Supplement for in Decomposing Modal Thought

Jonathan Phillips & Angelika Krazter

November, 2022

Normality constraints

We experimentally investigated the defeasible normality constraints that play a role in restricting the modal domain, and in particular focused on their defeasibility.

Methods

Participants We collected a sample of 254 participants (M_{age} = 37.31; SD_{age} = 11.07; 128 females) from Amazon Mechanical Turk (www.mturk.com).

Procedure Participants were randomly assigned to one of two conditions. In one condition, participants were first presented with the following question:

> A child was born two years ago. The child was born from its mother's first pregnancy and its mother died a year later before becoming pregnant again.
>
> Do you agree or disagree?
>
> ***The child must not have any siblings.***

In the other condition, participants first read a slight variation which the possibility of multiple births should be less likely to be excluded because of implicit normality constraints:

> A dog was born two years ago. The dog was born from its mother's first pregnancy and its mother died a year later before becoming pregnant again.
>
> Do you agree or disagree?
>
> ***The dog must not have any siblings.***

In both cases, participants' rated their agreement on a scale from 0 ('Disagree') to 100 ('Agree').

Then the possibility of there being multiple births from a single pregnancy was explicitly raised. In the condition with a child, we told participants:

> Now please carefully consider the following possibility.
>
> ***The mother could have had twins, triplets, or even more babies at once.***

and in the dog condition, we told participants:

> Now please carefully consider the following possibility.
>
> ***The mother could have had given birth to two, three, or even more puppies at once.***

Participants were then asked to rate their agreement with the must claim again, on the same scale as before.

Participants then completed a question which asked them to write down what the new possibility raised was, and then told us whether they had considered that possibility before we explicitly pointed it out. Lastly, participants completed a brief demographic questionnaire.

Results

We first analyzed participants agreement with the 'must not' claim using linear mixed effects models with participants as a random factor and Condition (child vs. dog) and Challenge (before challenge vs. after challenge) as fixed factors that were allowed to interact. This analysis revealed a main effect of Condition, $\chi^2(1)= 22.317$, $p < 0.001$, a main effect of Challenge, $\chi^2(1)= 53.651$, $p < 0.001$, and critically, a highly significant interaction effect as predicted, $\chi^2(1)= 13.967$, $p < 0.001$ (Figure 1).

We investigated this interaction effect with planned comparisons that separately focused only on the child and dog cases. Participants largely agreed with the claim that the child must not have any siblings before the challenge ($M = 65.662$; $SD = 35.739$), but this agreement was reduced drastically after the possibility of twins or triplets was raised explicitly ($M = 33.592$; $SD = 29.922$), $t(250.26) = 7.844$, $p < 0.001$, $d = 0.973$. By contrast, participants did not strongly agree with the claim that the dog must not have any siblings, even before the challenge, ($M = 37.863$; $SD = 41.851$), and explicitly raising the possibility of multiple children at birth had a small effect on their agreement ($M = 26.669$; $SD = 36.681$), $t(246) = 2.24$, $p = 0.026$, $d = 0.284$.

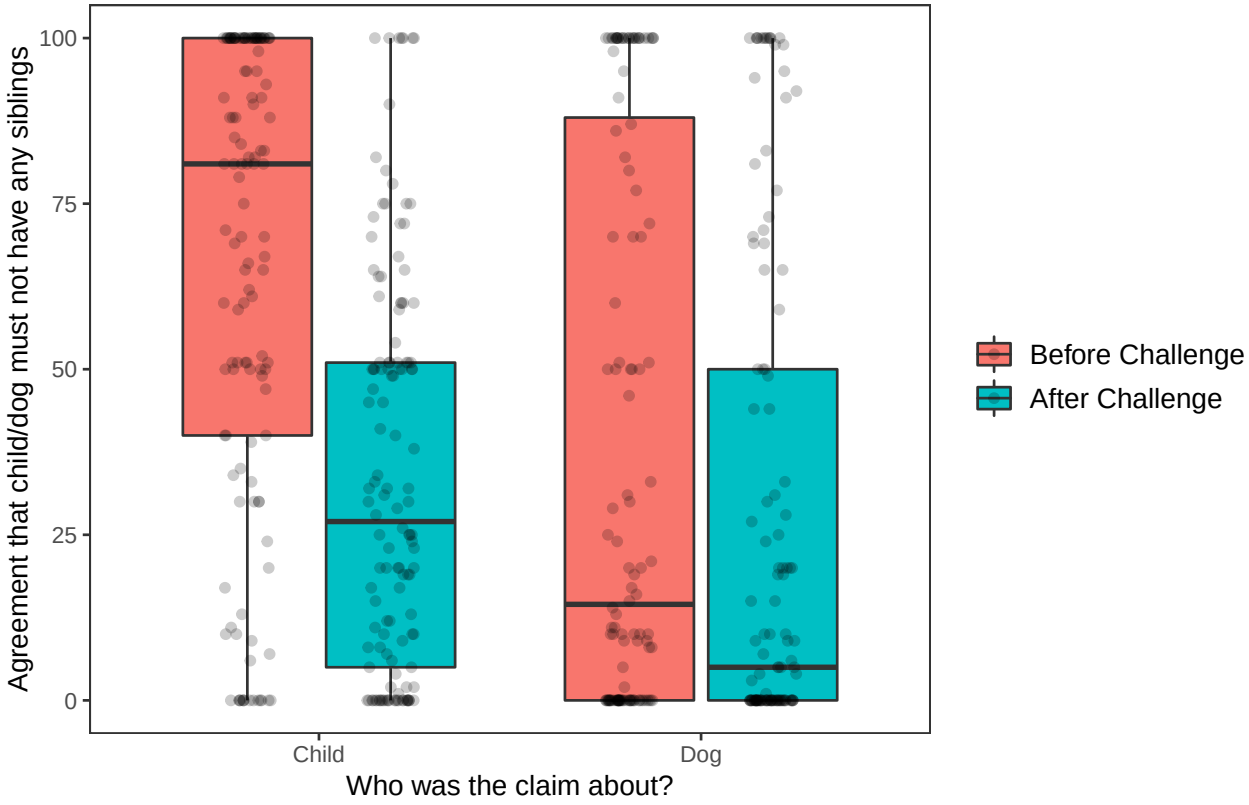

Figure 1: **Figure 1**. Agreement with 'must not' claim as a function of both condition and challenge.

We next investigated whether this difference arose from differences in participants' tendency to exclude the possibility of multiple births from a single pregnancy in the case of a child (where they are relatively abnormal), but not in the case of the dog (where this possibility is relatively normal). To do this, we first considered participants responses to the question that asked them whether they considered this possibility before we raised it, and found that this tendency shifted drastically between conditions, $\chi^2(1) = 40.202$, $p < 0.001$ (Table 1).

Table 1: **Table 1**.

Condition	No	Yes	Other
Child	93	34	3
Dog	39	82	3

While these data demonstrate that our manipulation worked as intended, we have not yet shown that this difference is what actually explains the difference in participants' agreement with the 'must' claim. Two further analyses suggest that this is the case. First, collapsing across Condition (child vs. dog), one finds whether or not participants considered the possibility of multiple births has the predicted effect on agreement with the 'must' claim, both before and after the possibility is explicitly raised (Fig. 2). Statistically, this effect can be captured by the interaction effect between whether or not the possibility had been explicitly raised, and whether participants considered it before it was raised, $\chi^2(1)= 60.651$, $p < 0.001$.

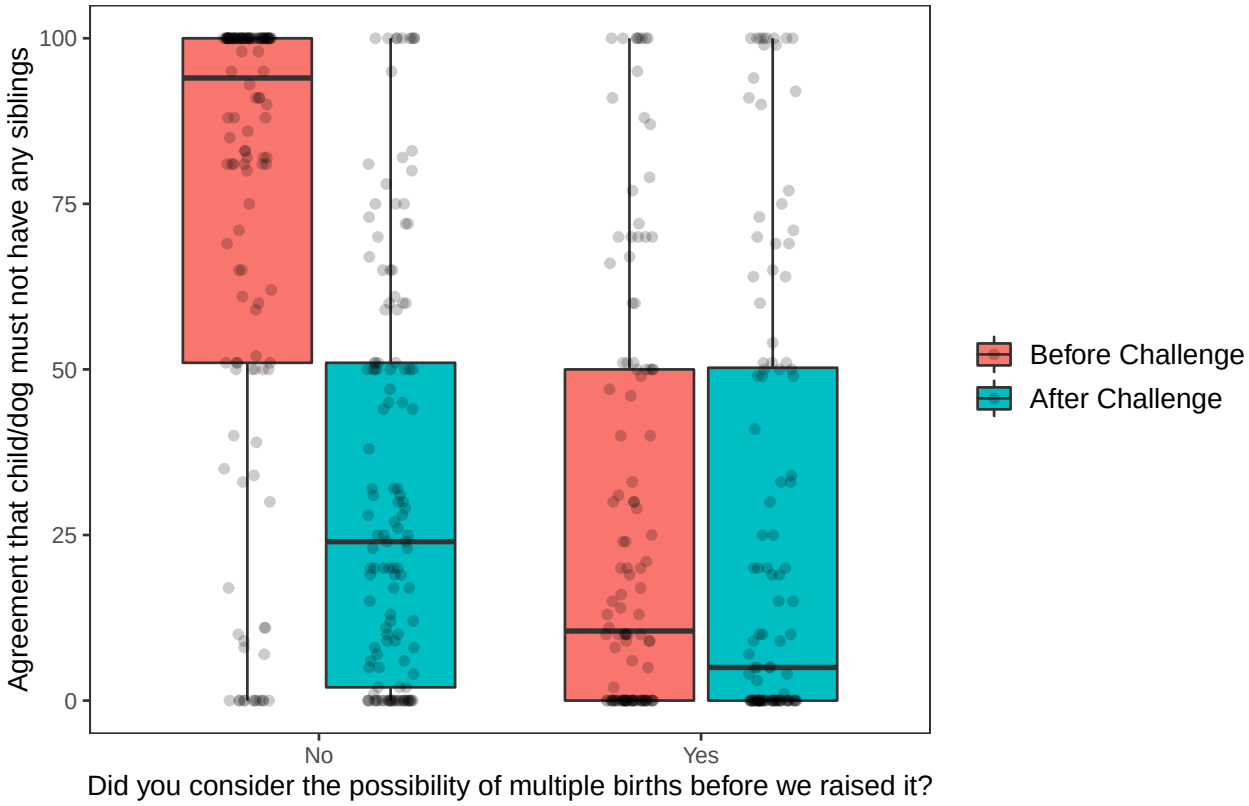

Figure 2: **Figure 2**. Agreement with 'must not' claim as a function of both challenge and whether the possibility was previously considered.

Second, we asked whether the effect of condition on participants' agreement with the 'must' claim before the possibility was raised was mediated by whether or not they considered the possibility of multiple births. A bootstrap mediation analysis revealed that whether the possibility was considered fully mediated effect of Condition on agreement with the 'must' claim: 95% *CI* of proportion mediated [0.44, 1.02], $p < 0.001$.

Modal Anchors

We empirically investigated whether ordinary truth-value judgments of sentences involving different modal terms differ in the predicted way in the context of Lewis's (1997) sorcerer context. We additionally assessed

participants' subjective reports of the relevant modal anchor when evaluating the truth of different modal terms.

Methods

Participants All participants were recruited through Amazon Mechanical Turk and the survey was administered through Qualtrics. In **Study 1**, we recruited 249 participants (M_{age}=36.1, SD_{age}=11.19; 110 females); in **Study 2**, we recruited 203 participants (M_{age}=37.4, SD_{age}=12.3; 95 females); in **Study 3**, we recruited 99 participants (M_{age}=37.31, SD_{age}=12.25; 34 females); in **Study 4**, we recruited 305 participants (M_{age}=38.78, SD_{age}=11.35; 144 females).

Procedure All participants read the following scenario:

> Imagine that there is a production line that produces a special kind of glass, each of which is very pretty but is also extremely fragile. Each glass is sure to break from even a slight bump.
>
> A magical sorcerer takes a liking to a glass from this production line and decides to protect the glass. He doesn't do anything to change the glass itself; he likes the glass exactly the way it is. He only watches and waits, resolved that if ever his glass is struck, then, quick as a flash, he will cast a spell that changes the glass, renders it no longer fragile, and thereby aborts the process of breaking.

After reading this scenario, participants rated their agreement with one of several possible sentences about the glass the sorcerer liked:

> The glass is fragile. [*Fragile*]
>
> The glass could break. [*Could*]
>
> The glass might break. [*Might*]
>
> The glass is vulnerable. [*Vulnerable*]

In all cases, participants indicated their agreement with the sentence on a scale from (0) "Completely disagree" to (100) "Completely agree.

After answering this question, participants were asked the following question about the judgment they just made:

> When you answered the last question, were you thinking only about the glass itself, or about the situation the glass was in and the sorcerer who wants to protect it?

They selected one of three options: (a) "Only about the glass", (b) "About the glass and the sorcerer", or (c) "Other" for which they could optionally write in an answer.

The four studies varied only in which set of questions participants were asked. In Study 1, participants evaluated *Could* and *Might*; in Study 2, participants evaluated *Fragile* and *Could*; in Study 3 participants evaluated *Might*; in Study 4, participants evaluated *Could*, *Fragile*, and *Vulnerable*.

Results

Given the identical nature of the study procedures and overlapping conditions across the four studies, the data were combined and analyzed jointly. Overall, there was a clear effect of question on participants' agreement $F(3)= 23.27$, $p < 0.001$, $\eta_p^2 = 0.076$ (see Fig 3.) More specifically, participants more agreed with *Fragile* than *Vulnerable* $t(167.33) = 2.94$, $p = 0.004$, $d = 0.384$ and more agreed with *Vulnerable* than *Could* $t(182.85) = 2.62$, $p = 0.009$, $d = 0.28$. The difference in agreement between *Could* and *Might* was not statistically significantly ($\alpha = .05$), $t(485.43) = 1.24$, $p = 0.217$, $d = 0.107$.

As reported in the main article, we additionally found that there was more variation in participants' agreement ratings with the *Could* claim, than the with the *Fragile* claim $F(328,201) = 1.892$, $p < 0.001$.

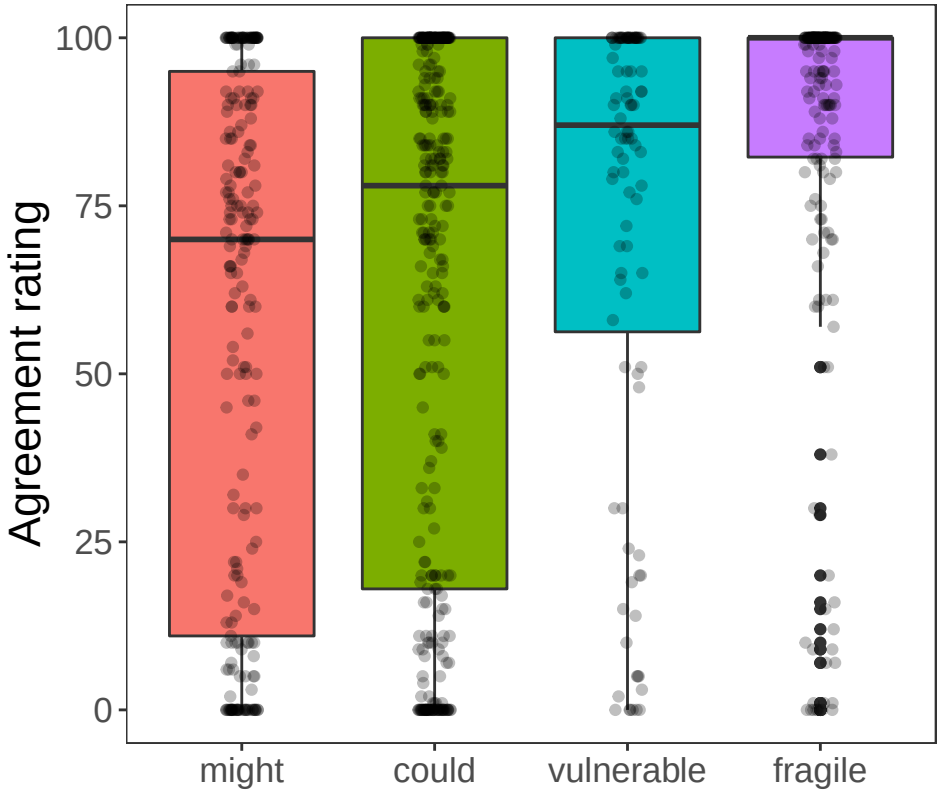

Figure 3: **Figure 3.** Boxplots of participants' agreement ratings with each of the four sentences about the Sorcerer's glass.

Participants' responses for whether they were thinking of the sorcerer when rating their agreement with each of the four sentences produced an inverse pattern (Fig. 4). Specifically, when rating their agreement with *Fragile*, most participants did not think about the sorcerer (0.32%). Moreover, compared to *Fragile*, participants were significantly more likely to think about the sorcerer when rating their agreement with *Vulnerable* (0.48%), $p = 0.005$; participants were even more likely to think about the sorcerer when rating their agreement with *Could* (0.65%), $p < 0.001$, and *Might* (0.65%), $p < 0.001$.

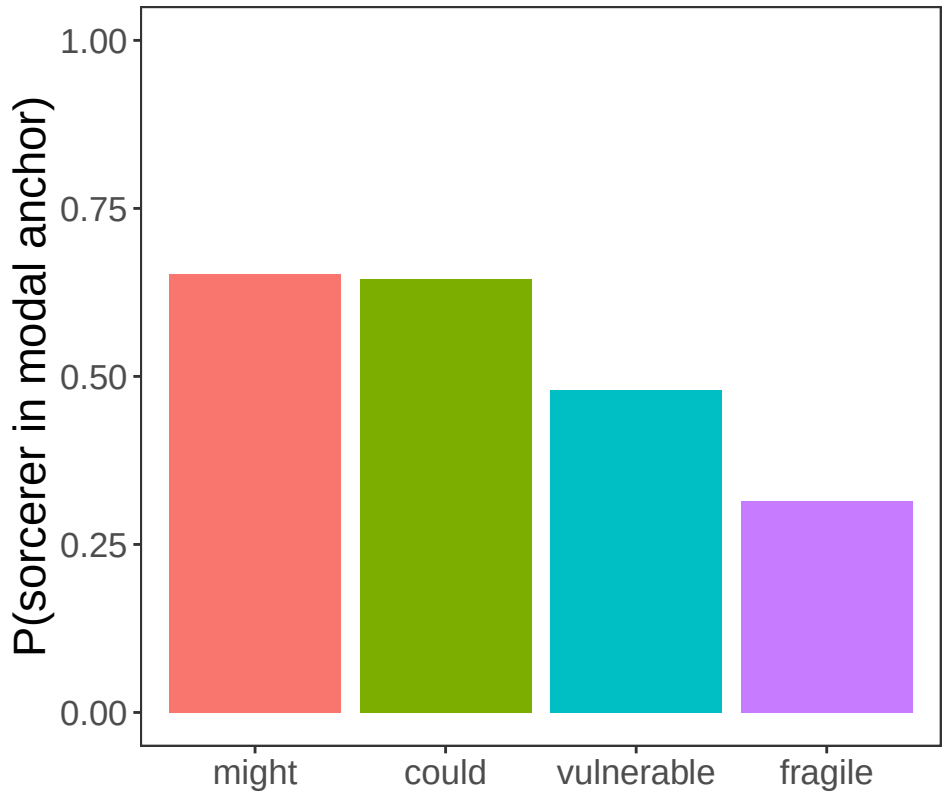

Figure 4: **Figure 4.** Probability that participants were *not* thinking about the scorcerer when rating their agreement with each of the four sentences about the glass.

Table 2: **Table 2**.

Modal Term	N	Mean Agreement	SD	SE	P(sorcerer)
could	329	61.58055	39.29012	2.166135	0.6452599
might	225	57.41333	38.78769	2.585846	0.6517857
fragile	202	84.11386	28.56644	2.009928	0.3150000
vulnerable	100	72.31000	34.69221	3.469221	0.4800000

We next demonstrated that whether or not participants reported that they were thinking of the sorcerer was highly predictive of their agreement rating overall, $F(1,843)=253.92$, $p = 0.001$. Further, we can ask whether the consideration of the sorcerer when evaluating the sentence differentially affected different modal judgments. We found that it did (Figure 5). Statistically, this pattern can be captured by an interaction between the different modal questions and whether or not the anchor included the sorcerer, $F(3,843)=3.35$, $p = 0.019$. More specifically, we found when the modal anchor included both the the glass and the sorcerer, answers to the different questions differed substantially $F(3,464)=7.21$, $p < 0.001$, while they became much more similar to one another when the modal anchor included only the glass but not the sorcerer $F(3,379)=3.02$, $p = 0.03$.

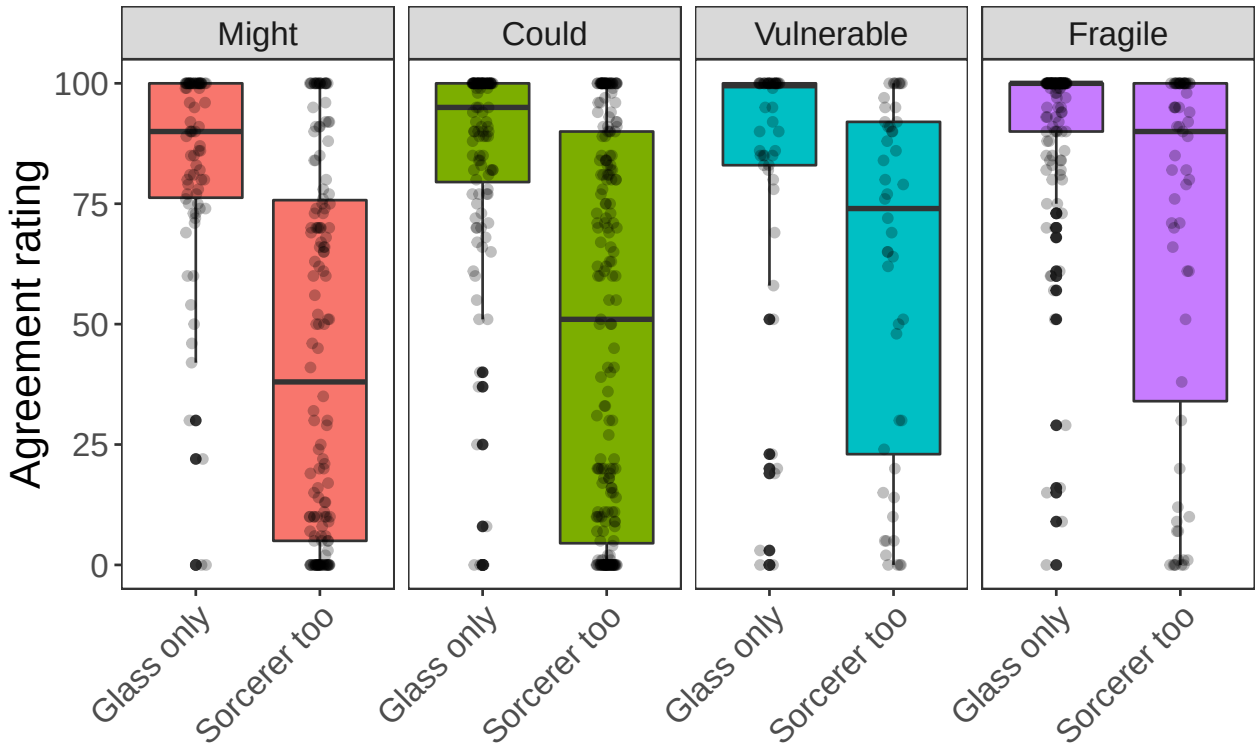

Figure 5: **Figure 5.** Boxplots of participants' agreement ratings with each of the four sentences about the Sorcerer's glass split by whether the modal anchor included the sorcerer.

As reported in the main paper, we also asked whether the difference in agreement with *Fragile* vs. *Could* can be explained by whether or not the sorcerer was included in the modal anchor. A bootstrap mediation analysis revealed that whether the sorcerer was considered largely mediated this difference: 95% *CI* of proportion mediated [0.32, 0.66], $p < 0.001$.

Supplmental discussion

Modal judgments like whether or not the glass 'might' or 'could' break take anchors that can vary dramatically in their contents and thus change the set of possibilities projected. Critically, they can take large situations that may include the sorcerer protecting the glass, and thus should be sensitive to whether the participant was considering the sorcerer when evaluating the sentence. As predicted, we find (i) large differences in agreement ratings with sentences involving these terms, (ii) correspondingly large differences in whether the modal anchor included the sorcerer, and (iii) that (i) is predicted by (ii).

In comparison, so-called "stage-level" predicates like 'vulnerable' will tend to exhibit somewhat less flexibility in their anchor argument, while still allowing for some variation because the anchor includes a spatiotemporal component that is not explicitly delimited. This aspect of 'vulnerable' suggests that we should find comparatively more agreement with the sentences involving 'vulnerable' and a correspondingly weaker tendency for the evaluations of 'vulnerability' to be affected by whether participants were considering the sorcerer. This is the pattern we observe.

Finally, for "individual-level" predicates like 'fragile', the modal anchor is provided by the syntactic subject and thus truth-value judgments of sentences involving such predicates should exhibit the least amount of variation, and correspondingly the highest level of agreement. We find the highest level of agreement in participants' evaluations of fragility and find that these judgments are the least affected by whether or not participants were considering the sorcerer when responding.

Lexically specified domain restriction

Here we empirically investigated how the difference in the lexical specified domain restrictions encoded by two modal auxiliaries, *Could* and *Might*.

Methods

Participants We collected a sample of 200 participants (M_{age} = 40.23; SD_{age} = 11.29; 99 females) from Prolific (www.prolific.com).

Procedure All participants read the following background context:

> FitzRoy is a notoriously ruthless pirate who decided to pose as a captain of a ship in order to steal several expensive sculptures that a museum needs to transport across the sea. After he posed as an ordinary ship captain who was taking some passengers across the sea, FitzRoy got the job.
>
> While sailing on the sea, a large storm came upon FitzRoy and his small ship. As the waves began to grow larger, FitzRoy realized that his small vessel was too heavy and the ship would flood if he didn't make it lighter. The only things on FitzRoy's small boat were the expensive art sculptures that he was stealing and the passengers he was transporting. He knew he had to throw something overboard to keep the ship from capsizing.

After reading, participants rated their agreement with the four following statements on a 100 point scale from 0 ('Completely Disagree') to 100 ('Completely Agree').

> FitzRoy could throw the sculptures overboard.
>
> FitzRoy might throw the sculptures overboard.
>
> FitzRoy could throw the passengers overboard.
>
> FitzRoy might throw the passengers overboard.

Participants then completed two comprehension questions. The first asked what FitzRoy was notorious for, and participants selected one of the three following options

- being generous
- being ruthless
- his name

The second asked what was *not* described as being on the ship, and participants selected on of the three following options:

- Passengers
- Sculptures
- FitzRoy's wife

Lastly, participants completed a brief demographic questionnaire.

Results

We excluded all participants who did not answer both comprehension check questions correctly (11 participants). For simplicity we focus on participants' first responses. Overall, we found the predicted interaction between the modal term used (*could* vs. *might*) and whether the modal claim concerned throwing passengers or sculptures overboard, $F(1,185)=54.73$, $p = 0.001$, $\eta_p^2 = 0.228$ (see Fig. 6). Specifically, participants more agreed that FitzRoy *could* throw the sculptures overboard ($M = 92.167$; $SD = 14.289$) than they agreed that he *might* throw the sculptures overboard ($M = 46.595$; $SD = 30.669$), $t(56.28) = 8.828$, $p < 0.001$, $d = 1.948$. By contrast, they less agreed that FitzRoy could throw the passengers overboard ($M = 65.13$; $SD = 37.043$) than they agreed that he *might* throw the sculptures overboard ($M = 79.889$; $SD = 22.967$), $t(90.08) = -2.422$, $p = 0.017$, $d = 0.469$

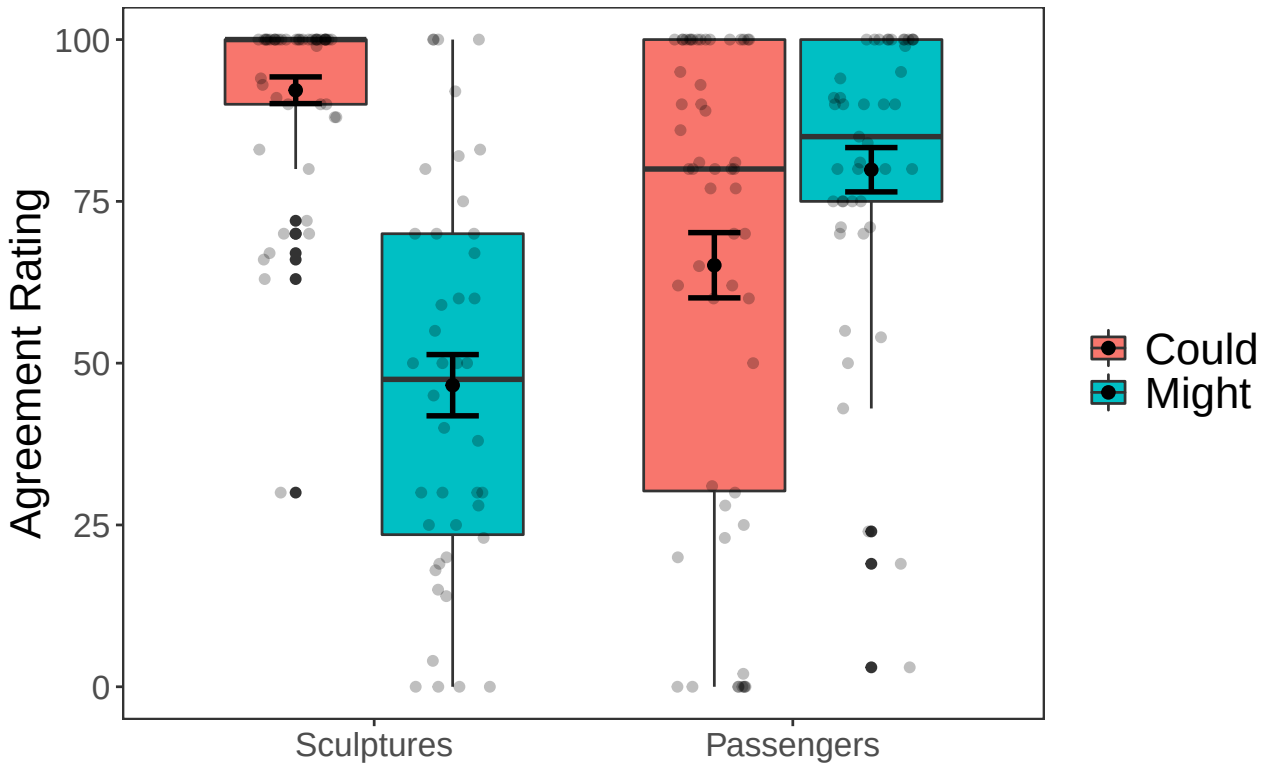

Figure 6: **Figure 6.** . Boxplots of participants' agreement the existential modal claims when they concerned ratings FitzRoy throwing the passengers overboard (left boxes and points) or the sculptures overboard (right boxes and points), both when the existential modal used was could (red boxes) and when it was might (blue boxes). The colored boxes depict the middle two quartiles of responses, the horizontal line depicts the median response, and the small grey dots depict individual participant responses. Thick black dots depict mean agreement and error bars represent +/- 1 SEM.