Erkenn
DOI 10.1007/s10670-017-9929-8

ORIGINAL RESEARCH

Epistemic Modals and Alternative Possibilities

John Turri[1]

Received: 30 August 2016 / Accepted: 19 July 2017

Abstract Indicative judgments pertain to what is true. Epistemic modal judgments pertain to what must or might be true relative to a body of information. A standard view is that epistemic modals implicitly quantify over alternative possibilities, or ways things could turn out. On this view, a proposition *must* be true just in case it is true in *all* the possibilities consistent with the available information, and a proposition *might* be true just in case it is true in *at least one* possibility consistent with the available information. I report three experiments testing this view of epistemic modals. The results show that although modal judgments are sensitive to information about alternative possibilities, the standard quantification theory mischaracterizes the ordinary meaning of modals. I then report two more experiments testing the hypothesis that epistemic modals express willingness to attribute knowledge based on the available information. The results support this hypothesis. The results also show that the difference between "inside" and "outside" probabilistic information, familiar from the judgment and decision-making literature, affects epistemic modal judgments.

1 Introduction

Mainstream news organizations and history textbooks teach us that James Earl Ray assassinated Martin Luther King Jr. In light of this, most of us would agree that Ray killed King. But would we agree that Ray *must* have killed King? Would we deny that Ray *might not* have killed King?

✉ John Turri
john.turri@gmail.com

[1] Philosophy Department and Cognitive Science Program, University of Waterloo, 200 University Avenue West, Waterloo, ON N2L3G1, Canada

Published online: 11 August 2017

Indicative judgments pertain to what is or is not true (e.g. "Ray killed King," "Sirhan did not kill King"), whereas modal judgments pertain to the way (the "mode") in which a proposition is true. Epistemic modals pertain to what *must* or *might* be true relative to a body of information (e.g. "Ray must/might not have killed King"). Epistemic modals are common in ordinary speech and research has shown that their use has social consequences. For example, their use affects decisions about whom to trust in both children and adults (Moore et al. 1990; Furrow and Moore 1990). Some experimental work has investigated the cognitive processes involved in modal reasoning (Bell and Johnson-Laird 1998), but this work has not focused on the cognitive processes underlying "must" and "might" modal judgments specifically. Linguists and philosophers have proposed formal semantic models for epistemic modals, and these have been motivated by researchers' linguistic intuitions and social observations about how others tend to use and interpret modal claims (e.g. Papafragou 2006; von Fintel and Gillies 2010; Dowell 2011; this fits the standard methodology in contemporary analytic philosophy of language, as found in, for instance, Austin 1962; Kripke 1972; Unger 1984; see Machery and Stich 2012 for critical discussion and insightful comparisons to methodological developments in linguistics). These models can provide a starting point for investigating the ordinary meaning and function of epistemic modals and the cognitive processes underlying modal judgments (compare Knobe and Yalcin 2014; Khoo 2014).

A standard view in linguistics, philosophy, and psychology is that in addition to being sensitive to information about how things actually are, epistemic modal judgments are sensitive to information about alternative possibilities in a way that indicative judgments are not. For example, the truth of a proposition, such as "Ray killed King," depends only on the way the world actually is. By contrast, the truth of related modal propositions, such as "Ray must have killed King" or "Ray might not have killed King," depends partly on alternative ways the world could be, consistent with the available information. Call this *the weak view* of epistemic modals (hereafter, I will typically refer to epistemic modals simply as "modals").

Many theorists have taken the weak view as their starting point in developing more detailed and stronger views. On one very popular approach, modal judgments implicitly quantify over possibilities in two very specific ways (Lewis 1979; von Fintel and Gillies 2007, 2008, 2010; Stephenson 2007; Yalcin 2007; Egan 2007; MacFarlane 2011; Schaffer 2011; Knobe and Szabó 2013). (Kratzer 1977 is often cited as holding this sort of view, but this is arguably an oversimplification). On the one hand, a proposition *must* be true just in case it is true in *all* the possibilities consistent with the available information. On the other hand, a proposition *might* be true just in case it is true in *at least one* possibility consistent with the available information. Thus, on this approach, "must" is likened to a universal quantifier and "might" is likened to an existential quantifier (compare Bell and Johnson-Laird 1998). Call this *the strong quantificational view* of modals.

A quantificational view typically involves two other important details. First, while modals quantify over a set of possibilities, it is not always the same set. The set can change across conversational contexts. Second, even within a single conversational context, some acknowledged possibilities can be *irrelevant*. For

example, although it is possible that members of an advanced alien species framed Ray for King's death, we treat this possibility as irrelevant and do not take it seriously. With these two details in mind, the most popular quantificational view is that "must" universally quantifies over the set of all relevant possibilities in a given context, and "might" existentially quantifies over that same set.

These views have consequences for how people will use epistemic modals. The weak view predicts that information about alternative possibilities will affect modal judgments more than it affects indicative judgments. The strong view makes additional, stronger predictions. The strong view predicts that if it is a relevant possibility that a proposition is false, then people will deny that it must be true, and people will agree that it might be false. In this paper, I report three experiments designed to test these predictions. To anticipate the findings, the results support the weak view but undermine the strong view. In order to explain this pattern of results, I propose that epistemic modals behave as they do because of their relationship to knowledge. In particular, I hypothesize that epistemic modals express willingness to attribute knowledge based on the available information. Two final experiments provide some initial evidence for this hypothesis.

2 Experiment 1: In and Out

2.1 Method

2.1.1 Participants

Three hundred fifty people participated (aged 18–71 years, mean age = 33 years; 139 female; 94% reporting English as a native language). Participants were U.S. residents, recruited and tested online using Amazon Mechanical Turk and Qualtrics, and compensated $0.40 for approximately 2–3 min of their time. Repeat participation was prevented.

2.1.2 Materials and Procedure

Participants were randomly assigned to one of ten conditions in a 2 (probability: outside, inside) × 5 (Probe: indicative, must, might not, necessity, possibly not) between-subjects design. Each participant read a single story, responded to test statements, then filled out a brief demographic survey.

The probability factor varied whether relevant probabilistic information pertained to a distribution or base rate (outside), or to a propensity of a specific item in the scenario (inside). Prior research shows that the outside/inside difference affects performance on a range of tasks, including planning, decision-making and social evaluations (Kahneman and Tversky 1982; Wells 1992; Lagnado and Sloman 2004; Friedman and Turri 2014; Turri et al. 2017; Turri ms.), so I used it to test modal judgments too. The basic story concerned Seth, who just turned in his final paper for a university course. In the outside version of the story, 95% of all university students

plagiarize their final paper. In the inside version, Seth's paper is 95% likely to be plagiarized. Here is the story (outside/inside manipulation in brackets):

> Seth recently handed in his final paper for a university course. According to a computerized analysis of papers handed in this semester, [95% of all university students plagiarize their final paper/Seth's final paper is 95% likely to be plagiarized].

The Probe factor varied which test statement participants responded to. Each participant rated their agreement with one of these statements:

(Indicative) Seth plagiarized his paper.
(Must) Seth must have plagiarized his paper.
(Might not) Seth might not have plagiarized his paper.
(Necessity) It's necessary that Seth plagiarized his paper.
(Possibly not) It's possible that Seth did not plagiarize his paper.

I included the indicative and alethic modal probes (i.e. "necessity" and "possibly not") as potentially useful points of comparison for interpreting the results of the epistemic modal probes (i.e. "must" and "might not"). Responses were collected on a standard Likert scale, 1 ("strongly disagree")–7 ("strongly agree"), left-to-right on the participant's screen.

After responding to the initial test statement, participants went to a new screen and completed a percentage task. All participants completed this same task and were instructed, "Your answer to the following two questions must sum to 100%."

How probable is it that Seth plagiarized his paper?
How probable is it that Seth did not plagiarize his paper?

The order of these two questions was randomly rotated. For each question, participants entered a numerical response into a text box. They could see the sum of their answers automatically tallied immediately below the two text boxes. Participants could not proceed unless the answers equaled 100%.

After completing the percentage task, participants went to a new screen and responded to a relevance statement:

It's a relevant possibility that Seth did not plagiarize his paper.

Responses were collected on the same 7-point Likert scale described above. All participants responded to this same statement.

2.2 Results

Response to all dependent measures was unaffected by participant gender, age, or socioeconomic status. The same is true for the other experiment reported below. These demographic variables will not be discussed further.

Response to the initial test statement was affected by Probe, $F(4, 340) = 17.61$, $p < .001$, $\eta_p^2 = .172$, marginally affected by probability, $F(1, 340) = 3.05$, $p = .082$, $\eta_p^2 = .009$, and affected by their interaction, $F(4, 340) = 11.50$, $p < .001$, $\eta_p^2 = .119$ (see Fig. 1a). Examining responses for each probe separately,

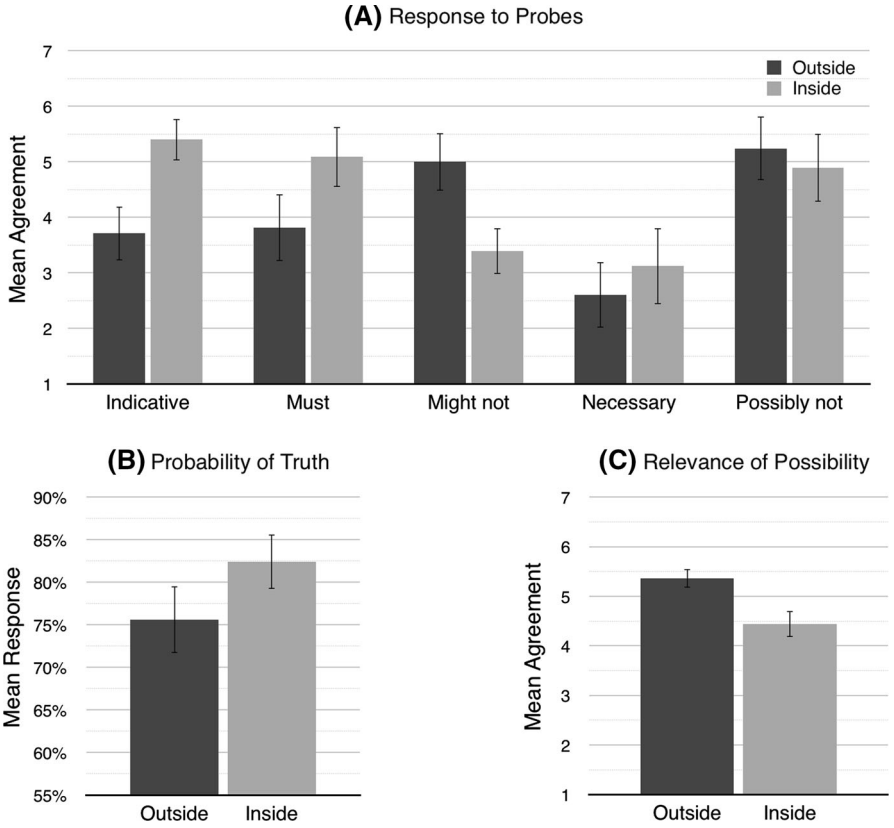

Fig. 1 Experiment 1. **a** Mean agreement to the five probes (between-subjects) in each probability condition (outside/inside); the scale ran 1 (SD)–7 (SA). **b** Mean rating of the probability that the proposition was true (collapsed across probe); the scale ran 0–100%. **c** Mean agreement that the proposition's being false is a relevant possibility (collapsed across probe); the scale ran 1 (SD)–7 (SA). *Error bars* represent 95% confidence intervals

we find that agreement was higher in the inside condition for the indicative probe (M = 5.40/3.71, SD = 1.06/1.41), t (68) = 5.66, $p < .001$, d = 1.38, and for the must probe (M = 5.09/3.81, SD = 1.56/1.79), t(68) = 3.19, $p = .002$, d = 0.77. By contrast, agreement was lower in the inside condition for the might-not probe (M = 3.39/5.00, SD = 1.80/1.63), t(69) = −4.15, $p < .001$, d = 1.00. Response to the necessity probe did not differ between inside and outside conditions (M = 3.12/2.60, SD = 1.97/1.72), t(67) = 1.12, $p = .248$. Similarly, response to the possibly-not probe did not differ between inside and outside conditions (M = 4.89/5.24, SD = 1.80/1.63), t(68) = −0.84, $p = .403$.

Response to the indicative probe was above midpoint in the inside condition, t(34) = 7.79, $p < .001$, and no different from midpoint in the outside condition, t(34) = −1.20, $p = .237$. Response to the must probe was above midpoint in the inside condition, t(33) = 4.06, $p < .001$, and no different from response to the indicative probe in the inside condition, t(67) = −0.97, $p = .335$. Response to the

must probe in the outside condition did not differ from midpoint, $t(35) = -0.65$, $p = .518$, or from response to the indicative probe in the outside condition, $t(66.15) = 0.24$, $p = .812$. Response to the might-not probe was significantly below midpoint in the inside condition, $t(35) = -2.10$, $p = .043$, and it was significantly above midpoint in the outside condition. Response to the necessity probe was significantly below the midpoint for both inside and outside conditions: inside, $t(33) = -2.62$, $p = .013$; outside, $t(34) = -4.81$, $p < .001$. Response to the possibly-not probe was significantly above the midpoint for both conditions: inside, $t(35) = 2.96$, $p = .005$; outside, $t(33) = 4.41$, $p < .001$.

Performance on the percentage task was affected by Probability, $F(1, 340) = 6.93$, $p = .009$, $\eta_p^2 = .020$, but not by Probe or their interaction. Participants rated the probability that Seth plagiarized the paper slightly lower in outside conditions ($M = 75.6\%$, $SD = 26.3$) than in inside conditions ($M = 82.4\%$, $SD = 21.8$), $t(336.1)$, $p = .008$, $d = 0.29$ (see Fig. 1b).

Similarly, response to the relevance statement—whether it was a relevant possibility that Seth did not plagiarize his paper—was affected by Probability, $F(1, 340) = 30.43$, $p < .001$, $\eta_p^2 = .082$, but not by Probe or their interaction. Participants were less likely to rate the possibility as relevant in inside conditions ($M = 4.44$, $SD = 1.76$) than in outside conditions ($M = 5.36$, $SD = 1.29$), $t(318.4) = 5.57$, $p < .001$, $d = 0.62$ (see Fig. 1c). Nevertheless, participants tended to agree that the possibility was relevant in both conditions: inside, $t(174) = 13.98$, $p < .001$; outside, $t(174) = 3.30$, $p < .001$.

To test the potential mediating role of relevant alternative possibilities and the perceived chance of error on modal judgments, I conducted bootstrap multiple-mediators analyses (Hayes 2013) for the must probe and might-not probe. For the first analysis, I used assignment to Probability condition as the independent variable (coded: 0 = outside, 1 = inside), response to the must probe as the outcome, and responses to the percentage task and relevance statement as potential mediators. This analysis showed that response to the relevance statement mediated the effect of condition on response to the must probe, 95% CI (i.e. confidence interval) for the indirect effect = [0.21, 0.98]. By contrast, response to the percentage task did not mediate the effect of condition, 95% CI for the indirect effect = [−0.14, 0.35] (see Fig. 2a). The second analysis was exactly the same as the first except that response to the might-not probe was the outcome. The analysis showed that response to the relevance statement mediated the effect of condition on response to the might-not probe, 95% CI for the indirect effect = [−1.10, −0.22]. By contrast, response to the percentage task did not mediate the effect of condition, 95% CI for the indirect effect = [−0.22, 0.22] (see Fig. 2b).

As a point of comparison, I also ran the same bootstrap multiple-mediators analysis with response to the indicative probe as outcome. The analysis did not find that response to the relevance statement mediated the effect of condition on response to the indicative probe, 95% CI for the indirect effect = [−0.01, 0.46], and the same was true for response to the percentage task, 95% CI for the indirect effect = [−0.03, 0.40] (see Fig. 2c). These null findings should not be interpreted too strongly, however, because the 95% confidence intervals extend only slightly past zero. Other things being equal, a larger sample size could have produced

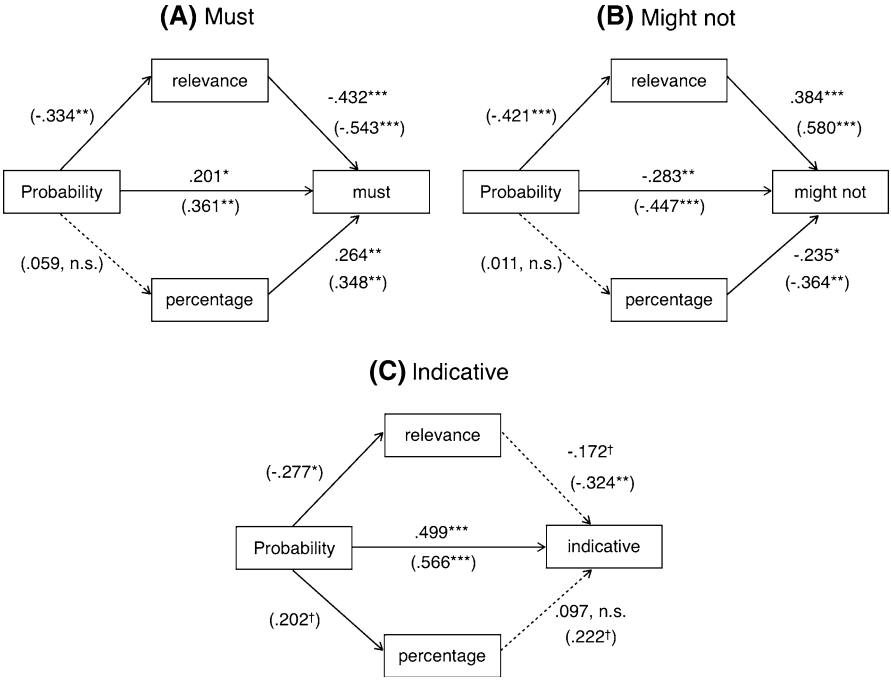

Fig. 2 Experiment 1. Mediation results for the two modal probes (must/might not) and the indicative probe. *Parenthetical values* represent the strength of a simple regression between the two variables; *values outside the parentheses* represent the strength of the relationships in a model used to test for mediation. $^{\dagger}p \leq .10$; $*p \leq .05$; $**p \leq .01$; $***p \leq .001$

confidence intervals entirely above zero. So although the current evidence suggests that modals and indicatives relate differently to judgments of relevance or percentage chances, the evidence does not show that the relationships are radically different. Instead, the difference might be one of degree.

2.3 Discussion

Three main findings emerge from this experiment. First, information about alternative possibilities affected modal judgments more strongly than it affected indicative judgments. This supports the weak view of epistemic modals, according to which epistemic modals are sensitive to information about alternative possibilities in a way that indicative judgments are not. Second, despite acknowledging, both quantitively and qualitatively, the possibility that a proposition is false, people were still willing to agree that the proposition must be true and deny that it might be false. This undermines the strong quantificational view of epistemic modals, according to which a proposition *must* be true just in case it is true in *all* the possibilities consistent with the available information, and a proposition *might* be true just in case it is true in at least one such possibility. Third, the outside/inside probability distinction significantly affected epistemic modal judgments.

It is worth emphasizing that people treated alethic modals (about what is necessary or possible) in roughly the way that the strong quantificational view predicted they would treat epistemic modals (see Fig. 1, Panel A, the "necessary" and "possibly not" probes). That is, people tended to disagree that a contingent outcome was necessary, and they tended to agree that it was possible for a contingent outcome to not occur. But people were randomly assigned to rate either an alethic or epistemic modal, and the procedures and stimuli were otherwise similar across conditions. Thus nothing in the procedures or scenario tested prevented people from responding in a way consistent with the strong view's predictions, and the observed differences between alethic and epistemic modal judgments are due differences in the probes used to measure them. (Interestingly, and perhaps relatedly, the outside/inside probability distinction did not significantly affect alethic modal judgments, although the observed numerical differences were in the same direction as for epistemic modal judgments, so perhaps there is a small outside/inside effect on alethic modal judgments that requires (much) more statistical power to detect.)

It might be suggested that the results unfavorable to the strong view are an artifact of the stimuli or procedures. On the one hand, perhaps the results were driven by something about the theme of plagiarism. Experiment 2 addresses this concern by using a completely different cover story and manipulation to test people's judgments. On the other hand, perhaps there is something odd about asking people to respond to modal statements in isolation and on a scale. Experiment 3 addresses this concern by using dichotomous response options allowing people to select between "must" and "might" modals to describe a situation.

3 Experiment 2: Up and Down

This experiment tests modal judgments using a different scenario and manipulation to see whether we continue observing results that undermine the strong view.

3.1 Method

3.1.1 Participants

Six hundred thirty-one new people participated (aged 18–76 years, mean age = 32 years; 247 female; 93% reporting English as a native language). Participants were recruited and compensated the same way as in Experiment 1. Repeat participation was prevented.

3.1.2 Materials and Procedure

Participants were randomly assigned to one of nine conditions in a 3 (probe: indicative, must, might not) × 3 (percent: 95 percent, 80 percent, 65 percent) between-subjects design. Each participant read a single story, responded to test statements, then filled out a brief demographic survey. The basic story concerned Tracy, who just purchased a new smartphone. Tracy's uncle, a smartphone engineer,

conducts a test and the result indicates a certain percentage chance that Tracy's smartphone has spyware installed on it. The Percent factor manipulated the percentage chance. Here is the story (percent manipulation in brackets):

> Tracy just bought a new smartphone. Her uncle is a smartphone engineer. He conducted a test and discovered that Tracy's smartphone is [95/80/65]% likely to have spyware installed.

The Probe factor varied which test statement participants responded to initially. Each participant rated their agreement to one of these statements:

> (Indicative) Tracy's smartphone has spyware.
> (Must) Tracy's smartphone must have spyware.
> (Might not) Tracy's smartphone might not have spyware.

Responses were collected on the same 7-point Likert scale used in Experiment 1.

After responding to the initial test statement, participants went to a new screen and completed a percentage task similar to the one used in Experiment 1, regarding these two statements:

> How probable is it that Tracy's smartphone has spyware?
> How probable is it that Tracy's smartphone does not have spyware?

After completing the percentage task, participants went to a new screen and responded to a relevance statement:

> It's a relevant possibility that Tracy's smartphone does not have spyware.

Responses were collected on the same 7-point Likert scale described above.

3.2 Results

Response to the initial test statement was affected by Probe $F(2, 622) = 21.0$, $p < .001$ $\eta_p^2 = .063$, marginally affected by Percent, $F(2, 622) = 2.89$, $p = .056$, $\eta_p^2 = .009$, and affected by their interaction, $F(4, 622) = 5.60$, $p < .001$, $\eta_p^2 = .035$ (see Fig. 3a).

Conducting pairwise comparisons between percent conditions within each probe, we find that agreement with the indicative probe was higher in the 95 percent condition ($M = 5.36$, $SD = 1.25$) than in the 65 percent condition ($M = 4.75$, $SD = 1.07$), $t(138) = 3.14$, $p = .002$, $d = 0.54$, but it did not differ between the 95 percent condition and the 80 percent condition ($M = 5.03$, $SD = 1.19$), $t(137) = 1.61$, $p = .109$, or between the 80 percent condition and the 65 percent condition, $t(139) = 1.48$, $p = .140$. Agreement with the must probe was higher in the 95 percent condition ($M = 5.25$, $SD = 1.42$) than in the 80 percent condition ($M = 4.64$, $SD = 1.29$), $t(137) = 2.63$, $p = .010$, $d = 0.45$, higher in the 95 percent condition than in the 65 percent condition ($M = 4.39$, $SD = 1.24$), $t(136) = 3.77$, $p < .001$, $d = 0.65$, but it did not differ between the 80 percent condition and the 65 percent condition, $t(137) = 1.17$, $p = .242$. Agreement with the might-not probe did not differ between the 95 percent condition ($M = 3.93$, $SD = 2.00$) and the 80 percent condition ($M = 3.91$, $SD = 1.78$), $t(140) = .051$, $p = .959$, but it was lower in the 90 percent condition than in the 65 percent

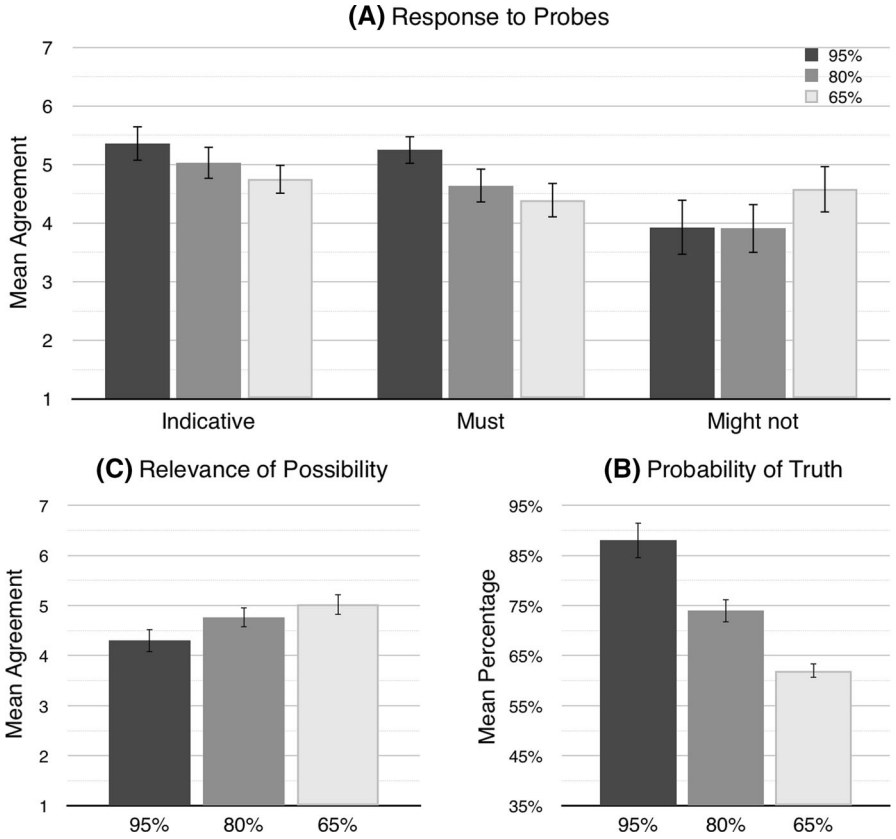

Fig. 3 Experiment 2. **a** Mean agreement to the three probes (between-subjects) in each percentage condition (95/80/65 percent); the scale ran 1 (SD)–7 (SA). **b** Mean rating of the probability that the proposition was true (collapsed across probe); the scale ran 0–100%. **c** Mean agreement that the proposition's being false is a relevant possibility (collapsed across probe); the scale ran 1 (SD)–7 (SA). *Error bars* represent 95% confidence intervals

condition, $t(137.7) = -2.09$, $p = .038$, $d = 0.36$, and lower in the 80 percent condition than in the 65 percent condition, $t(139) = -2.27$, $p = .025$, $d = 0.39$.

Response to the indicative probe was above midpoint in all three percent conditions: 95 percent, $t(68) = 9.07$, $p < .001$; 80 percent, $t(69) = 7.22$, $p < .001$; 65 percent, $t(70) = 5.91$, $p < .001$. Similarly, response to the must probe was above midpoint in all 3 percent conditions: 95 percent, $t(68) = 7.30$, $p < .001$; 80 percent, $t(68) = 4.18$, $p < .001$; 65 percent, $t(68) = 2.62$, $p = .011$. Response to the might-not probe did not differ from midpoint in either the 95% condition, $t(71) = -0.29$, $p = .769$, or the 80 percent condition, $t(69) = -0.40$, $p = .689$, but it was above the midpoint in the 65 percent condition, $t(70) = 2.88$, $p = .005$.

Performance on the percentage task was affected by Percent, $F(2, 622) = 129.50$, $p < .001$, $\eta_p^2 = .294$, but not by Probe or their interaction. Participants rated the probability that the phone had spyware installed higher in the 95 percent condition ($M = 88.0$ percent, $SD = 19.3$) than in the 80 percent condition ($M = 74.3\%$,

SD = 16.5), t(418) = 7.79, $p < .001$, d = 0.76, higher in the 95 percent condition than in the 65 percent condition, t(311.8) = 16.34, $p < .001$, d = 1.87, and higher in the 80 percent condition than in the 65 percent condition (M = 63.2%, SD = 9.9), t(341.8) = 8.34, $p < .001$, d = 0.90.

Similarly, response to the relevance statement—whether it is a relevant possibility that the smartphone does not have spyware installed—was affected by Percent, F(2, 622) = 12.01, $p < .001$, $\eta_p^2 = .037$, but not by Probe or their interaction. Participants were less likely to rate the possibility as relevant in the 95 percent condition (M = 4.30, SD = 1.73) than in both the 80 percent condition (M = 4.76, SD = 1.47), t(407.7) = −2.95, $p = .003$, d = 0.29, and the 65 percent condition (M = 5.02, SD = 1.36), t(396.6) = 4.77, $p < .001$, d = 0.48. Participants were also marginally less likely to rate the possibility as relevant in the 80 percent condition than in the 65 percent condition, t(419) = −1.89, $p = .059$, d = 0.19. Participants tended to agree that the possibility was relevant in all 3 percent conditions: 95 percent, t(209) = 2.48, $p = .014$; 80 percent, t(209) = 7.46, $p < .001$; 65 percent, t(210) = 10.86, $p < .001$.

3.3 Discussion

The results from this experiment further undermine the strong quantificational view of epistemic modals. Replicating a main finding from Experiment 1, despite acknowledging the possibility that a proposition is false, people still agreed that it must be true. As in Experiment 1, people rated the chance of a proposition's truth at about 80% and still judged that it must be true. In this experiment, people were willing to do the same thing when rating the chance of truth at about 60%. Moreover, despite rating the chance of truth at 80 or 70%, people did not agree that the proposition might be false. It was not until the chance of truth reached around 60% that people tended to agree that the proposition might be false. Interestingly, when the chance of truth reached around 60%, people tended to agree that the proposition must be true, and that it might be false (between-subjects).

4 Experiment 3: Might or Must

One potential worry about Experiment 1 is that it might be odd to ask people to respond to modal statements in isolation and on a Likert scale. Instead, it might be more natural to think of "must" and "might" modals as either/or possibilities to be considered together: either something must be true or it might false. Up till now, participants have not considered these options in tandem. This experiment addresses this concern by using a dichotomous probe allowing people to select between "must" and "might" to describe a situation.

4.1 Method

4.1.1 Participants

Two hundred new people participated (aged 19–67 years, mean age = 33 years; 92 female; 94% reporting English as a native language). Participants were recruited

and compensated the same way as in earlier experiments. Repeat participation was prevented.

4.1.2 Materials and Procedure

Participants were randomly assigned to one of four conditions in a 2 (Probability: inside, outside) × 2 (Probe: modal, indicative) between-subjects design. Each participant read a single story, responded to test statements, then filled out a brief demographic survey. The story and Probability manipulation were the same as in Experiment 1. The Probe factor manipulated whether participants initially responded to an indicative or modal probe. Participants chose an affirmative ("did" or "must") or negative ("did not" or "might not") response to best describe the case.

(Indicative) Seth _____ plagiarize his paper. [did/did not]
(Modal) Seth _____ have plagiarized his paper. [must/might not]

After responding to the initial probe, participants went to a new screen and responded to the same relevance statement as in Experiment 1, "It's a relevant possibility that Seth did not plagiarize his paper." Responses were collected on the same 7-point Likert scale as earlier.

4.2 Results

Response to the relevance statement was affected by Probability, $F(1, 196) = 11.83$, $p < .001$, $\eta_p^2 = .057$, but not by Probe or their interaction (see Fig. 4). In response to the indicative probe, 90% of participants answered affirmatively ("did"), compared to 80% in the outside probability condition. This difference was not significant, $\chi^2(1, 101) = 1.16$, $p = .281$. In response to the modal probe, 54% of

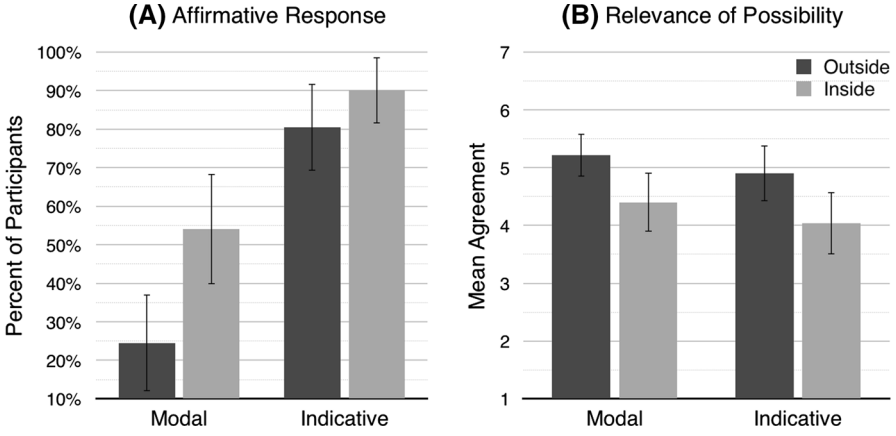

Fig. 4 Experiment 3. **a** Percentage answering affirmatively in the four conditions ("must have" for modal probes; "did" for indicative probes). **b** Mean agreement that the proposition's being false is a relevant possibility; the scale ran 1 (SD)–7 (SA). *Error bars* represent 95% confidence intervals

participants answered affirmatively ("must") in the inside probability condition, compared to 24% in the outside probability condition. This difference was significant, $\chi^2(1, 99) = 7.83$, $p = .005$, Cramer's V = .302.

To assess the importance of relevant alternative possibilities on response to the modal and indicative probes, I conducted a logistic regression model for each probe separately. Each model contained response to the relevance statement as the predictor and response to the probe as the outcome. Response to the relevance statement was significantly predictive for both the modal probe, Wald = 22.71, $p < .001$, and the indicative probe, Wald = 4.17, $p = .041$. Nevertheless, response to the relevance statement explained far more of the variance for the modal probe (between 30% and 41%) than for the indicative probe (between 5% and 8%).

4.3 Discussion

Three main findings emerge from this experiment, replicating and generalizing results from above. First, information about alternative possibilities affected modal judgments more strongly than it affected indicative judgments. This further supports the weak view of epistemic modals. Second, even though there was a salient chance that a proposition was false, when given a choice to say that the proposition "must" be true or that it "might" be false, more than half of people answered that it "must" be true. This further undermines the strong quantificational view of epistemic modals. Third, the outside/inside probability distinction again affected modal judgments.

5 Experiment 4: Must and Know

The results thus far support the weak view of epistemic modals but undermine the strong view. What could explain these findings? Why are epistemic modals sensitive to information about alternative possibilities despite resisting treatment in terms of universal or existential quantification over sets of possibilities? To begin answering this question, one strategy is to identify another category exhibiting a similar pattern of sensitivity to and tolerance of alternative possibilities. One such category is *knowledge*. Knowledge attributions are also sensitive to the difference between inside and outside probability (Friedman and Turri 2014; Turri et al. 2017; Turri ms.), and people are willing to attribute knowledge despite salient error possibilities (Turri and Friedman 2014; Colaço et al. 2014; Turri 2016). One hypothesis, then, is that epistemic modals behave as they do because of their relationship to knowledge. One version of this hypothesis is that modals express willingness or unwillingness to attribute knowledge based on the information. On this view, saying that a proposition must be true expresses willingness to attribute knowledge of it based on the available information. The present experiment tests this hypothesis. The prediction is that people's judgments about "must" modals will mediate the effect of condition on knowledge judgments.

5.1 Method

5.1.1 Participants

One hundred fifty new people participated (aged 18–62 years, mean age = 31 - years; 59 female; 96% reporting English as a native language). Participants were recruited and compensated the same way as in earlier experiments. Repeat participation was prevented.

5.1.2 Materials and Procedure

Participants were randomly assigned to one of two conditions, inside and outside, in a between-subjects design. Each participant read a single story, responded to test statements, then filled out a brief demographic survey. The outside/inside manipulation was the same as in Experiment 1 and the story was very similar, except this time the first sentence was changed to naturally allow an assessment of a third party's knowledge. Instead of the first sentence being, "Seth recently handed in his final paper for a university course," it was, "Seth's professor is marking final papers for a university course." After reading the story, participants responded to a knowledge statement:

The professor knows that Seth plagiarized his final paper.

Participants then went to a new screen and responded to a must modal:

It must be that Seth plagiarized his final paper.

Responses to the knowledge and modal statements were collected on the same 7-point Likert scale as earlier. Participants then went to a new screen and performed the same percentage task as in Experiment 1.

5.2 Results

Assignment to condition affected response to the knowledge statement and the modal statement but not performance on the percentage task (see Table 1). To test the mediating role of modal judgments on knowledge judgments, I conducted a bootstrap mediation analysis (Hayes 2013) with assignment to condition as the independent variable (coded: 0 = outside, 1 = inside), knowledge judgment as the

Table 1 Experiment 4

Measure	Mean (SD)		t	*df*	*p*	MD	d
	Inside	Outside					
Know	4.84 (1.67)	3.23 (1.78)	5.72	148	<.001	1.61	0.94
Must	4.57 (1.74)	2.96 (1.68)	5.77	148	<.001	1.61	0.95
Percent	83.2% (20.8)	80.4% (29.9)	0.81	148	.419	2.80	0.13

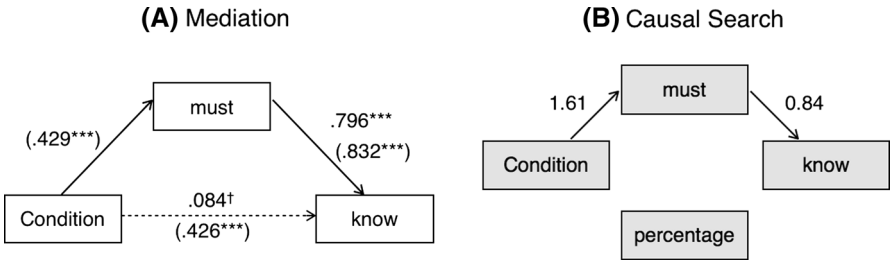

Fig. 5 Experiment 4. **a** Mediation results. *Parenthetical values* represent the strength of a simple regression between the two variables; values outside the *parentheses* represent the strength of the relationships in a model used to test for mediation. $^{\dagger}p \leq .10$; $^{***}p \leq .001$. **b** Causal search results. *Arrows* represent directional causal relations; *path coefficients* represent the strength of the causal relation

outcome, and modal judgment as potential mediator. This analysis showed that modal judgments mediated the effect of condition on knowledge judgments (see Fig. 5a). The 95% CI for the indirect effect was [0.79, 1.83]. The 95% CI for the direct effect was [−0.06, 0.69].

Because the knowledge judgments and modal judgments were strongly correlated, the reverse mediation model was also significant. Of course, this is perfectly consistent with the hypothesis being tested. Nevertheless, in order to gain insight into the underlying psychological processes involved, I conducted a causal search with with the Greedy Equivalence Search (GES) algorithm to identify the model that best fits the data. The causal search was conducted with Tetrad 5 (http://www.phil.cmu.edu/projects/tetrad/). GES considers all possible models available given the different variables. Each variable is treated as a node. GES assigns an information score to the model in which all the nodes are disconnected—the "null model." GES then evaluates adding causal arrows—"edges"—between the nodes (Meek 1997 provides the edge orientation rules). GES adds edges that best improve the model's information score, if such edges exist, until adding more edges does not improve the information score. At this point, GES evaluates whether deleting any edges will further improve the information score, and it deletes any such edges until deleting more edges does not improve the information score. Whereas regression assumes a causal direction, GES does not. GES is preferable to mediation models because it provides an overall measure of model fit, does not assume a causal direction, and tends to deliver more accurate models (Chickering 2002; Iacobucci et al. 2007).

I entered assignment to condition and response to the three dependent measures (know, must, percentage) into a causal search using GES. The model was constrained so that assignment to condition could not be caused by any other variable in the model. Figure 5b depicts the best fitting model, which fit the data well, $\chi^2(4) = 7.96$, $p = .09$, BIC $= -12.08$. In the model, modal judgments cause knowledge judgments.

5.3 Discussion

This experiment tested the hypothesis that must modals express willingness to attribute knowledge based on the available information. Results from mediation

analysis support the hypothesis. People's modal judgments mediated the effect on knowledge attributions of inside versus outside probabilistic information. Results from a causal search corroborated the mediation analysis. In the best fitting causal model of the data, people's modal judgments caused their knowledge attributions. The next experiment tests whether the same is true for people's modal judgments that a proposition might be false.

6 Experiment 5: Might and Know

This experiment tests whether saying that a proposition might be false expresses unwillingness to attribute knowledge based on the available information. The prediction is that people's judgments about "might" modals will mediate the effect of condition on knowledge judgments.

6.1 Method

6.1.1 Participants

One hundred fifty-two new people participated (aged 18–65 years, mean age = 30 years; 58 female; 99% reporting English as a native language). Participants were recruited and compensated the same way as in earlier experiments. Repeat participation was prevented.

6.1.2 Materials and Procedure

The story and procedures were exactly the same as in Experiment 4, except that the must modal was replaced with a might modal, "Seth might not have plagiarized his final paper."

6.2 Results

Assignment to condition affected response to the knowledge statement and the modal statement but not performance on the percentage task (see Table 2). To test the mediating role of modal judgments on knowledge judgments, I conducted a bootstrap mediation analysis (Hayes 2013) with assignment to condition as the

Table 2 Experiment 5

Measure	Mean (SD)		t	*df*	*p*	MD	d
	Inside	Outside					
Know	4.73 (1.77)	3.19 (1.72)	5.44	150	<.001	1.54	0.89
Might not	4.13 (1.89)	5.67 (1.04)	−6.19	119.2	<.001	−1.54	1.13
Percent	82.1% (27.3)	75.7% (28.4)	1.41	150	.161	6.37	0.23

independent variable (coded: 0 = outside, 1 = inside), knowledge judgment as the outcome, and modal judgment as potential mediator. This analysis showed that modal judgments mediated the effect of condition on knowledge judgments (see Fig. 6a). The 95% CI for the indirect effect was [0.70, 1.42]. The 95% CI for the direct effect was [0.02, 1.05].

Once again, the reverse mediation model was also significant, so I conducted a causal search with with the Greedy Equivalence Search (GES) algorithm to identify the model that best fits the data (see the "Results" section of Experiment 4 for more details on the algorithm). I entered assignment to condition and response to the three dependent measures (know, might not, percentage) into a causal search using GES. The model was constrained so that assignment to condition could not be caused by any other variable in the model. Figure 6b depicts the best fitting model, which fit the data well, $\chi^2(4) = 6.54$, $p = .16$, BIC $= -13.55$. In the model, modal judgments cause knowledge judgments.

6.3 Discussion

This experiment tested the hypothesis that might modals express (un-)willingness to attribute knowledge based on the available information. Results from mediation analysis support the hypothesis. People's modal judgments mediated the effect on knowledge attributions of inside versus outside probabilistic information. Results from a causal search corroborated the mediation analysis. In the best fitting causal model of the data, people's modal judgments caused their knowledge attributions.

7 General Discussion

According to the weak view of epistemic modals, modal judgments are sensitive to information about alternative possibilities in a way that indicative judgments are not. According to the strong quantificational view, epistemic modals quantify over alternative possibilities in very specific ways: a proposition *must* be true just in case it is true in *all* relevant possibilities consistent with the available information; and a

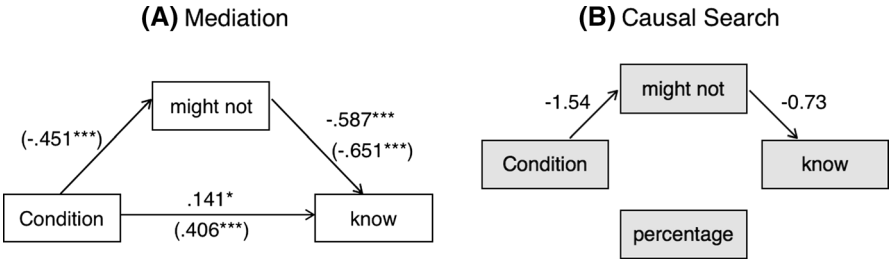

Fig. 6 Experiment 5. **a** Mediation results. *Parenthetical values* represent the strength of a simple regression between the two variables; values outside the *parentheses* represent the strength of the relationships in a model used to test for mediation. *$p \leq .05$; ***$p \leq .001$. **b** Causal search results. *Arrows* represent directional causal relations; *path coefficients* represent the strength of the causal relation

proposition *might* be true just in case it is true in *at least one* relevant possibility consistent with the available information. These views are very popular among theorists, but they have not been tested previously. To test these views, I conducted three experiments. Overall the results supported the weak view but undermined the strong view. Information about alternative possibilities affected modal judgments more strongly than it affected indicative judgments, which supports the weak view. However, despite acknowledging the possibility that a proposition is false, people were still willing to agree that the proposition must be true and deny that it might be false, which undermines the strong quantificational view. To make sense of these findings, I conducted two further experiments that tested a hypothesis. According to the hypothesis, epistemic modals behave as they do because of their relationship to knowledge. The results supported this hypothesis.

Several of the experiments used a manipulation shown by prior research to affect decision-making and social evaluations: the difference between "inside" and "outside" probabilities. When probabilistic information pertained to a propensity of a specific item in the scenario (inside) rather than base rates (outside), people were more willing to judge that the target proposition must be true and less likely to judge that the proposition might be false (Experiments 1, 4 and 5). Similarly, when people received inside probabilistic information that a proposition was likely to be true, most people answered that the proposition must be true, rather than that it might be false. By contrast, when receiving outside information that the proposition was likely to be true, the vast majority of people answered that the proposition might be false, rather than that it must be true (Experiment 3). These results show the outside/inside probability distinction affects modal judgments, adding to a growing list of tasks that exhibit an outside/inside asymmetry (see Lagnado and Sloman 2004; Turri et al. 2017).

Overall the results support the weak view of epistemic modals. The outside/inside effect on modal judgments was mediated by the estimated chance that the proposition was false, and by the perceived relevance of the possibility that the proposition was false. By contrast, even though the outside/inside difference also affected people's indicative judgments about the proposition, I did not find clear evidence that this effect was mediated by estimations of chance or judgments about relevant possibilities (Experiment 1). Similarly, regression analysis showed that the perceived relevance of error possibilities explained much more of the variance in people's modal judgments than in their indicative judgments (Experiment 3).

By contrast, the results undermine the strong quantificational view. People agreed that a proposition must be true even though the proposition's chance of being false was estimated at up to nearly 40%, and even though they agreed that its falsity was a relevant possibility (Experiments 1, 2 and 4). For example, people rated the chance that a smartphone had spyware on it at around 60%, but they still tended to agree that the smartphone must have spyware, and they were unwilling to agree that the smartphone might not have spyware. Similarly, when presented information that a proposition was very likely, but not guaranteed, to be true, most people said that the proposition must be true, rather than that it might be false. In this respect, the overall patterns observed in modal judgments were surprisingly similar to the patterns observed for indicative judgments. Prior work in formal semantics has argued

against the strong view on theoretical grounds (Lassiter 2011), and this relates to discussion among linguists on how to interpret "must" (Lassiter 2014). The present results are the first empirical evidence that the strong view fails to capture the ordinary meaning of epistemic modals.

Epistemic modals are sensitive to information about alternative possibilities despite resisting treatment in terms of universal or existential quantification over sets of relevant possibilities. But why do epistemic modals behave this way? Prior work has shown that knowledge attributions are also sensitive to the difference between inside and outside probability, and that people are willing to attribute knowledge despite salient error possibilities (Turri and Friedman 2014; Colaço et al. 2014; Turri 2016). This suggests the hypothesis that epistemic modals behave as they do because of their relationship to knowledge. More specifically, epistemic modals might express people's willingness to attribute knowledge based on the available information. Saying that a proposition must be true expresses willingness to attribute knowledge of it based on the available information, and saying a proposition might be false expresses unwillingness to do so. Otherwise put, saying that a proposition must be true indicates that the available information suffices to know it, and saying that it might be false indicates that the available information does not suffice. Two experiments provided initial evidence supporting this hypothesis (Experiments 4 and 5). In these experiments, the outside/inside effect on knowledge attributions was mediated by judgments about what must be true or might be false.

Theorists have previously proposed a link between epistemic modality and knowledge. As one commentator put it, "Clearly, epistemic modals have something to do with knowledge" (MacFarlane 2011, p. 144). And one primary source of inspiration for the strong view informally introduces the view in terms of "what is known" (Kratzer 1977, p. 341 ff) (for alternative theoretical approaches linking knowledge and modality, see Hacking 1967; DeRose 1991; and Hawthorne 2004, pp. 25–27). However, when formalizing this notion, theorists imposed a quantificational structure that does not fit the way epistemic modals are actually used and evaluated. Contrary to the received view, it turns out not to be a "key insight" that "must" works like a universal quantifier, or that "might" works like an existential quantifier (cf. Schaffer 2011, p. 207). For people will agree that a proposition "must" be true, and they will disagree that it "might" be false, while acknowledging its falsity as a relevant possibility. My hypothesis regarding the expressive function of epistemic modals accommodates the behavioral data while preserving the intuitive link between epistemic modals and knowledge.

Nevertheless, taken as a whole, the above results might not uniformly support my hypothesis. In particular, one pair of results might raise questions. In Experiment 2, when people rated the probability of a proposition at around 60%, they were inclined to agree that the proposition must be true, and that the proposition might not be true. One might doubt that this is because people were inclined to both attribute and to deny knowledge. Nevertheless, it is possible that the case approached a tipping point between attributing and denying knowledge, in which case people might well harbor contrary inclinations. Moreover, the results in question were from a between-subjects study, so no single group of people exhibited

both of these tendencies overall. Future work could investigate whether this same pattern of results recurs within-subjects. If it does, then my simple hypothesis regarding epistemic modals will likely have to be amended.

My argument against the strong quantificational view may be qualified in at least two respects. On the one hand, as noted in the introduction, the view's proponents distinguish between relevant and irrelevant possibilities, and they claim that "must" acts as a universal quantifier over only *relevant* possibilities. On their view, a "must" modal could still be true even if an *irrelevant* possibility of falsehood is consistent with the available information. I attempted to accommodate this by asking participants to record judgments about whether the possibility of falsehood was relevant. This is clearly the simplest and most straightforward way to measure perceptions of relevance, it was effective enough to demonstrate an outside/inside effect on relevance judgments, and it was effective enough to detect that relevance judgments mediate effects on modal judgments. Nevertheless, it is possible that there are better ways to measure relevance judgments and that the results would then be more favorable to the strong view. On the other hand, some proponents of the strong view propose specific, complicated accounts of how relevance is determined by "intersecting sets of possibilities compatible" with what is known by all members in a contextually determined group of people, and also by a "speaker's publicly manifestable" linguistic intentions (Dowell 2011, p. 5). The present experiments were not designed to test accounts of this sort. If such accounts capture the ordinary meaning of modals and thereby help explain "language use" (Dowell 2011, p. 1), then the proposed mechanisms determining relevance are operative in the psychology of ordinary language users. It could advance our collective understanding of the underlying issues if proponents of such accounts explained how to operationalize and test key features of their theories. Perhaps the results from the present experiments could be helpful toward that end. For instance, perhaps there is a way to relate the outside/inside effect on modal judgments to the more complicated machinery envisioned by such theories.

My findings and hypothesis on epistemic modals could have implications for legal practice and decision-making. For example, in common law countries, criminal conviction requires proving guilt "beyond a reasonable doubt," but it is notoriously difficult to say precisely what this means (Whitman 2008; Diamond 1990). One natural way to understand it, though, is in terms of epistemic modals. When forced to apply the "reasonable doubt" rule, jurors might conclude that doubt is reasonable if the defendant might be innocent. Conversely, jurors might conclude that doubt is unreasonable if the defendant must be guilty. If my hypothesis about epistemic modals is correct, then jurors who reason this way implicitly apply a *knowledge standard* in criminal proceedings. Accordingly, a successful prosecution would have to convince these jurors that they know the defendant is guilty, and a successful defense would have to convince jurors that they do not know. In turn, the emergent study of folk epistemology in general, and knowledge attributions in particular, could be relevant to the criminal justice system.

Acknowledgements For helpful feedback and discussion, I thank James Beebe, Wesley Buckwalter, Ori Friedman, Joshua Knobe, Daniel Lassiter, Lisa Matthewson, Jonathan Phillips, David Rose, and Angelo

Turri. Thanks also to anonymous reviewers for this journal. This research was supported by the Social Sciences and Humanities Research Council of Canada, the Ontario Ministry of Economic Development and Innovation, and the Canada Research Chairs program.

References

Austin, J. L. (1962). *How to do things with words*. J. O. Urmson (Ed.). Oxford: Oxford University Press.

Bell, V. A., & Johnson-Laird, P. N. (1998). A model theory of modal reasoning. *Cognitive Science, 22*(1), 25–51.

Chickering, D. M. (2002). Optimal structure identification with greedy search. *The Journal of Machine Learning Research, 3*(1), 507–554.

Colaço, D., Buckwalter, W., Stich, S., & Machery, E. (2014). Epistemic intuitions in fake-barn thought experiments. *Episteme, 11*(02), 199–212.

DeRose, K. (1991). Epistemic possibilities. *Philosophical Review, 100*(4), 581–605.

Diamond, H. A. (1990). Reasonable doubt: To define, or not to define. *Columbia Law Review, 90*(6), 1716.

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

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

Friedman, O., & Turri, J. (2014). Is probabilistic evidence a source of knowledge? *Cognitive Science, 39*(5), 1062–1080. doi:10.1111/cogs.12182.

Furrow, D., & Moore, C. (1990). Gender differences in differentiating terms expressing certainty. *Journal of Psycholinguistic Research, 19*(6), 375–385.

Hacking, I. (1967). Possibility. *The Philosophical Review, 76*(2), 143.

Hawthorne, J. (2004). *Knowledge and lotteries*. Oxford: Oxford University Press.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York: Guilford Press.

Iacobucci, D., Saldanha, N., & Deng, X. (2007). A meditation on mediation: Evidence that structural equations models perform better than regressions. *Journal of Consumer Psychology, 17*(2), 139–153.

Kahneman, D., & Tversky, A. (1982). Variants of uncertainty. *Cognition, 11*(2), 143–157.

Khoo, J. (2014). *Modal disagreements*. Unpublished Manuscript. Massachusetts Institute of Technology. Retrieved August 29, 2014 from http://web.mit.edu/jkhoo/www/MD4.pdf.

Knobe, J., & Szabó, Z. G. (2013). Modals with a taste of the deontic. *Semantics and Pragmatics, 6*(1), 1–42.

Knobe, J., & Yalcin, S. (2014). Epistemic modals and context: Experimental data. *Semantics and Pragmatics, 7*(11), 1–21.

Kratzer, A. (1977). What must and can must and can mean. *Linguistics and Philosophy, 1*(3), 337–355.

Kripke, S. (1972). *Naming and necessity*. Cambridge, MA: Harvard University Press.

Lagnado, D., & Sloman, S. A. (2004). Inside and outside probability judgments. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment & decision making* (pp. 157–176). Malden, MA: Blackwell.

Lassiter, D. (2011). *Measurement and modality: The scalar basis of modal semantics*. Ph.D. Thesis. New York University.

Lassiter, D. (2014). The weakness of must: In defense of a mantra. *Proceedings of SALT, 24,* 597–618.

Lewis, D. (1979). Scorekeeping in a language game. *Journal of Philosophical Logic, 8*(1), 339–359.

Lewis, D. (1996). Elusive knowledge. *Australasian Journal of Philosophy, 74*(4), 549–567.

MacFarlane, J. (2011). Epistemic modals are assessment-sensitive. In A. Egan & B. Weatherson (Eds.), *Epistemic modality* (pp. 144–178). Oxford: Oxford University Press.

Machery, E., & Stich, S. (2012). The role of experiment in the philosophy of language. In G. Russell & D. G. Fara (Eds.), *The Routledge companion to the philosophy of language* (pp. 495–512). New York: Routledge.

Meek, C. (1997). *Graphical models: Selecting causal and statistical models*. Carnegie Mellon University.

Moore, C., Pure, K., & Furrow, D. (1990). Children's understanding of the modal expression of speaker certainty and uncertainty and its relation to the development of a representational theory of mind. *Child Development, 61*(3), 722–730.

Papafragou, A. (2006). Epistemic modality and truth conditions. *Lingua, 116*(10), 1688–1702.
Schaffer, J. (2011). Perspective in taste predicates and epistemic modals. In A. Egan & B. Weatherson (Eds.), *Epistemic modality* (pp. 179–226). Oxford: Oxford University Press.
Stephenson, T. (2007). Judge dependence, epistemic modals, and predicates of personal taste. *Linguistics and Philosophy, 30,* 487–525.
Turri, J. (2016). A new paradigm for epistemology: From reliabilism to abilism. *Ergo, 3*(8), 189–231.
Turri, J. (ms). Knowledge attributions and lottery cases: A review and new evidence. University of Waterloo.
Turri, J., & Friedman, O. (2014). Winners and losers in the folk epistemology of lotteries. In J. R. Beebe (Ed.), *Advances in experimental epistemology* (pp. 45–69). London: Bloomsbury.
Turri, J., Friedman, O., & Keefner, A. (2017). Knowledge central: A central role for knowledge attributions in social evaluations. *Quarterly journal of experimental psychology, 70*(3), 504–515.
Unger, P. (1984). *Philosophical relativity*. Oxford: Oxford University Press.
von Fintel, K., & Gillies, A. S. (2007). An opinionated guide to epistemic modality. In T. S. Gendler & J. Hawthorne (Eds.), *Oxford studies in epistemology* (Vol. 2, pp. 32–62). Oxford: Oxford University Press.
von Fintel, K., & Gillies, A. S. (2008). CIA Leaks. *Philosophical Review, 117*(1), 77–98.
von Fintel, K., & Gillies, A. S. (2010). Must… stay… strong! *Natural Language Semantics, 18*(4), 351–383.
Wells, G. L. (1992). Naked statistical evidence of liability: Is subjective probability enough? *Journal of Personality and Social Psychology, 62*(5), 739.
Whitman, J. Q. (2008). *The origins of reasonable doubt: Theological roots of the criminal trial*. New Haven: Yale University Press.
Yalcin, S. (2007). Epistemic modals. *Mind, 116*(464), 983–1026.
Yalcin, S. (2011). Nonfactualism about epistemic modality. In A. Egan & B. Weatherson (Eds.), *Epistemic modality* (pp. 295–332). Oxford: Oxford University Press.