Eavesdropping: What is it good for?*

Jonathan Phillips
Dartmouth College

Matthew Mandelkern
New York University

Abstract Judgments about truth, retraction, and consistency across contexts have been used in recent years to argue both for and against the revisionary theses of relativism about truth and expressivism about apparently truth-apt expressions like epistemic modals. We show that we find the same patterns that have been observed for epistemic modal claims like ⌜Might p⌝ when it comes to first-person attitude claims with the form ⌜I think that p⌝. This poses a serious challenge to many extant accounts of eavesdropping judgments — whether relativist, expressivist, or contextualist in nature — because extending these treatments to the corresponding 'thinks' judgments is *prima facie* implausible. Moreover, we argue, it suggests that eavesdropping judgments will not play an essential role in deciding between these views.

> agreement and disagreement is possible in regard to any aspect of a man's view and need not take the simple form of 'p', '$\overline{p}$'

1 Introduction

Consider this 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 assessment across contexts, which we call *eavesdropping judgments* — can be 'yes', and that this reveals something striking about truth. Namely, relativists argue that this shows that the truth of claims like Sally's depends on the *assessor's* evidence, not the speaker's evidence. Such a position requires a radical adjustment in our thinking about truth: on this way of thinking, the truth of a sentence is not fixed simply by the context in which it is asserted and the world where it is asserted; it also depends on a context where it is *assessed*.[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 hold that the context of assertion and the world of evaluation together suffice to fix the truth-value of claims like Sally's. They have

---

[1] E.g. Egan, Hawthorne & Weatherson 2005, Egan 2007, Kölbel 2009, Egan 2011, MacFarlane 2014, Beddor & Egan 2018, Khoo & Phillips 2019. Cf. Lasersohn 2009, Stephenson 2007a,b who also defend relativism, but not motivated by eavesdropping judgments.

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

pushed back against these revisionary claims, aiming to account for the judgments that have motivated relativism and expressivism within 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 played a central role in this debate — in fact do not help us decide between these views at all. Our argument, which builds on observations in von Fintel & Gillies 2008, is simple: eavesdropping 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 eavesdropper judgments about attitude constructions will essentially involve relativist or expressivist resources; and so it looks unlikely that the judgments in the modal 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 indeed 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, following similar claims in von Fintel & Gillies 2008, is that it can be reasonable to answer 'yes' to these questions, and, indeed, that this can be reasonable to just the same degree

---

3 E.g. von Fintel & Gillies 2008, 2011, Dowell 2011, 2017.

that it is reasonable to answer 'yes' to the parallel questions in the 'might' variant.

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 a context where it is assessed (on a relativist line), or is non-truthvalued (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 be determined by what the *assessor* of the report believes; or that it will generally be truthvalueless; or that it will generally be highly sensitive to what information is salient in the context of assertion (in the way that epistemic modals might be).

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 of-fer a single explanation for the general pattern. Such an explanation is un-likely 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 specifically rela-tivist/contextualist/expressivist features of epistemic modals. A more plau-sible 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 eavesdropping judg-ments are probably not helpful in distinguishing between relativism, contex-tualism, 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 eavesdropping judgments (especially armchair judgments, but also to a certain degree experimental work, especially Beddor & Egan 2018) 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 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 to date, namely those in Knobe & Yalcin 2014, Beddor & Egan 2018, and Khoo & Phillips 2019. These results comprise a mixed bag for relativism and expressivism: Beddor & Egan (2018) and Khoo & Phillips (2019) both take their results to support some forms of relativism, while Knobe & Yalcin (2014) take their results to undermine arguments for relativism on the basis of eavesdropper judgments. We choose the most critical experiment from each of those papers and conducted a variant of each experiment which simply replaces the relevant epistemic modal claim with a 'think'-claim (by *'think'-claim* we mean a first person belief ascription with the form ⌜I think that p⌝). In each case, we find that subjects' judgments pattern the same way in the 'think'-variant as in the original epistemic modal variant, supporting our central claim: that we need an explanation of eavesdropper judgments in the case of epistemic modals that extends to parallel judgments about 'thinks'-claims.[4] 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 essentially relativist or expressivist resources 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

---

4 Some relativists also argue on the basis of eavesdropper judgments for relativism about deontic modals, predicates of personal taste, knowledge, and other terms; see MacFarlane 2014 for the most extensive survey of applications. In this paper we will focus on epistemic modals, largely because those judgments have been studied the most systematically, though we expect our central points to extend broadly: insofar as these terms all show similar patterns of eavesdropper judgments to 'thinks'-claims, we should look for a unified explanation of them all.

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 bare claim about Joe's location (a *bare* claim is one that contains neither modals nor attitude predicates):

> **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:
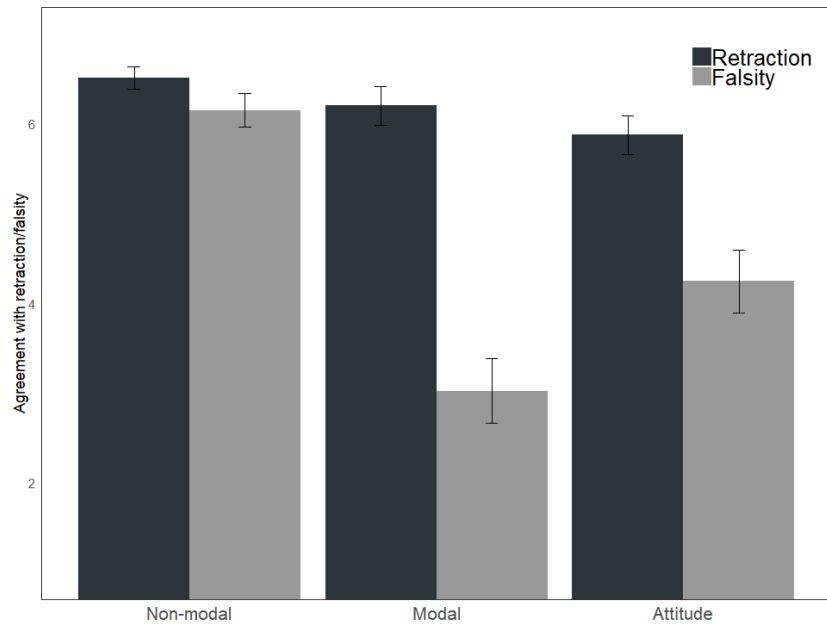
Figure 1   Graph of participants' agreement ratings that retraction would be appropriate (dark bars) or that the claim was false (light bars) as a function of whether the utterance involved a bare non-modal assertion, an epistemic modal claim, or an attitude report. Errors bars indicate +/- 1 *SEM*.

- 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 Phillips & Mandelk-ern 2020.

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.[5] Simi-lar to Knobe & Yalcin (2014), we observed the critical interaction effect be-tween question-type (Retraction vs. Falsity) and statement-type (Non-modal vs. Modal), $F(1, 156) = 34.64$, $p < .001$, $\eta_p^2 = 0.182$. This interaction was driven by the fact that we observed no significant difference between judg-ments 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 ex-tended 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,

---

5 An overall analysis of variance of all responses 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.

$t(80) = -2.45$, $p = 0.016$, $d = 0.541$. We return to this 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 a context where they are assessed. 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 Experiment 5 in Beddor & Egan 2018, which we think controls for some independent confounds in a helpful way as discussed by Beddor and Egan. (We come back to one these potential confounds, namely prejacent targeting, in §5.) We replicated their experiment and added 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 **Modal / QUD-Prejacent** condition, the continuation focused on whether the prejacent is true—i.e., on whether John has strep throat:

> **Modal / 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 **Modal / QUD-Competence** condition, the continuation focused instead on the competence of John's primary care physician:

> **Modal / 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 & 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 **Modal / QUD-Prejacent** condition selected 'No, it's not', while only 28.81% of the participants in the **Modal / QUD-Competence** condition did (see Fig. 2).

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).

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, $p \geq 0.121$.
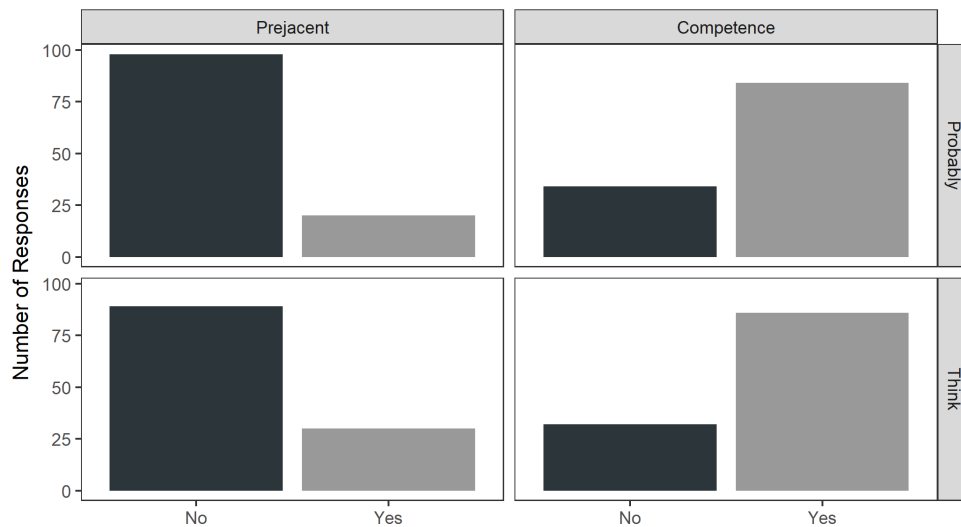
Figure 2   Graph of the number of times each response option was selected as a function of both whether the utterance involved a modal claim (**Probably**) or an attitude report (**Think**), and whether the assessment QUD targeted the **Prejacent** or the doctor's **Competence**.

In sum, Beddor & 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 stan-

dard relativist approach (according to which, in particular, ⌜Might p⌝ and ⌜Not p⌝ / ⌜Must not p⌝ are inconsistent if they are *assessed* at the same context, even if they are *asserted* in different contexts) or a standard contextualist approach (on which ⌜Might p⌝ and ⌜Not p⌝ are consistent), 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 (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 an 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 **not** 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 assessed 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) (see Fig. 3)

Next, we analyzed participants' compatibility judgments in the **Modal** and **Non-Modal** conditions with a 2 (Statement: Bare vs. Modal) × 2 (Condition: Utterances vs. Assessments) ANOVA. Replicating Khoo & Phillips (2019), participants' 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, $F(1, 199) = 0.091$, $p = .764$, $\eta_p^2 < 0.001$, 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

We then 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, $F(1, 202) = 0.213, p = .644, \eta_p^2 = 0.001$, 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.

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 judgments that have been at the heart of the debate between relativism, contextualism, and expressivism. We think that this, in turn, suggests that these
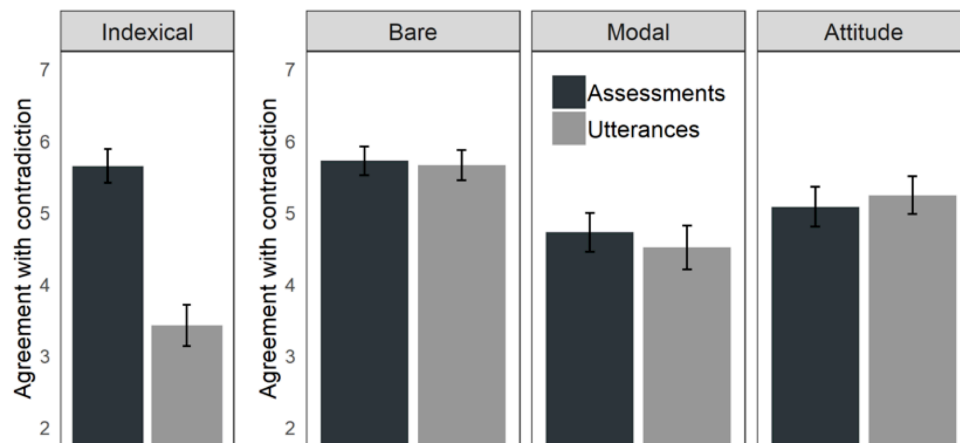
Figure 3   Participants' mean level of agreement that at least one of the in-spectors' claims must be false. Errors bars indicate +/- 1 *SEM*.

phenomena cannot play a central role in deciding between these views. Let us lay out our argument for this conclusion.

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 each separate 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 remarkably 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, start by thinking 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 *prima facie* unreasonable to think that sensitivity to skeptical scenario-raising varies with contexts of assessment rather than the context of 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 a context of assessment rather than the context of assertion.

But none of this 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 '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 assessing subject's attitude state in a context of assessment: the fact that we know in the later context that Joe is not in Boston would somehow have to suffice to make Sally's earlier claim 'I think Joe is in Boston' false, as judged at a context where it is assessed. But this idea has not been defended in the literature, and it strikes us as 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'. 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, as assessed in our context, 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 implausible. *We* can't change facts about what the doctor thought at some point in the past simply by finding out that it was probably false.

Analogously, to account for the 'think' phenomena in a parallel way to the modal phenomena, an expressivist approach would have to say something along the following lines: 'I think Joe is in Boston' expresses a state of mind that entails (or lends sufficient credence to) the proposition that Joe is in Boston, but 'I think Joe is in Boston' is neither true nor false; so you should

accept 'I think Joe is in Boston', as asserted by Sally in c, just in case *your* state of mind entails that Joe is in Boston. The problem is that it is intuitively clear that, while a claim like this *does* of course express the state of mind in question, it *also* has truth conditions, which depend on whether Sally is in c in fact in the relevant state of mind.

The upshot of this is that giving a unified explanation of eavesdropping judgments which covers both epistemic modals and 'think'-claims in a way that essentially uses relativist or expressivist resources would face substantial challenges. We do not know of any accounts in the literature that could help face those challenges, and we are pessimistic about their prospects for success, given the points just outlined.

By contrast, the forms of relativism (or expressivism) that have been put forward about epistemic modals, predicates of taste, and so on are at least *prima facie* much more plausible than the kinds of relativism (or expressivism) about 'thinks' that we are considering. To say that the truth of ⌜Might p⌝ or ⌜That's tasty⌝ depends on the assessor's information or standards, or that these are truth-valueless, seems like a possibility worth exploration, in a way that these relativist or expressivist hypotheses about 'think' simply don't.

And what about contextualism? On standard contextualist approaches to eavesdropping, like those in von Fintel & Gillies 2008, 2011, Dowell 2011, 2017, the aim is essentially to use contextualist resources to make sense of eavesdropping judgments about epistemic modals. Extending these kinds of approaches directly to the corresponding 'thinks' data does not seem promising to us, for similar reasons to those given above vis-à-vis relativism and expressivism. For instance, these accounts generally appeal to flexibility in supplying the source of evidence relative to which epistemic modal claims are assessed. On these accounts, this flexibility plays an essential role in accounting for eavesdropping judgments. But there is intuitively not similar flexibility in the interpretation of 'thinks'-claims: whether ⌜I think that p⌝, as asserted by Sally in context c, is true intuitively depends just on whether Sally's state of mind at the time of c supports p (make sense of 'support' with whatever theory of belief you prefer); we do not have flexibility in saying, for instance, that this claim is false if Sally's state of mind at the time of c supports p but there is some other very salient body of evidence, say our evidence, or our doctor's evidence, or whatever, that doesn't support p.

We can distinguish this sort of contextualist response to eavesdropping data from ones that try to explain eavesdropping judgments by way of inde-

pendently motivated, general considerations about how people think about disagreement, retraction, and so on (as in e.g. Khoo 2015). We are very sympathetic to this latter kind of approach, and will build on it presently. However, on this approach, context-sensitivity is not *itself* being proposed as an explanation of these phenomena; rather, a more general explanation, consistent with contextualism, is being proposed. In short, we do not think that appeal to the context-sensitivity of 'think' can account for eavesdropping judgments about 'think'.

In sum, then, it looks like, if we want a unified account of eavesdropping judgments involving attitude reports and epistemic modals, that account will not essentially involve the resources of relativism, expressivism, or context-sensitivity, but rather will come from independent considerations about eavesdropping judgments more generally.

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 argument is that eavesdropping judgments do not on their own 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 *substantially motivated* by considerations about eavesdropping 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 eavesdropping 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 eavesdropping judgments don't on their own provide any evidence for or against relativism, expressivism, or contextualism.[6]

---

6 Both relativism and expressivism have also been motivated on the basis of embedding data (for the former, see Stephenson 2007a,b, Lasersohn 2009; for the latter, see Yalcin 2007). But contextualist theories have been developed which account for these data (at least when it comes to epistemic modal; see in particular Ninan 2016, Mandelkern 2019).

5.2 Positive accounts

We have now stated our central point, which is negative. Still, a natural question to ask is how one *could* give a unified account of the eavesdropping judgments concerning epistemic modals and attitude reports. Answering this question is not our main goal. But we do want to say enough here to make it plausible that a unified account *can* be given. This is dialectically important because a key point in our negative argument is that we should aim to give a unified account of these judgments; if there was no satisfying unified account of these judgments on offer, it would be more tempting to give a disjunctive account that dealt with modal judgments in one way and 'think'-claims in a different way. We will briefly sketch two different ways one might account for the data above in a unified way. While we find the first account more appealing, both accounts offer uniform explanations of the data and should be further explored in future work.

5.2.1 Leaving p open

The key idea behind the first approach comes from the suggestion in Khoo 2015 that disagreement can target the *proposed update* that an assertion makes to the common ground rather than its truth-value.[7] To motivate the idea, Khoo gives the following example. If someone says 'The bank is open', but you think this is not supported by the evidence, you could reasonably disagree by replying 'No, the bank *might* be open.' Intuitively, it's not that you think what the person said is *false*, but rather that you want to resist the proposed update they have made: you think that the proposal they have made about how to update the conversation's common ground — namely, the proposal to accept that the bank is open — is overly strong, since it would not leave open the possibility that the bank is closed. Khoo's idea is that modal disagreements are often structurally like this: when a modal assertion is made, you can evaluate the *proposal* that was made, and express agreement

---

7 These ideas are tentatively endorsed in Knobe & Yalcin 2014. Roberts (2017) develops a similar proposal, arguing that 'retractions and disavowals are basically about one's former beliefs, rather than about the truth of the statements that reported them', which echoes similar ideas in von Fintel & Gillies 2008. There are, however, differences between the proposals: for instance, Khoo's proposal predicts a difference in eavesdropper judgments between first- and third-personal belief ascriptions, which have different update effects, whereas it is not obvious that the belief-based idea does.

or disagreement with that proposal, which might come apart from your judgment about the truth or falsity of the proposition that was asserted.

Khoo originally applied this idea to expressions of (dis)agreement, but it is natural to extend Khoo's suggestion to expressions involving truth-value, retraction, and consistency. 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, and thus their intuitive judgments may be about the *proposed change to the common ground* rather than the proposition asserted.[8]

The final piece needed to give a unified explanation of our data is a theory on which modal- and 'think'-claims have similar characteristic update effects. If they do, and if eavesdropping judgments can target the proposed update rather than the asserted proposition, then we will have an explanation of why eavesdropping judgments about modal claims and 'think'-claims are so similar.

To flesh this out, we'll sketch one approach that seems natural to us. First, following much literature (going back to Stalnaker 1970, Groenendijk & Stokhof 1975), we adopt the following observation about modal updates. As Khoo (2015) puts it:

> *The Update Observation:* Generally, assertively uttering an epistemic possibility sentence involves proposing that it not be common ground that its prejacent is false.

---

8 Of course, there are limits to this. For instance, 'I disagree' or 'That's not true' can't be used as a response to an assertion which is, say, merely irrelevant or rude. An important question is exactly what the limits are; generally speaking, these reactions seem appropriate only as epistemic evaluations.

Our claim 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 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 proposed update it made was a good one. (Compare Tversky & Kahneman (1983)'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'.) This is especially so in cases where the questioner makes clear that what they *care* about is whether the proposal in question was a good one, as illustrated in the QUD paradigm.

To spell this idea out, begin with the pragmatic framework from Stalnaker 1974, 1978, who models the common ground of a conversation as the set of propositions which are believed by all the conversants at the time of the conversation, believed to be believed, and so on. We write $CG_{\mathcal{S}}^{t,w}$ for the common ground of group $\mathcal{S}$ at time t in world w. Given a conversation with participants $\mathcal{S}$, Stalnaker proposes that an assertion of p at a time t is a proposal for the conversants to make p common ground at the prospective time t' at which the assertion has been considered and either accepted or rejected. *The Update Observation* thus says that, in asserting something like 'You [might/probably] have strep throat', you are, inter alia, proposing that it *not* be common ground that John does *not* have strep throat. Put differently, you are proposing that it remain compatible with the common ground that John does have strep throat—that you leave this possibility *open* in the sense that it both remains compatible with what you commonly accept, and it is commonly accepted that it remains compatible with what you accept.

There are a variety of ways to capture the *The Update Observation*. A simple one, following the theory developed in Stalnaker 1970, 2014, Mandelkern 2020, Kratzer 2020, says that ⌜Might p⌝, as asserted at t, simply says that p is compatible with the common ground at the prospective time t'. This is a minimal way to account for the *Update Observation*, since, on this account, if ⌜Might p⌝ is asserted at t and accepted at t', then it will be common ground at t' that p is compatible with the common ground, which guarantees that p will be compatible with the common ground at t', thanks simply to the logic of common ground.[9]

Intuitively, 'think'-claims *also* have the effect of proposing that their prejacent remain compatible with the common ground. If you say to John, 'I think you have strep throat', then you are, inter alia, proposing that it be left open, in your conversation, that John has strep throat. More generally, suppose that S asserts ⌜I think that p⌝ to a group $\mathcal{S}$, and that this assertion is accepted at t'.[10] We assume, moreover, the simple Hintikkan semantics for 'think', on which ⌜I think that p⌝ is true, as asserted by S at world w and time t, just in case all the worlds compatible with S's beliefs at w and t entail p (Hintikka 1962). That means that, at t', it is common ground that S believes

---

9 Assuming that the common ground is itself consistent; see Mandelkern 2020 for more careful exposition.

10 We assume, moreover, that it is common ground that nothing has happened between t and t' to change S's belief that p, i.e., it is common ground that, if ⌜S thinks p⌝ is true at t, then it is also true at t'.

that p. But that, in turn, means that, at t', p is consistent with the common ground (and that this fact, moreover, is common ground); once again, this is guaranteed simply by the logic of the common ground. So asserting ⌜I think that p⌝, like asserting ⌜[Might/probably] p⌝, has the characteristic effect of ensuring that p remains an open possibility.

This simple contextualist approach thus accounts for the *Update Observation* for epistemic modals together with the parallel observation for 'think'-claims, and thus lays the foundation for a unified way of explaining the similarity in patterns of judgments involving them. To make this a bit more concrete, suppose that the doctor says 'I think you have strep throat'. On the present approach, whether this is true or false depends just on the doctor's mental state. But suppose that eavesdropping judgments can target the proposed update effect, rather than just the truth-value of what was asserted. If you think that, at that stage in the conversation, the conversants *should not have left open the possibility that John had strep throat*, then you could resaonably dispute the doctor's assertion—even if you think that she is correctly reporting her mental state. Similar points go for epistemic modals, accounting for the parallels between them.

5.2.2   Speech act modifier

A different approach regards both epistemic modals and 'I think' as speech-act modifiers which serve to hedge an assertion, rather than as part of the assertion itself. On this approach, ⌜I think p⌝ and ⌜[Might/probably] p⌝ can both be used to update the common ground *with p* itself, albeit in a more cautious manner than an assertion of p itself. Krifka (2019: p. 85) gives a characteristic statement of the view: 'the function of subjective epistemics is to put the non-modalized proposition into the common ground in a more cautious manner'. 'Subjective epistemics' here includes both modal and 'think'-claims. On this view, as Krifka puts it, subjective epistemics 'are not part of the proposition to be put into the common ground, but rather belong to the machinery by which this is done'. And thus *both* ⌜I think that p⌝ and ⌜[Might/probably] p⌝ can be proposals that p itself should become part of the common ground.

This general kind of approach obviously has the potential to predict the parallel judgments between epistemic modal and 'think'-claims, since, on this kind of account, both are essentially proposals to add p to the common ground (with slightly different hedges). This kind of account could be coupled with Khoo's idea that eavesdropping judgments target the proposal,

rather than the truth-value, of the assertion. However, a more natural way to flesh out this approach is to simply say that eavesdropping judgments can *either* target the prejacent *or* the whole claim, including the hedge. That is, ⌜I think that p⌝ puts forward two propositions, on this view: both p, and that the speaker thinks that p; and subsequent judgments might target either one.[11]

5.3 Accounting for QUD effects

To have a full account of the patterns we have brought out, we need to explain the striking QUD sensitivity that Beddor and Egan report about modal claims, and the corresponding observations 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.

The first approach we outlined above can naturally explain these data. Both of these assertions amount to, inter alia, a proposal for the conversants to leave it open that John doesn't have strep throat. There are two natural ways to judge such a proposal, depending on what your main concerns are. If you are primarily concerned with the *facts of the matter about John's case*, then these were objectively misleading proposals: the conversants objectively *shouldn't* have left it open that John didn't have strep throat, since, in fact, he probably did. On the other hand, if we are primarily concerned with *whether John's doctor is reliable*, then these update proposals look reasonable, since they were *justified* by the doctor's evidence. Thus, there are different respects in which a proposed update can be evaluated, depending on the cooperative aims of the conversants, and shifting focus from one of those respects to another can naturally shift our judgments about the proposed update.

The second approach can also explain the patterns above by positing that, when subjects judge that 'I think you have strep throat' is false, they are targeting the prejacent rather than the sentence as a whole; while when they judge it to be true, they are instead taking the whole sentence (including

---

11 One can find roots of this general picture in Lyons 1977, Recanati 1987, von Fintel & Gillies 2007, 2008, Dowell 2011, Krifka 2014, Murray 2014, Greenberg & Wolf 2018, Krifka 2019, Incurvati & Schlöder 2020.

the attitude ascription) into consideration. Indeed, Murray (2014: pp. 16, 42) (citing Simons et al. 2010) argues that whether a prejacent or a whole modal clause is targeted depends on what QUD is salient. We might posit that, in general, the prejacent is a salient target just in case it answers the QUD. Thus when the salient QUD is whether or not John has strep, subjects will more likely target the prejacent, while when the QUD is whether or not John's doctor is good, the prejacent is less available as a target.[12]

5.4   Future directions

Recall that our purpose in sketching these two accounts was to demonstrate that a unified account of the patterns for 'might'- and 'think'-claims is possible. And indeed, a range of different unified accounts looks likely to be available. Future theoretical and empirical work should test and refine these theories. In closing this section, we want to briefly point to important places for future exploration.

One helpful point for further exploration concerns the difference between first- and third-personal attitude reports, as Bob Beddor and Andy Egan have pointed out to us. Consider a variant on our key example. Instead of the doctor saying 'I don't think you have strep throat', a third party, Bob, is asked what John's mother thinks about John, and says: 'John's mom thinks that John doesn't have strep throat'. Intuitively, the attitude ascription here does not have the characteristic update effect of first-personal 'think'- or 'might'-claims of ensuring that it is common ground that its prejacent is left open. This intuition aligns with the predictions of the update-proposal account: eavesdroppers should agree with the assertion just to the degree that they think the proposed update *about John's mother's beliefs* is appro-

---

12 Beddor & Egan (2018: Section 6) argue against a prejacent-targeting account, and we will briefly explain their worry. One of their experiments — replicated above in Section 3 — was specifically designed to test a prejacent-targeting account by ensuring that subjects do not actually find out that John *does* have strep, they only find out that it is *likely* that he has strep. In the prejacent condition, however, most still judge the doctor's assertion 'You probably don't have strep' (or 'I don't think you have strep', in our variant) to be false. This is *prima facie* surprising from the perspective of the prejacent-targeting view, since subjects don't know that John has strep. However, given that they face a forced choice between 'Yes' and 'No', perhaps they are just choosing the most likely option vis-à-vis the prejacent. More exploration is needed here: for example, one could add a non-modal condition ('You don't have strep') and test whether subjects similarly judge it to be false given their information in this case. To the extent that they do not, this evidence would be problematic for a prejacent-targeting approach.

priate, and should *not* be inclined to disagree with the assertion if, say, it becomes clear that John's mom thought John didn't have strep throat, but in fact she was wrong. Additionally, however, there are also cases in which the update-proposal account *does* predict that third-person attitude ascriptions will pattern like first-person belief reports. *If* the third party is treated as an epistemic authority on the matter in question, then third-personal attitude ascriptions should pattern like first-personal ones. Why? Because, if it is common ground that the speaker defers to S on the matter of p, then if it is common ground that the speaker thinks that S thinks that p, then it will be common ground that p is an open possibility.[13] As for the speech-act modifier account, to work out its predictions in cases like these, we would need to couple it with a general theory of how epistemic modal and attitude verbs are interpreted when embedded on the speech-act modifier approach (see e.g. Krifka 2014 for relevant exploration).

In our discussion thus far, we have emphasized the commonalities between 'might' and 'think', but there are obviously important differences between them, as well as between 'might' and 'probably'. Plausibly, 'probably' and 'think' are stronger than 'might' in some important sense in terms of the update they effect. While focusing on the commonalities above has simplified our discussion, these differences are important to keep in mind. For instance, they might help to explain the fact that, in our variant on the Knobe and Yalcin case (Section 2), subjects were even more likely to reject the 'think'-claim as false than they were to reject the 'might'-claim as false. If 'think' is stronger than 'might', this fits nicely with Khoo's update proposal account. More work, theoretical and experimental, is needed to explore these differences in more detail.

Finally, we have focused on epistemic modals throughout, which, again, are just one area where relativism has been motivated on the basis of eavesdropping judgments. Future work should explore similarities and differences with other kinds of language which have been argued to be relativist, like predicates of personal taste and future contingents.

6   Conclusion

What is eavesdropping good for? Our main point in this paper has been negative: eavesdropping isn't worth much in deciding between different theories

---

13 Parallel points apply to restricted epistemic modal claims like 'For all John's mother knows, he probably doesn't have strep throat.

of the 'post-semantics' of epistemic modals, or, more generally, different theories of truth. Our argument has been the following. We find the same patterns of eavesdropping judgments for assertions with the form ⌜I think p⌝ as for assertions of the form ⌜Might p⌝ and ⌜Probably p⌝. Proposals essentially involving relativism, expressivism, and context-sensitivity have been made to explain the latter kinds of patterns. But parallel moves in the case of attitude reports are just not plausible. Moreover, as we have argued, we *do* want a unified account of the two phenomena: the patterns we find across the three experiments reported here are too similar for a disjunctive strategy to be attractive.

We have combined this negative point with a sketch of two possible positive accounts, building on existing proposals in the literature. We should emphasize that the positive proposals we outlined are 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 exist some 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 found.

The particular positive proposals we looked at are consistent with a variety of different underlying theories of epistemic modals, including contextualist, expressivist, and relativist ones. 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 eavesdropping judgments; and our central claim has been that eavesdropping judgments do not motivate relativism and expressivism. So, insofar as one sees contextualism as a default position that we should only depart from because of these puzzling eavesdropping judgments, our negative point may specifically put argumentative pressure on the motivation for relativism and expressivism.

Let us close by noting a potentially broader upshot of our discussion for the interpretation of intuitions about truth-value, (dis)agreement, and consistency in general, stemming from the first positive proposal we sketched above. That proposal is very much in line with a broadly Gricean perspective on conversation. Grice emphasized that speakers tend to be cooperative, even when this entails being non-literal in various 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 which are apparently about truth-value may constitute just one instance of this general Gricean phenomenon, in which speakers who are asked about one thing (truth or consistency) may respond in a slightly non-literal, but manifestly more helpful, way, about something slightly different (in this case, the proposed update); which may have broad upshots for the interpretation of results in experimental semantics and pragmatics.

References

Beddor, Bob & Andy Egan. 2018. Might do better: Flexible relativism and the QUD. *Semantics & Pragmatics*. https://doi.org/10.3765/sp.11.7.

Dowell, Janice. 2011. A flexible contextualist account of epistemic modals. *Philosophers' Imprint* 11(14). 1–25. http://hdl.handle.net/2027/spo.3521354.0011.014.

Dowell, Janice. 2017. Contextualism about epistemic modals. In Jonathan Jenkins Ichikawa (ed.), *The Routledge handbook of epistemic contextualism*. New York: Routledge.

Egan, Andy. 2007. Epistemic modals, relativism, and assertion. *Philosophical Studies* 113(1). 1–22. https://doi.org/10.1007/s11098-006-9003-x.

Egan, Andy. 2011. Relativism about epistemic modals. In Steven Hales (ed.), *The Blackwell companion to relativism*, 219–241. Oxford: 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. 2007. An opinionated guide to epistemic modality. In Tamar Gendler & John Hawthorne (eds.), *Oxford studies in epistemology 2*, 32–62. Oxford: Oxford University Press.

von Fintel, Kai & Anthony Gillies. 2008. CIA leaks. *Philosophical Review* 117(1). 77–98. https://doi.org/10.1215/00318108-2007-025.

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

Greenberg, Yael & Lavi Wolf. 2018. Gradable assertion speech acts. *North East Linguistic Society (NELS)* 48(1). 271–280.

Groenendijk, Jeroen & Martin Stokhof. 1975. Modality and conversational information. *Theoretical Linguistics* 2(1-3). 61–112. https://doi.org/10.1515/thli.1975.2.1-3.61.

Hintikka, Jaakko. 1962. *Knowledge and belief: An introduction to the logic of two notions*. Ithaca, NY: Cornell University Press.

Incurvati, Luca & Julian J. Schlöder. 2020. Epistemic multilateral logic. *Review of Symbolic Logic*. https://doi.org/10.1017/S1755020320000313.

Khoo, Justin. 2015. Modal disagreements. *Inquiry* 58(5). 511–534. https://doi.org/10.1080/0020174X.2015.1033005.

Khoo, Justin & Jonathan Phillips. 2019. New horizons for a theory of epistemic modals. *Australasian Journal of Philosophy* 97(2). 309–324. https://doi.org/10.1080/00048402.2018.1484499.

Knobe, Joshua & Seth Yalcin. 2014. Epistemic modals and context: Experimental data. *Semantics and Pragmatics* 7(10). 1–21. https://doi.org/10.3765/sp.7.10.

Kölbel, Max. 2009. The evidence for relativism. *Synthese* 166. 375–95. https://doi.org/10.1007/s11229-007-9281-7.

Kratzer, Angelika. 2020. *What's an epistemic modal anyway?* Manuscript, University of Massachusetts at Amherst.

Krifka, Manfred. 2014. Embedding illocutionary acts. In Tom Roeper & Margaret Speas (eds.), *Recursion: Complexity in cognition*, vol. 43 (Studies in Theoretical Psycholinguistics). Cham: Springer. https://doi.org/10.1007/978-3-319-05086-7.

Krifka, Manfred. 2019. Commitments and beyond. *Theoretical Linguistics* 45(1-2). https://doi.org/10.1515/tl-2019-0006.

Lasersohn, Peter. 2009. Relative truth, speaker commitment, and control of implicit arguments. *Synthese* 166(2). 359–374. https://doi.org/10.1007/s11229-007-9280-8.

Lyons, John. 1977. *Semantics*. Cambridge: Cambridge University Press.

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

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

Mandelkern, Matthew. 2019. Bounded modality. *The Philosophical Review* 128(1). 1–61. https://doi.org/10.1215/00318108-7213001.

Mandelkern, Matthew. 2020. How to do things with modals. *Mind & Language* 35(1). 115–138. https://doi.org/10.1111/mila.12231.

Moss, Sarah. 2015. On the semantics and pragmatics of epistemic vocabulary. *Semantics and Pragmatics* 8(5). 1–81. https://doi.org/10.3765/sp.8.5.

Murray, Sarah. 2014. Varieties of update. *Semantics and Pragmatics* 7(2). 1–53. https://doi.org/10.3765/sp.7.2.

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

Phillips, Jonathan & Matthew Mandelkern. 2020. *Repository for* Eavesdropping: What is it good for? Version V1. Harvard Dataverse. https://doi.org/10.7910/DVN/BGA2BA.

Ramsey, Frank P. 1931/1978. General propositions and causality. In D.H. Mellor (ed.), *Foundations: Essays in philosophy, logic, mathematics and economics*, 133–51. London: Routledge & Kegan Paul.

Recanati, Francois. 1987. *Meaning and force: The pragmatics of performative utterances*. Cambridge: Cambridge University Press.

Roberts, Craige. 2017. Agreeing and assessing: Epistemic modals and the question under discussion. Presentation at Oslo Workshop on Non-At-Issue Meaning and Information Structure.

Simons, Mandy, Judith Tonhauser, David Beaver & Craige Roberts. 2010. What projects and why. *Semantics and Linguistic Theory (SALT)* 20. 309–327. https://doi.org/10.3765/salt.v20i0.2584.

Stalnaker, Robert. 1970. Pragmatics. *Synthese* 22(1-2). 272–289. https://doi.org/10.1007/BF00413603.

Stalnaker, Robert. 1974. Pragmatic presuppositions. In Milton K. Munitz & Peter Unger (eds.), *Semantics and philosophy*, 197–213. New York: New York University Press. https://doi.org/10.1093/0198237073.003.0003.

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

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

Stephenson, Tamina. 2007a. A parallel account of epistemic modals and predicates of personal taste. *Sinn und Bedeutung* 11. 583–97. https://doi.org/https://doi.org/10.18148/sub/2007.v11i0.666.

Stephenson, Tamina. 2007b. Judge dependence, epistemic modals, and predicates of personal taste. *Linguistics and Philosophy* 30(4). 487–525. https://doi.org/10.1007/s10988-008-9023-4.

Swanson, Eric. 2015. The application of constraint semantics to the language of subjective uncertainty. *Journal of Philosophical Logic* 45(121). 121–146. https://doi.org/10.1007/s10992-015-9367-5.

Tversky, Amos & Daniel Kahneman. 1983. Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review* 90(4). 293–315. https://doi.org/10.1037/0033-295X.90.4.293.

Yalcin, Seth. 2007. Epistemic modals. *Mind* 116(464). 983–1026. https://doi.org/10.1093/mind/fzm983.

Jonathan Phillips
Program in Cognitive Science
Dartmouth College
01 Reed Hall
Hanover, NH 03755
Jonathan.S.Phillips@Dartmouth.edu

Matthew Mandelkern
Department of Philosophy
New York University
5 Washington Place
New York, NY 10003
mandelkern@nyu.edu