying unified account of these judgments, then a disjunctive strategy which deals with modal judgments in one way and attitude judgments in a different way might start to look more plausible.

What we want to say here is inspired both by earlier work on modal disagreement in Khoo 2015, as well as the particular relativist proposal given in Beddor & Egan 2018. Although we have argued that Beddor and Egan's QUD-relativism cannot be plausible extended to 'think', and so is unlikely to be the correct account of modal judgments either, the observation that cross-contextual judgments vary with the QUD of assessment naturally suggests a different route, one rooted in the Gricean

5 Both relativism and expressivism have also been motivated on the basis of embedding data (for the former, see Stephenson 2007b,a, Lasersohn 2009; for the latter, see Yalcin 2007). But contextualist theories have been developed which account for most or all of these data (see in particular Ninan 2016, Mandelkern 2019), so it is not clear that there remains an embedding-based case for either relativism or expressivism as compared with contextualism.

perspective on conversation as a cooperative enterprise (Grice 1989) together with a Stalnakerian perspective on conversation as a series of proposed updates to a common ground (Stalnaker 1978).

To see our idea, start by recalling Beddor and Egan's scenario. Think about what statements like 'You probably don't have strep throat' and 'I don't think that you have strep throat', as asserted by the doctor to John, *do*. Intuitively, these are proposals to put little credence in the proposition that John has strep throat. If either one is accepted, then the interlocutors will no longer spend much energy thinking about possibilities in which John has strep throat. These intuitions can be made precise in a variety of different ways. A helpful tool here is a Stalnakerian model of conversation as a series of updates to a *common ground*: a record of the common commitments of the conversants. Given this picture, an assertion of ⌜I don't think p⌝ will be very naturally taken as a proposal to rule out *p*,[6] or at least to lend it little credence, since the common ground aims to coordinate on the speakers' beliefs. Likewise, there are a variety of accounts of epistemic modals, in both contextualist (Stalnaker 2014, Mandelkern 2018) and expressivist (Yalcin 2007, 2012, Swanson 2015) frameworks, on which an assertion of ⌜Probably p⌝ will have the effect of ensuring that the conversants jointly assign little credence to *p* (plausibly relativist frameworks could be developed in that direction as well).

Second, consider the proposal from Khoo 2015 about modal disagreements. Khoo's key idea is that disagreement can target the *proposal* made in the original context, rather than the truth-value of what was said. So, for instance, even if it is strictly true that the doctor *thought* that John had strep throat, her claim was not just an articulation of this fact but also a proposal to make a certain kind of update to the common ground—and one can *disagree* with the proposed update to the common ground, even without thinking that the literal content communicated is false (as a simple example of this kind of disagreement, Khoo considers the exchange 'The bank is open today.' 'No, the bank *might* be open today.') Our proposal is to extend Khoo's idea to retrospective assessments in general, not just about agreement but also about truth-value, retraction, and consistency. The idea is that subjects asked for their intuitions about any of these can naturally focus on evaluating the proposal which was made by the speech act in question, rather than on the asserted content itself. In other words, subjects may interpret questions which are explicitly about disagreement, retraction, truth, or consistency as being about the *proposal* in question rather than about the asserted *semantic content*. The point is not that subjects are completely unaware of the distinction between, say, considering a sentence to be false and disagreeing with the update it proposes. Indeed, awareness (at some level) of just that kind of distinction is presumably what accounts for Khoo's finding that

6 At least on its most prominent, neg-raised reading.

subjects in some conditions are more inclined to disagree with a sentence than to judge it false, and likewise with Knobe and Yalcin's finding, replicated above, that subjects are more willing to retract a modal claim (and, likewise, an attitude claim) than to judge it false. Nonetheless, we propose that subjects tend to move somewhat freely between, say, a question about whether a claim was true and whether the proposal it made was a good one.[7] This is especially so in cases where, say, the subject is asked explicitly about truth, but the questioner makes clear that what they *care* about is a more general question, say, about the epistemic reliability of the person who made the claim.

To make this more concrete, think about the cross-contextual evaluations that Beddor and Egan report about 'probably', and the corresponding ones that we report about 'think'. These findings show that, when the salient question is whether John has strep throat, subjects are much more likely to report that the doctor's assertions of 'You probably don't have strep throat' and 'I don't think that you have strep throat' were false. When the salient question is instead whether the doctor is a good doctor, they are much more likely to report these to be true. We think this is naturally explained by way the idea that these responses can target the *proposals* that these make, rather than their truth-conditions. These are, again, plausibly both proposals to lend little credence to the proposition that John has strep. Well, were these proposed updates good ones? It depends what features of the situation you care about. Were they good with respect to moving us towards the facts of the matter about John's case? Clearly not. Were they good with respect to evidencing iatric virtues? Clearly so. There are different respects in which a proposed update can be a good one, and shifting focus from one of those respects to another can naturally shift our retrospective judgments.

To see this point, consider a variant of this example, where John's doctor has very good evidence that John has condition X, treatable only by drug Y. The doctor says: 'Take Y'. The doctor did exactly the right thing given her evidence; but it turns out that John's tests for condition X were a very rare false positive, and he doesn't have X. Now consider the question: was the doctor's prescription a good one? Well, it depends on what sense of goodness we focus on: it was not good for John, insofar as it was the wrong prescription; but it was evidence of good doctoring, insofar as it was the best reaction to her evidence.

A virtue of the present proposal is that it also accounts for an intuitive contrast between first- and third-personal attitude reports, pointed out to us by Bob Beddor and Andy Egan (p.c.). Consider a variant on the present example, in which instead of the doctor saying 'I think you have strep throat', a third party, Bob, says: 'John's

7 This proposal may have an underlying similarity to Kahneman and Tversky's suggestion that 'the answer to a question can be biased by the availability of an answer to a cognate question–even when the respondent is well aware of the distinction between them' (Tversky & Kahneman 1983).

mom thinks that John has strep throat'. Now suppose we find out that John doesn't have strep throat, and we are in a context where we are interested in the question of whether John actually has strep throat. Intuitively, there is very little inclination to say that what Bob said is false, unlike in the first-person case.

This is naturally accounted for on the present approach, since third-personal belief reports plausibly have a very different update effect than first-personal ones. Insofar as conversation aims to coordinate on the beliefs of the interlocutors, an assertion of ⌜I think that p⌝ will standardly be taken to be a proposal to commonly accept p; whereas an assertion of ⌜S thinks that p⌝ will not typically be taken this way. (Of course, it *could* be taken this way, insofar as the interlocutors treat S as a relevant authority; but it need not be.) And so we predict that retrospective judgments about third-personal attitude reports will have a quite different profile than retrospective judgments about first-personal attitude reports. This is a prediction we hope will be borne out in future experimental work.

Similar points can be made about the other cases we explored above, and about judgments about retraction and consistency as well as judgments about truth. Consider the first result reported above—our variation on Knobe and Yalcin's case—where Sally says either 'Joe might be in Boston' or 'I think Joe is in Boston'. Assume that ⌜Might p⌝ is a proposal to make p compatible with the context—again, an idea that has been spelled out in different ways in both contextualist and expressivist frameworks. Then the first of these is a proposal to leave it open that Joe is in Boston; the second a proposal to lend high credence to this proposition. Was Sally making a good proposal in making these proposals? Well, in one sense, perhaps yes: insofar as they conformed with her current evidence, they were perfectly good proposals. From another, more objective, sense, they were bad proposals: since Joe wasn't in Boston, it was, in some objective sense, wrong to leave it open that he was in Boston. Judgments about retraction, and perhaps to a lesser degree truth, might track either one of these senses. One thing that speaks in favor of this account in this particular case is that it explains the fact that subjects were even more likely to reject the 'think' claims as false than they were to reject the 'might' claim as false. This fits nicely with our account, since the former of these makes a stronger proposal than the latter—namely, to lend high credence to the proposition that Joe is in Boston, rather than simply to leave this open. So there is an obvious sense in which, from an objective point of view, the first of these is a less good proposal than the latter, and so more natural to reject.

Finally, the observed judgments about consistency fit naturally into this framework. Recall that subjects are somewhat inclined (a bit above a midpoint) to judge that p and ⌜Might not p⌝ are inconsistent, and likewise that p and ⌜I think not p⌝ are inconsistent. This is naturally accounted for if what they are judging to be inconsistent are sometimes the associated proposed updates: plausibly, one cannot update

the same common ground with both members of either of these pairs. Judgments of inconsistency were not as high as for non-modal claims, however, which is also what we would expect if those judgments can also sometimes target plain old semantic consistency, rather than update consistency (since p is semantically consistent with ⌜I think not p⌝, on any reasonable account, and likewise with ⌜Might not p⌝, on many reasonable accounts).

We want to emphasize that we are not saying much theoretically new here: our account primarily generalizes existing insights from the literature. At the same time, we think our results provide strong new evidence in favor of an account along the lines we've proposed. This account—on which retrospective assessment can focus on the proposal made, rather than the literal truth-value of what was said—gives a unified explanation of both the modal data and the 'think' data. And it nicely accounts for the QUD sensitivity that Beddor and Egan observe for 'might' and that we replicate for 'think': changing what question we are currently concerned about will naturally lead to changes in which aspect of the proposal we are judging, and so to changes in retrospective judgment. Given the matching 'think' data, we think that the current proposal is a far more plausible one of the epistemic modal data than one that makes central use of relativist or expressivist resources. [8]

6 Conclusion

What is eavesdropping good for? Our main point in this paper has been negative: eavesdropping isn't good for much in deciding between different theories of epistemic modals—and hence different theories of truth. Our argument has been the following. We find the same patterns of cross-contextual judgments for claims with the form ⌜I think p⌝ as for claims of the form ⌜Might p⌝ and ⌜Probably p⌝. Relativist and expressivist views have been proposed to account for the latter kinds of claims.

8 A different kind of approach would develop a suggestion from von Fintel & Gillies 2008: that retrospective judgments can sometimes target the prejacent of a modal claim, rather than the claim as a whole. Extending this proposal to 'think' is straightforward. However, we think the QUD-sensitivity we observe in the latter case speaks strongly against this approach. On this approach, what would explain our tendency to judge 'I think you have strep throat' false when we are concerned with whether John has strep throat would be a tendency, in this case, to target the prejacent ('You have strep throat') rather than the sentence as a whole. But what would explain our tendency to judge the sentence *true* when we are concerned with whether the doctor is good? On a prejacent-targeting story, we would say that, in the latter case, we tend more to take the whole sentence into consideration, rather than just its prejacent. But this doesn't make much sense when we want to know whether John's doctor is good: after all, being a good self-reporter of beliefs isn't very good evidence that one is a good doctor. This approach would also not, as far as we can tell, make sense of the contrast between first- and third-personal reports noted above. So we think something more needs to be said here: a prejacent-targeting story may well be part of the picture, but we doubt that, once QUD sensitivity is taken into account, it can account for the bulk of the judgments reported above.

But what the relativist or expressivist would have to say to account for the attitude report cases in a parallel fashion are just not plausible. Moreover, as we have argued, we *do* want a unified account: the patterns we find across the three experiments reported here are too similar for a disjunctive strategy to appear attractive. So it does not look like a relativist or expressivist account is promising.

We have combined this negative point with a sketch of a positive account, building on existing proposals in the literature, in particular that of Khoo 2015, on which cross-contextual judgments are often based on the appropriateness of the proposal that has been made, rather than the literal truth-value of what was said. We have argued that this provides a promising account of both the modal data and the 'think' data, and that its plausibility is further augmented by the QUD-sensitivity that we observe for both. We should emphasize that our positive proposal is separable from our central negative point, and that one could try to develop a different unified account of the phenomena. It is, however, important for our negative point that there exists a unified positive account; if there were not one, then a disjunctive approach would look much more plausible, and our negative point would be undermined. We hope to have said enough to show that there is indeed good reason to think that a unified positive account can be had.

Our particular positive proposal is consistent with a variety of different underlying theories of epistemic modals, including contextualist, expressivist, and relativist ones—provided they account for the characteristic update effects of modal claims. And so our negative point should not be interpreted as an anti-relativist or anti-expressivist argument on its own. But, as we have noted, relativism and expressivism have often been centrally motivated on the basis of cross-contextual judgments; and our central claim has been that cross-contextual judgments do not motivate relativism and expressivism. So, insofar as we see contextualism as a default position that we have only departed from because of these puzzling cross-contextual judgments, our negative point may specifically put argumentative pressure on the motivation for relativism or expressivism.

Let us close by noting that our positive proposal is very much in line with a broadly Gricean perspective on conversation. Grice's work emphasized that speakers tend to be cooperative, even when this entails being non-literal in a variety of ways. Beddor and Egan's results, and our corresponding development of them, show that eavesdropping judgments are very sensitive to the conversation's QUD, and thus to the manifest goals of the conversation. This, in turn, suggests that judgments in this area constitute just one instance of this general phenomenon, in which speakers who are asked about truth may respond in a slightly non-literal, but manifestly more helpful, way; corresponding caution is therefore needed before we interpret those responses as telling us anything directly about truth.

References

Beddor, Bob & Andy Egan. 2018. Might do better: Flexible relativism and the QUD. *Semantics & Pragmatics* .

Dowell, Janice. 2011. A flexibly contextualist account of epistemic modals. *Philosophers' Imprint* 11(14). 1–25.

Egan, Andy. 2007. Epistemic modals, relativism, and assertion. *Philosophical Studies* 113(1). 1–22.

Egan, Andy. 2011. Relativism about epistemic modals. In Steven Hales (ed.), *Blackwell companion to relativism*, 219–241. Wiley-Blackwell.

Egan, Andy, John Hawthorne & Brian Weatherson. 2005. Epistemic modals in context. In Gerhard Preyer & Georg Peter (eds.), *Contextualism in philosophy: Knowledge, meaning and truth*, chap. 6, 131–169. Oxford University Press.

von Fintel, Kai & Anthony Gillies. 2008. CIA leaks. *Philosophical Review* 117(1). 77–98.

von Fintel, Kai & Anthony Gillies. 2011. 'Might' made right. In Andy Egan & Brian Weatherson (eds.), *Epistemic modality*, Oxford University Press.

Grice, Paul. 1989. *Studies in the way of words*. Harvard.

Khoo, Justin. 2015. Modal disagreements. *Inquiry* 58(5). 511–534.

Khoo, Justin & Jonathan Phillips. 2019. New horizons for a theory of epistemic modals. *Australasian Journal of Philosophy* .

Knobe, Joshua & Seth Yalcin. 2014. Epistemic modals and context: Experimental data. *Semantics and Pragmatics* 7(10). 1–21.

Kölbel, Max. 2009. The evidence for relativism. *Synthese* 166. 375–95.

Lasersohn, Peter. 2009. Relative truth, speaker commitment, and control of implicit arguments. *Synthese* 166(2). 359–374.

MacFarlane, John. 2011. Epistemic modals are assessment sensitive. In Andy Egan & Brian Weatherson (eds.), *Epistemic modality*, Oxford University Press.

MacFarlane, John. 2014. *Assessment sensitivity: Relative truth and its applications*. Oxford: Oxford University Press.

Mandelkern, Matthew. 2018. How to do things with modals. To appear in *Mind & Language* .

Mandelkern, Matthew. 2019. Bounded modality. *The Philosophical Review* 181(1). 1–61. http://dx.doi.org/10.1215/00318108-7213001.

Moss, Sarah. 2015. On the semantics and pragmatics of epistemic vocabulary. *Semantics and Pragmatics* 8(5). 1–81.

Ninan, Dilip. 2016. Relational semantics and domain semantics for epistemic modals. *Journal of Philosophical Logic* 47(1). 1–16. http://dx.doi.org/10.1007/s10992-016-9414-x.

Stalnaker, Robert. 1978. Assertion. In Peter Cole (ed.), *Syntax and semantics*, vol. 9, 315–322. New York: Academic Press.

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

Stephenson, Tamina. 2007a. Judge dependence, epistemic modals, and predicates of personal taste. *Linguistics and Philosophy* 30(4). 487–525.

Stephenson, Tamina. 2007b. A parallel account of epistemic modals and predicates of personal taste. In Puig-Waldmueller (ed.), *Sinn und bedeutung*, vol. 11, 583–97.

Swanson, Eric. 2015. The application of constraint semantics to the language of subjective uncertainty. *Journal of Philosophical Logic* 45(121). 121–146.

Tversky, Amos & Daniel Kahneman. 1983. Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review* 90(4). 293–315.

Yalcin, Seth. 2007. Epistemic modals. *Mind* 116(464). 983–1026.

Yalcin, Seth. 2012. Bayesian expressivism. *Proceedings of the Aristotelian Society* CXII(2). 123–160.