Epistemic modals and context: Experimental data[*]

Joshua Knobe
Yale University

Seth Yalcin
University of California, Berkeley

Abstract

Recently, a number of theorists (MacFarlane (2003, 2011), Egan et al. (2005), Egan (2007), Stephenson (2007a,b)) have argued that an adequate semantics and pragmatics for epistemic modals calls for some technical notion of relativist truth and/or relativist content. Much of this work has relied on an empirical thesis about speaker judgments, namely that competent speakers tend to judge a present-tense bare epistemic possibility claim true only if the prejacent is compatible with their information. Relativists have in particular appealed to judgments elicited in so-called *eavesdropping* and *retraction* cases to support this empirical thesis. But opposing theorists have denied the judgments, and at present there is no consensus in the literature about how the speaker judgments in fact pattern. Consequently there is little agreement on what exactly a semantics and pragmatics for epistemic modals should predict about the pattern of judgments in these cases. Further theorizing requires greater clarity on the data to be explained. To clarify the data, we subjected eavesdropping and retraction cases to experimental evaluation. Our data provide evidence against the claim that competent speakers tend to judge a present-tense bare epistemic possibility claim true only if the prejacent is compatible with their information. Theories designed to predict this result are accordingly undermined.

Keywords: epistemic modality, context-sensitivity, assessment-sensitivity, relativism, experimental semantics, update semantics

1 Introduction

In recent work, a number of theorists (MacFarlane (2003, 2011), Egan et al. (2005), Egan (2007), Stephenson (2007a,b); see Egan 2011, Weatherson & Egan 2011 for overviews) have defended the following thesis:

1

(R) An adequate semantics and pragmatics for epistemic modals calls for some technical notion of relativist truth and/or relativist content.

Here are two examples of ways that (R) has been concretely developed, both widely cited.

First example: Egan (2007) holds that present-tense bare epistemic possibility sentences express centered worlds propositions:

> *It might be the case that P* is true relative to a centered world $\langle w, t, i \rangle$ iff it's compatible with everything that's within i's epistemic reach at t in w that P. (8)

This is a relativist theory of the content of epistemic modal sentences inasmuch as the propositions these sentences express will generally be variable in truth value with respect to distinct centers within the same world.

Second example: MacFarlane (2011) explores the idea of packaging the following compositional semantic value for epistemic possibility modals:

> $\ulcorner Might : \phi \urcorner$ is true at $\langle c, w, i, a \rangle$ iff for some w' in i, ϕ is true at $\langle c, w', i, a \rangle$ (164)

with a general definition of truth for occurrences of sentences at pairs of contexts along the following lines:

> An occurrence of sentence ϕ at context [of utterance] c_U is true as assessed at a context c_A iff ϕ is true at every point of evaluation $\langle c_U, w_{c_U}, i_{c_A}, a \rangle$, where:
>
> - w_{c_U} = the world of c_U,
>
> - i_{c_A} = the set of worlds that aren't excluded by the information that is relevant at c_A,
>
> - a = an assignment of objects from the domain of c to the variables. (176)

MacFarlane describes the view as relativist inasmuch as "it is features of the context of assessment, not the context of use, that determine which information state is relevant for the evaluation of epistemic modals" (176). When we speak generically of "relativist views" below, we have Egan 2007 and MacFarlane 2011 in mind as the primary exemplars.

Of course, (R) is vague, and it can be precisified in many ways. We are not interested in counting the ways. Rather, we are interested in one kind of motivation

one could have for pursuing some version of (R) in the first place. This is a certain thesis about speaker judgments, namely:

(J) Competent speaker/hearers tend to judge a present-tense bare epistemic possibility claim (BEP) true only if the prejacent is compatible with their information (whether or not they are the producer of that utterance); otherwise the BEP is judged false.[1]

(J) is an empirical claim about competent speaker judgments. To see the content of (J), consider some BEP appropriately produced by a subject *A* and evaluated for truth by distinct subjects *B* and *C*. Suppose further that *B* and *C* are both fully informed about what information was in the epistemic reach (in the sense of Egan 2007) of *A* at the time of the BEP. (J) entails that *B* and *C* will nevertheless tend to diverge concerning their truth value judgments about the BEP in the case where one but not the other knows that the prejacent is false. Normally we tend to think that when two subjects diverge in their truth value judgments about a single utterance, only one of them can be correct. But the relativist views just described open up an alternative avenue of response. If the truth value of a BEP does in fact vary with the information relevant at the context of assessment (MacFarlane), or if the content of any given BEP varies in truth value depending on what is in the epistemic reach of the centered world at which it is evaluated (Egan), then the divergence in truth value judgment between *B* and *C* is to be expected: we need not say that either is going wrong, or judging inaccurately.

The idea that (J), or something closely approximating it, is true has played a significant role in recent attempts to motivate (R). Egan and MacFarlane are both clearly animated by the thought that "people tend to assess epistemic modal claims for truth in light of what *they* (the assessors) know, even if they realize that they know more than the speaker (or relevant group) did at the time of utterance" (MacFarlane 2011: 160; see also Egan 2007: 2-5, the section entitled 'Motivation for relativism: eavesdroppers'). Their respective theoretical developments of (R) are designed to predict this putative fact about the way subjects are prone to assess epistemic modal claims.[2] The objective of this paper is to clarify what exactly the speaker judgments

---

1 Cf. Hawthorne 2004, p. 29, fn. 69. We interpret (J) to concern just those cases where the truth judgment occurs immediately after the BEP is produced. We are setting aside cases where there is a substantial time lag between the production of a BEP and its assessment.

2 A general point of clarification: we take it that the technical notion(s) of truth devised in the context of particular compositional semantic theories need not be such as to map onto ordinary naive speaker judgments of truth and falsity in any self-evident way. Ordinary speaker truth-value judgments are standardly assumed to be among the explananda of natural language semantics, but their relations to technical notions may well be highly indirect. (For additional discussion, see Yalcin 2014 forthcoming.) We read the relativists highlighted as positing a relatively direct relation between

are in this domain. Our approach is to frame a precise version of the empirical thesis, viz. (J), and to investigate experimentally whether it is correct.

To be clear, note that we are not claiming that (J) is the only motivation that theorists could offer, or have in fact offered, for relativist views. That would of course be mistaken. Stephenson (2007a), for example, defends a version of (R), but not in order to predict (J). One might of course try to defend (R) in myriad ways. We are not engaged in any general attack on (R). Rather, we are keying into exactly one kind of motivation that has loomed large among developers of (R)-like views, namely the idea that a semantics and pragmatics for the language of epistemic modality needs to predict, or at least sit naturally with, (J).

Why might anyone think something like (J) is correct? In motivating (J), relativists have drawn particular attention to two sorts of intuition-pumping cases, what we will call *extracontextual assessment* cases and *retraction* cases.

In an extracontextual assessment case, a scenario is constructed where a BEP is produced by an agent in some normal way, but it is stipulated that the prejacent of the BEP is, after all, false. Here is an example:[3]

> FAT TONY
>
> Fat Tony is a mobster who has faked his own death in order to evade the police. He secretly plants highly compelling evidence of his murder at the docks. The evidence is discovered by the authorities, and word gets out about his apparent death. The next evening, from his safehouse, Fat Tony watches a panel of experts on the news discussing the question of whether he is dead.
>
> Expert A has had a good look at the evidence found at the scene. "Fat Tony is dead," he says. Expert B has also had a good look at the evidence, but his assessment is more cautious. "Fat Tony might be dead," B says.

In the story, Expert B says 'Fat Tony might be dead', but Fat Tony is alive. Call Expert B's utterance token MIGHT. The idea is to use this case to examine judgments about the truth-value of MIGHT, as issued by *extracontextual assessors*: competent speakers who (a) are outside of the discourse context in which the BEP is generated, and (b) know that the prejacent of that BEP is false. In the FAT TONY case, one kind of extracontextual assessor is Fat Tony himself. He is an *eavesdropper* on the discourse where MIGHT occurs in the sense we are interested in. Another kind

---

the technical notions of truth they postulate and ordinary speaker truth value judgments, at least in the relevant cases. We interpret this as a substantive dimension of the relativists' theoretical position.

3 For examples in the literature, see for instance Dialogue 3 of MacFarlane 2003: 5, or the James Bond eavesdropping case described by Egan 2007: 2-5.

of extracontextual assessor is the reader of the vignette. Although the reader does not "eavesdrop" on the imagined discourse in any normal sense, she is not in the discourse context of MIGHT, and she possesses the (stipulated) information that the prejacent is false (relative to the world of the vignette). This qualifies her for the status of extracontextual assessor.[4]

What is the judgment of extracontextual assessors concerning the truth value of MIGHT? Relativists like Egan (2007) have generally taken it that these sorts of cases provoke judgments in line with (J). It is helpful to frame two specific empirical claims one might make in this vein:

(1) Extracontextual assessors do not tend to judge that MIGHT is true.

(2) Extracontextual assessors tend to judge that MIGHT is false.

If (1) and (2) were correct, this would indeed support (J).

Consider next retraction cases. Here a scenario is constructed where a BEP is sensibly produced by some agent in a discourse, but at some later time, that agent acquires the information that the prejacent of the BEP is in fact false. The agent then responds by retracting his BEP — for example, by saying something to the effect

---

4 For another example, Hacking (1967) imagines a salvage crew looking for a treasure in a ship that sank long ago. A mate says "It is possible that we shall find the treasure here," but in the scenario it is stipulated (indirectly) that the treasure is not in the relevant bay. So the crew will not find it there. So the prejacent of the modal claim produced by the mate is false. The case thus fits our description of an extracontextual assessment case: the reader of this vignette is provided with the information that the prejacent of an imagined BEP is false, and we may then probe the intuition of the reader about the truth value of that BEP.

(Hacking also stipulates as part of his example that the unbeknownst to the mate, his own log data actually rules out the possibility that the ship had sank in the relevant bay. Hacking's main purpose with the example is not to develop a relativist view, but to defend the idea that BEPs may be truth-conditionally sensitive to evidence going beyond what is in fact known (by anyone). But a relativist might offer an alternative account of the putative judgments in this case. Roughly, she could hold that (i) in evaluating fictional utterances within a story, subjects pretend to be in a state of information that corresponds to the content of the story; and (ii) fictional utterances having relativist content will be evaluated by subjects for truth and falsity relative to this pretend state of information. Since in the story, it is given that the ship is not in the relevant bay, that information will be part of the pretend state of information that subjects will use to evaluate the BEP produced by the mate; and thus they will judge it false.)

More recent work in the relativist vein has emphasized extracontextual assessment cases which involve an explicit eavesdropper in the imagined scenario (thus "eavesdropping cases"). Readers of these vignettes are not directly asked for their truth value judgments about the hypothetical BEP; rather, the eavesdropper in the scenario is stipulated to vocalize a truth-value judgment about a BEP, and readers are then asked about their judgments about the eavesdropper's judgments. Below we experimentally evaluate both species of extracontextual assessment. Thus we are conceiving of the eavesdropping cases as a subclass of the extracontextual assessment cases.

that he was wrong in saying what he said. To give an example, take the FAT TONY vignette above and extend it with the following text:

> Shortly thereafter, new evidence comes to light, and everyone now agrees that Fat Tony is actually alive.
>
> Expert A then says, "I was wrong — Fat Tony was actually alive."
>
> Expert B also says, "I was wrong — Fat Tony was actually alive."

Call this extended vignette FAT TONY WITH RETRACTION. It is natural to say that Expert A is right to retract his earlier assertion. Do competent speakers of English think it is also right for Expert B to retract *his* earlier assertion (viz., MIGHT)? Some relativists have thought that this kind of retraction would indeed be proper (notably, MacFarlane (2011)). They have in effect hypothesized that given the extended vignette,

(3) Competent speakers tend to judge that it is appropriate for Expert *B* to retract MIGHT.

If (3) is correct, this would indeed lend support to (J), given the simple thought that, other things being equal, one ought to retract a statement one has recently made when it becomes clear that it is false.

The robustness of the judgments relativists have cited in support of their theories, in connection with these two kinds of cases, has been questioned (Hacquard 2006, Hawthorne 2007, Stephenson 2007a, Wright 2007, von Fintel & Gillies 2008, 2011, Dietz 2008, Yalcin 2008, 2011, Portner 2009, Willer 2010, 2013, Dowell 2011, Swanson 2011, Braun 2012, Yanovich 2014). What is reasonable to assume about the truth values of (1), (2), and (3)? More generally, how do the judgments in fact pattern in extracontextual assessment and retraction cases? There is no consensus in the literature on these matters. In particular, there is no consensus in the literature about whether (J) is correct. Correspondingly, there is no consensus about whether it should constrain theorizing about the semantics and pragmatics of epistemic modals.

To make further progress in this fragment of the literature, we require greater clarity on the data to be explained. Our objective in this note is not to develop a new theory that explains the patterns in speaker judgments about epistemic modal sentences, but rather to engage in the prior task of clarifying just what these patterns are. Below we present experimentally collected data for a number of extracontextual assessment and retraction cases. We bring these data to bear on (J), and the larger question of (R).

Our data focuses on *un*embedded occurrences of epistemic possibility modals, the locus of attention in this fragment of the literature. In this respect it complements

to some extent Hacquard & Wellwood 2012, a recent corpus-based study of the distribution of epistemics in embedded contexts.

Our main result can be summarized briefly: (J) is incorrect. Theorists who have supposed otherwise are, we suspect, tracking something theoretically interesting and important for our understanding of epistemic modality; it is just that what they are tracking is not the truth of (J). Our data hint that a certain related generalization, distinct from (J) but easily confused or conflated with it, may well be correct. We frame a candidate for this alternative generalization in our concluding section.

2 Experiment 1

In our first experiment, we examined people's intuitions about the truth value of the BEP produced by Expert B in the FAT TONY vignette, MIGHT. We wanted to know whether people would indeed agree with the claim that this sentence was false and disagree with the claim that the sentence was true, given that the prejacent of MIGHT is plainly false according to the vignette. In this experiment, we are in effect treating the experimental participant (the reader of the vignette) as the extracontexual assessor.

2.1 Methods

One hundred thirty participants completed a questionnaire online through Amazon Mechanical Turk (AMT). All participants were given the vignette about Fat Tony presented above. All participants were then asked whether they agreed or disagreed with a single statement. Each participant was randomly assigned to receive one of four different statements, as follows:

(NONMODAL-TRUE)	What Expert A said is true.
(NONMODAL-FALSE)	What Expert A said is false.
(MODAL-TRUE)	What Expert B said is true.
(MODAL-FALSE)	What Expert B said is false.

Each participant was asked to rate the relevant statement on a Likert scale from 1 ('completely disagree') to 7 ('completely agree').

2.2 Results and Discussion

Results for each condition are displayed in Figure 1.[5]

---

5 Within experimental linguistics, there is a certain amount of controversy as to whether Likert scale responses should be treated as continuous data or as ordinal data (see Schütze & Sprouse forthcoming).
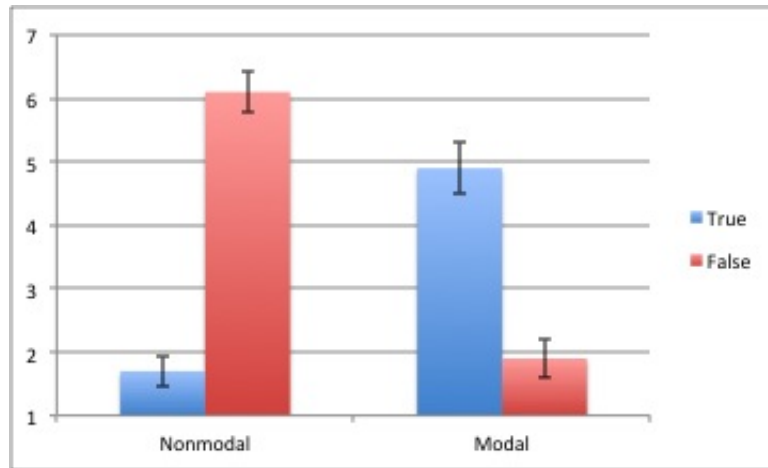
Figure 1. Mean ratings by condition for Experiment 1. Error bars show standard error of the mean.

In this particular case, (J) predicts that participants should give the same responses in the modal conditions that they do in the nonmodal conditions, and we therefore used the nonmodal conditions as a baseline for comparison. As the figure shows, participants' responses in the modal conditions were strikingly different from those in the nonmodal conditions.[6] Participants were far more inclined to agree with the claim that the modal statement was true than they were to agree that the nonmodal statement was true[7] and far less inclined to agree that the modal statement was false than they were to agree that the nonmodal statement was false.[8] Most importantly, within the modal conditions, participants were significantly more inclined to agree that the statement was true than they were to agree that it was false.[9]

In short, the pattern of people's judgments was almost exactly the opposite of

---

For completeness, we report means and standard errors in the figures accompanying each experiment, but we perform only nonparametric statistical analyses.

6 Before participating in the actual experiment, participants received a 'catch question' designed to make sure that they were paying careful attention. (The paragraph of instructions explained that they should not fill out the question written below but should instead simply write 'I have read the instructions.') In all, 151 participants were recruited for the study, but 21 failed this catch question. In addition, 2 participants reported that they were not native speakers of English and were therefore excluded from the analyses.

The data were analyzed using ordinal regression, with participants' responses as the dependent variable and statement type (1 for modal, 0 for nonmodal), truth value (1 for true, 0 for false) and the interaction as predictors. There was a significant effect of statement type (b = -3.0, SE = .54, $p < .001$), a significant effect of truth value (b = 4.1, SE = .62, $p < .001$) and a significant interaction (b = 6.5, SE = .86, $p < .001$).

7 Mann-Whitney U, $p < .001$

8 Mann-Whitney U, $p < .001$

9 Mann-Whitney U, $p < .001$

what one would predict given (J). Participants could see that the prejacent of the BEP was plainly false, but they were nonetheless significantly more inclined to say that the BEP itself was true than that it was false.

One might worry that the results of this first study are due to some experimental artifact involving either (a) the precise way in which the question was posed or (b) the idiosyncratic details of the Fat Tony vignette itself. We address these worries in the remaining studies. First, we introduce an alternative way of posing the question. Then, in Experiment 4, we shift to an entirely different vignette.

3 Experiment 2

Experiment 1 focuses on the experimental participants themselves as the extra-contextual assessors. But we might also focus on Fat Tony as the extracontextual assessor. In other words, rather than directly asking particpants for their judgments about MIGHT, we could instead have Fat Tony verbalize a truth-value judgment about MIGHT, and then ask participants to what extent they agree with what Fat Tony says. This kind of task is rather more complicated — in effect, it asks for a judgment about a judgment — but cases with this structure have nevertheless been used by relativists to motivate (J) (see, e.g., Egan et al. 2005, Egan 2007, MacFarlane 2011). Indeed, this paradigm is the standard kind of extracontextual assessment case in the literature. Experiment 2 is designed to discover the pattern in judgments in this kind of extracontextual assessment case.

3.1 Methods

One hundred twenty-three participants completed a questionnaire online using AMT. All participants received the vignette about Fat Tony used in the previous two studies. However, in this new study, an additional sentence was included at the end of the vignette. The sentence always took the form: "Watching this discussion on television, Fat Tony says to his henchmen, '(STATEMENT)'" The exact statement included in this sentence then varied by condition as follows:

(NONMODAL-TRUE)	"What Expert A said is true."
(NONMODAL-FALSE)	"What Expert A said is false."
(MODAL-TRUE)	"What Expert B said is true."
(MODAL-FALSE)	"What Expert B said is false."

All participants were then asked to indicate whether they agreed or disagreed with what Fat Tony had said on a scale from 1 ('completely disagree') to 7 ('completely agree').

3.2 Results and Discussion

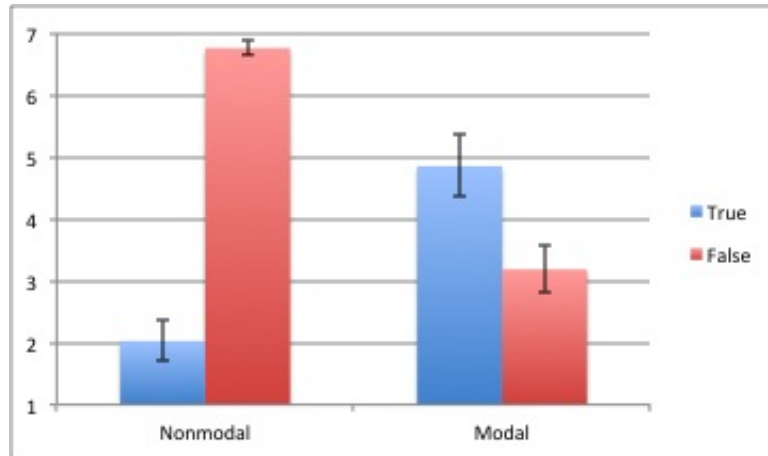Results for each condition are displayed in Figure 2.[10]

Figure 2. Mean ratings by condition for Experiment 2. Error bars show standard error of the mean.

As the figure shows, the results of this second experiment came closer to the pattern that would be expected given (J). Ratings for the claim that the modal was true were significantly lower than those provided for the obviously true statement in (NONMODAL-FALSE),[11] and ratings for the claim that the modal was false were significantly higher than those provided for the obviously false statement in (NONMODAL-TRUE).[12]

Yet, once again, the more salient result was participants' continuing tendency to go against the predictions of (J). Participants were far more inclined to say that the modal claim was true than they were to say that the nonmodal assertion was true[13] and far less inclined to say that the modal claim was false than they were to say that the nonmodal claim was false.[14] Most importantly, they were significantly more inclined to say that the modal claim was true than they were to say that it was

---

10 Three participants reported that they were not native speakers of English and were therefore excluded from the analyses. The data were analyzed using ordinal regression, with participants' responses as the dependent variable and statement type (1 for modal, 0 for nonmodal), truth value (1 for true, 0 for false) and the interaction as predictors. There was a significant effect of statement type (b = -2.3, SE = .56, $p < .001$), a significant effect of truth value (b = -1.4, SE = .50, $p < .001$) and a significant interaction (b = 6.2, SE = .90, $p < .001$).

11 Mann-Whitney U, $p < .001$

12 Mann-Whitney U, $p < .05$

13 Mann-Whitney U, $p < .001$

14 Mann-Whitney U, $p < .001$

false.[15]

In sum, the results of our first two experiments tell a fairly consistent story. In cases where (J) predicts that people will regard the statement MIGHT as false, experimental participants do show at least some reluctance to say that it is true and at least some inclination to say that it is false. Yet on the whole, participants are inclined to say that MIGHT is true and not false. Thus, the overall pattern of people's responses departs nontrivially from what would be expected given (J).

4 Experiment 3

In this next experiment, we looked at judgments about the appropriateness of retraction. That is, we looked at judgments about whether it would be right for a speaker who had produced a BEP to subsequently say "I was wrong" after learning that the prejacent of the BEP is false.

4.1 Methods

Eighty-two participants completed a questionnaire study online using AMT. All participants received the FAT TONY WITH RETRACTION vignette presented above. Thus participants were told that Expert A made an unmodalized assertion while Expert B used an epistemic modal, and then when further evidence became available, both experts said 'I was wrong.'

Participants were randomly assigned either to the nonmodal condition or to the modal condition. Participants were asked to evaluate the sentence:

Expert A [B] was right to say "I was wrong."

In other words, the difference between conditions was simply that participants in the nonmodal condition were asked about Expert A's retraction, while participants in the modal condition were asked about Expert B's retraction. Participants were asked to indicate whether they agreed with what Expert A [B] said on a scale from 1 ('completely disagree') to 7 ('completely agree').

4.2 Results

Results for each condition are displayed in Figure 3.

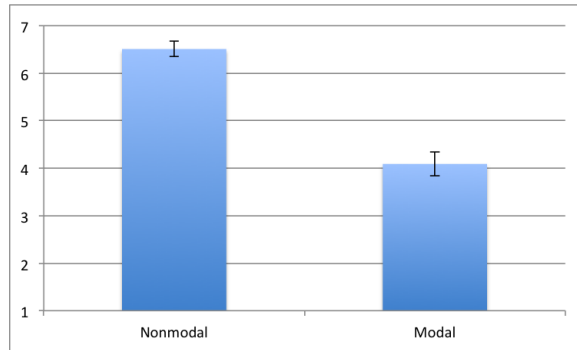
15 Mann-Whitney U, $p = .01$

Figure 3. Mean ratings by condition for Experiment 3. Error bars show standard error of the mean.

As the figure shows, participants in the nonmodal condition agreed that it was right for the speaker to retract, while participants in the modal condition gave the statement an intermediate rating. This rating was significantly lower than the one obtained for the nonmodal condition.[16]

4.3 Discussion

Participants did not strongly agree with the claim that it would be correct in this case for the speaker to retract, but at the very least, they did not specifically disagree with this claim. This result is the closest we have been able to get so far to the data that might be expected given (J).

The pattern observed here in people's judgments about retraction is certainly an interesting and important one. The key question now is whether it provides evidence for (J). We take up this question in the next experiment.

5 Experiment 4

In this final experiment, we further explore the issue of retraction. When one learns that people think that it would be appropriate for a speaker to retract an utterance, it is natural to conclude that people regard the content of that utterance as false. But perhaps this conclusion is too hasty. The relationship between judgments about retraction and judgments about falsity might, in fact, be considerably more complex. Our aim in this experiment is to more clearly distinguish these two sorts of judgments. Accordingly, participants were randomly assigned to be asked either (a) whether it would be appropriate for the speaker to retract or (b) whether the speaker's original assertion was false.

---

16 Three participants reported that they were not native speakers of English and were therefore excluded. Mann-Whitney U, $p < .001$.

In addition, we turned to a slightly different methodology. One might worry that there is something idiosyncratic about our Fat Tony vignette which has been distorting the results in the experiments reported thus far. We therefore thought it worthwhile to test an example offered by an advocate of relativism, rather than an example contrived by us. Specifically, we used an example given by MacFarlane (2003, 2011) in support of a relativist analysis.

5.1 Methods

One hundred fifty-nine participants completed a questionnaire study online using AMT. The experiment used a 2x2 design in which each participant was assigned to receive a particular statement (epistemic modal vs. nonmodal) and a particular question about that statement (falsity vs. retraction).

Participants in the epistemic modal condition received the following vignette (closely modeled on MacFarlane 2011):[17]

> 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 might be 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."

Participants in the nonmodal condition received a vignette that was exactly the same, except that Sally says "Joe is in Boston."

Participants in the retraction conditions then received the 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.

Participants in the falsity conditions received the question:

---

17 The vignette was slightly modified so as to make it clear that, at the time of the utterance, the relevant evidence was not within the speaker's 'epistemic reach' (see, e.g., Yanovich 2014). This issue is largely orthogonal to the one under investigation in the present paper.

> 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 is false.

All participants were asked to indicate their level of agreement on a scale from 1 ('completely disagree') to 7 ('completely agree').

5.2   Results and Discussion

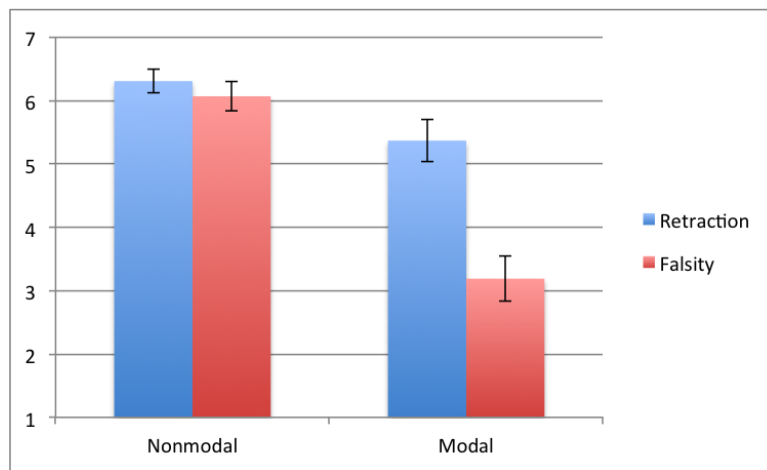Results for each condition are displayed in Figure 4.

Figure 4. Mean ratings by condition for Experiment 4. Error bars show standard error of the mean.

As the figure shows, the pattern of responses obtained for the modal statement was quite different from the pattern obtained for the nonmodal statement.[18] For judgments about the nonmodal statement, there was no significant difference between agreement with the claim that the assertion was false and agreement with the claim that the speaker should retract.[19] By contrast, for judgments in the modal conditions, there was a difference between responses to the two questions. Specifically, participants were more inclined to agree that the speaker should retract the BEP than they

---

18 Five participants reported that they were not native speakers of English and were therefore excluded from the analyses. The data were analyzed using ordinal regression, with participants' responses as the dependent variable and statement type (1 for modal, 0 for nonmodal), question type (1 for retraction, 0 for falsity) and the interaction as predictors. There was a significant effect of statement type (b = 0.9, SE = .44, $p < .05$), a significant effect of question type (b = -1.9, SE = .46, $p < .001$) and a significant interaction (b = 1.9, SE = .64, $p < .01$).

19 Mann-Whitney U, $p = .84$

were to agree that the BEP was false.[20] Moreover, agreement with the claim that the BEP was false was significantly lower than the baseline provided by agreement with the claim that the nonmodal assertion was false.[21]

In short, this final study showed a significant divergence between people's judgments about retraction and their judgments about falsity. We found strong agreement with the claim that it would be appropriate for the speaker to retract the BEP, but we did not find strong agreement with the claim that the BEP was false. In other words, there is clearly something that makes people think it would be appropriate for the speaker to retract, but whatever that something is, we cannot say in general that it owes to a judgment that the original BEP itself is false.

6 General Discussion

In four experiments, we examined judgments about cases in which a speaker utters a BEP whose prejacent is known to be false. When participants were given a case of this type, they disagreed with the claim that the BEP was false and actually tended to agree that it was true (Experiment 1). This finding arose even when we switched to an 'eavesdropper' case in which participants are asked to imagine an extracontextual assessor who has more evidence than the original speaker (Experiment 2). Participants did show a certain inclination to say that the speaker should retract a BEP when the prejacent is revealed to be false (Experiment 3), but when judgments about the appropriateness of retraction are disentangled from judgments about the falsity of the original statement, it becomes clear that whatever it is that makes participants regard retraction as appropriate, it is not a judgment that the original statement is false (Experiment 4). Taken together, these four experiments tell a consistent story. What they suggest is that (J) is mistaken.

In making this claim, we do not take ourselves to have refuted (R). Again, the dialectic is this. Relativists like Egan (2007) and MacFarlane (2011) have defended versions of (R). Their defenses have rested in nontrivial part on (J). Our data undermine (J). Thereby our data undermine attempts to support (R) from (J). We have not shown that (R) could not be defended in other ways. Perhaps it can be; see again Stephenson 2007a.

Egan (2007) and MacFarlane (2011) have, we think, brought to attention interesting and important facts about the language of epistemic modality. What we have critiqued is the idea that (J) accurately characterizes the relevant empirical facts. In seeking fruitful alternative descriptive generalizations, we might begin by looking again at the results of our last experiment. Here we observed that truth value

---

20 Mann-Whitney U, $p < .001$
21 Mann-Whitney U, $p < .001$

judgments and retraction judgments can come apart in the case of epistemic modal language. This surprising fact suggests that retraction is not — or not generally — a way of manifesting a view about the truth value of a claim. We might therefore seek some other kind of theoretical understanding of retraction. One possible approach would be to view retraction as a phenomenon whereby speakers are primarily indicating that they no longer want a conversational common ground incorporating the update associated with a sentence that they previously uttered. On this approach, *what is retracted* is a certain conversational update; retraction is in part a means of undoing or disowning the context change or update performed by a speech act. One attractive feature of this conception of retraction is that it is sufficiently general to apply to nonassertoric speech acts, like commands, questions, and expressives. There is an intuitive sense in which one can retract a speech act of any of these forms.

If we understood retraction along these general lines, it would make sense to shift some of our attention to a different question. Specifically, following a suggestion from Khoo (2014), we might seek to explain these data not in terms of the truth conditions of epistemic modal sentences but rather in terms of their conversational dynamics. What is the characteristic conversational effect or update of epistemic possibility sentences, such that it might make sense for speakers to disown or retract that update in the sort of context provided in experiment 4? There has, of course, been much theorizing about the conversational effect and dynamics of epistemic modal sentences (Groenendijk et al. 1996, Veltman 1996, Beaver 2001, Egan 2007, von Fintel & Gillies 2007, 2010, 2011, Yalcin 2007, 2011, 2012a,b, Willer 2010, 2013, Yablo 2011, among others). It would be natural to seek some illumination about retraction in this work.

In this connection, we note that a common theme in this work is the following rather natural idea: a conversation which incorporates the update associated with an epistemic possibility sentence is (*inter alia*) a conversation that can be felicitously updated with its prejacent. If so, perhaps retraction, in situations like the one provided by experiment 4, chiefly targets this dimension of the conversational effect of epistemic possibility sentences. Perhaps retraction is a way of flagging that one no longer wishes to be in a conversational context having this property. For example, in Experiment 4, participants say that the speaker should retract his assertion of "Joe might be in Boston". The hypothesis under consideration here would be that participants are not thereby indicating that the speaker should maintain that his original assertion was false; rather, they are indicating that he should flag that he no longer wishes to be in a particular kind of conversational context (namely, a context that can be felicitously updated with "Joe is in Boston").

This suggestion requires further investigation. We offer it as a possible direction for further inquiry. And we note that even if existing accounts of the dynamics of

epistemic modal sentences can make sense of the retraction data along these lines, it would remain to be seen whether such accounts can also predict the distribution of truth value judgments observed in the experiments reported above.

References

Beaver, David. 2001. *Presupposition and assertion in dynamic semantics*. Stanford, CA: CSLI Publications.

Braun, David. 2012. An invariantist theory of 'might' might be right. *Linguistics and Philosophy* 35(6). 1–29. http://dx.doi.org/10.1007/s10988-012-9124-y.

Dietz, Richard. 2008. Epistemic modals and correct disagreement. In Manuel Garcia-Carpintero & Max Kolbel (eds.), *Relative truth*, 225–38. Oxford: Oxford University Press.

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

Egan, Andy. 2007. Epistemic modals, relativism and assertion. *Philosophical Studies* 133(1). 1–22.

Egan, Andy. 2011. Relativism about epistemic modals. In Stephen D. Hales (ed.), *A companion to relativism*, 219–241. Hoboken, NJ: Wiley-Blackwell.

Egan, Andy, John Hawthorne & Brian Weatherson. 2005. Epistemic modals in context. In G. Preyer & G. Peter (eds.), *Contextualism in philosophy*, 131–169. Oxford: Oxford University Press.

von Fintel, Kai & Anthony Gillies. 2007. An opinionated guide to epistemic modality. *Oxford Studies in Epistemology* 2. 32–63.

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

von Fintel, Kai & Anthony Gillies. 2010. Must... stay... strong! *Natural Language Semantics* 18(4). 351–383. http://dx.doi.org/10.1007/s11050-010-9058-2.

von Fintel, Kai & Anthony Gillies. 2011. *Might* made right. In Andy Egan & Brian Weatherson (eds.), *Epistemic modality*, 108–30. Oxford: Oxford University Press.

Groenendijk, Jeroen, Martin Stokhof & Frank Veltman. 1996. Coreference and modality. In Shalom Lappin (ed.), *Handbook of contemporary semantic theory*, 179–213. Oxford: Blackwell. http://dx.doi.org/10.1111/b.9780631207498.1997.00010.x.

Hacking, Ian. 1967. Possibility. *Philosophical Review* 76(2). 143–68.

Hacquard, Valentine. 2006. *Aspects of modality*. Cambridge, MA: Massachusetts Institute of Technology dissertation.

Hacquard, Valentine & Alexis Wellwood. 2012. Embedding epistemic modals in

English: A corpus-based study. *Semantics and Pragmatics* 5(4). 1–29. http://dx.doi.org/10.3765/sp.5.4.

Hawthorne, John. 2004. *Knowledge and lotteries*. Oxford: Oxford University Press.

Hawthorne, John. 2007. Eavesdroppers and epistemic modals. *Philosophical Issues* 17(1). 92–101. http://dx.doi.org/10.1111/j.1533-6077.2007.00124.x.

Khoo, Justin. 2014. Modal disagreements. Unpublished manuscript, Massachusetts Institute of Technology. http://web.mit.edu/jkhoo/www/MD4.pdf.

MacFarlane, John. 2003. Epistemic modalities and relative truth. Unpublished manuscript, Berkeley. http://johnmacfarlane.net/epistmod-2003.pdf.

MacFarlane, John. 2011. Epistemic modals are assessment-sensitive. In Andy Egan & Brian Weatherson (eds.), *Epistemic modality*, 144–78. Oxford: Oxford University Press.

Portner, Paul. 2009. *Modality*. Oxford: Oxford University Press.

Schütze, Carson T. & Jon Sprouse. forthcoming. Judgment data. In Devyani Sharma & Rob Podesva (eds.), *Research Methods in Linguistics*, Cambridge, UK: Cambridge University Press.

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

Stephenson, Tamina C. 2007b. *Towards a theory of subjective meaning*. Cambridge, MA: Massachusetts Institute of Technology dissertation.

Swanson, Eric. 2011. How not to theorize about the language of subjective uncertainty. In Andy Egan & Brian Weatherson (eds.), *Epistemic modality*, 249–69. Oxford: Oxford University Press.

Veltman, Frank. 1996. Defaults in update semantics. *Journal of Philosophical Logic* 25(3). 221–61. http://dx.doi.org/10.1007/BF00248150.

Weatherson, Brian & Andy Egan. 2011. Epistemic modals and epistemic modality. In Andy Egan & Brian Weatherson (eds.), *Epistemic modality*, 1–18. Oxford: Oxford University Press.

Willer, Malte. 2010. *Modality in flux*: University of Texas dissertation.

Willer, Malte. 2013. Dynamics of epistemic modality. *Philosophical Review* 122(1). 45–92. http://dx.doi.org/10.1215/00318108-1728714.

Wright, Crispin. 2007. New age relativism and epistemic possibility: The question of evidence. *Philosophical Issues* 17(1). 262–283. http://dx.doi.org/10.1111/j.1533-6077.2007.00133.x.

Yablo, Stephen. 2011. A problem about permission and possibility. In Andy Egan & Brian Weatherson (eds.), *Epistemic modality*, 270–94. Oxford: Oxford University Press.

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

Yalcin, Seth. 2008. *Modality and inquiry*. Cambridge, MA: Massachusetts Institute of Technology dissertation.

Yalcin, Seth. 2011. Nonfactualism about epistemic modality. In Andy Egan & Brian Weatherson (eds.), *Epistemic modality*, 295–332. Oxford: Oxford University Press.

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

Yalcin, Seth. 2012b. Context probabilism. *Eighteenth Amsterdam Colloquium* 12–21. http://dx.doi.org/10.1007/978-3-642-31482-7_2.

Yalcin, Seth. 2014 forthcoming. Semantics and metasemantics in the context of generative grammar. In Alexis Burgess & Brett Sherman (eds.), *Metasemantics*, Oxford: Oxford University Press.

Yanovich, Igor. 2014. Standard contextualism strikes back. *Journal of Semantics* 31(1). 67–144. http://dx.doi.org/10.1093/jos/ffs022.