Modal auxiliary verbs are used to give a judgment or interpretation about an action or state. Here we consider deontic and epistemic modalities, which show systematic ambiguity in modal verbs in many languages. Consider (1). The deontic interpretation gives the meaning “He has permission to be there,” whereas the epistemic interpretation gives the meaning “It is a possibility that he’s there.” The deontic reading is one of obligation or permission (relative to the rules of the situation), and the epistemic reading is one of necessity or possibility (relative to the speaker’s knowledge about the situation). This is an example of how intrinsic (lexical) modal force (e.g. that “must” has intrinsically stronger implications than “may”) can be modified by conversational backgrounds (Kratzer 1981, 1991).

(1) He may be in the other room.

The context around the utterance provides the laws and facts of the world upon which the obligations rest, and/or premises from which deductions about possibilities may be drawn. The epistemic reading is associated with a background of what the speaker knows and what evidence the speaker has, while the deontic reading is associated with a background of what the circumstances are or what the obligations/laws are. I.e. the deontic is associated with world-truths irrespective of speaker knowledge, while the epistemic is associated with speaker knowledge irrespective of world-truths.

How are these differences encoded in syntax? The classical analysis, due to Jackendoff (1972) and Brennan (1993), suggests that modality is mediated by syntactic structure: the deontic interpretation is mediated by control syntax whereas the epistemic interpretation is mediated by raising syntax. In (2), PRO results in the deontic interpretation, but the trace in (3), originating in the subordinate clause and raising to the main clause, results in the epistemic interpretation.

(2) John must PRO be there by 5 pm. \(\rightarrow\) deontic/obligation
(3) John must John-\textit{t} be there by 5 pm. \(\rightarrow\) epistemic/possibility
(4) There may be singing but no dancing on the premises.
A more recent analysis, due to Wurmbrand (1999), Bhatt (1997) and others suggests that all modals involve raising syntax. Here the intrinsic modal force can be modified by conversational backgrounds (following Kratzer). Consider (4) (from Wurmbrand, 1999) which involves expletive “there” that is indicative of raising. Both epistemic and deontic interpretations are possible. On an epistemic reading, it is possible that people at the party will sing but not dance. On the deontic reading, singing is allowed, but dancing is not allowed.

Many acquisition studies find that the deontic use of modals is acquired earlier than the epistemic. Deontic meanings are the first to appear in spontaneous production, by 2-3 years. Epistemic meanings follow later, and full mastery is achieved by 8 years (e.g. Bascelli and Barbieri, 2002).

In comprehension experiments, children are shown to know relative modal force by about 5-6 years (e.g. Hirst & Weil, 1982). In such experiments children are presented with two characters, X and Y, where each has a box, but only one has a toy hidden under it. X says “There MUST be a peanut under the box,” and Y says “There MAY be a peanut under the box.” The question to the child is: “Who do you believe? Where is the toy?” Children can detect the relative modal forces, and by 5-6 years they reliably understand that “must” is relatively stronger than “may,” and correctly find the hidden toy. Notably, few studies explicitly control for epistemic or deontic contexts, independently of the intrinsic modal force.

What can account for the acquisition patterns? One possible line of analysis is syntactic. On the classical account of modal syntax where the deontic interpretation is mediated by control syntax and the epistemic interpretation is mediated by raising syntax, children should have trouble with epistemic interpretations, since raising structures are very late developing, coming in at 7-8 years of age (Hirsch & Wexler, 2007). In contrast, given this analysis, they should have no trouble with deontic interpretations, since children acquire control structures early (Wexler, 1992). On the more recent account of modal syntax, where all modals involve raising syntax, children ought to have trouble with all modals, unless they reinterpret raised structures as control structures, in which case they could only get the deontic interpretation, because the epistemic is incompatible with control syntax. Given arguments that children reinterpret unaccusative verbs as unergatives (Borer & Wexler, 1992; Babyonyshev et al, 2001), and children reinterpret verbal passive structures as adjectival passives (Borer & Wexler, 1987; Hirsch & Wexler, 2006), it is possible that they also reinterpret raising structures as control. Hirsch & Wexler (2007), in fact, argued that children interpret sentences with raising seem as control structures (their think analysis), even when there is no experiencer phrase. Regardless of what the true syntax of modals is, children’s underdeveloped grammars leave them unable to entertain epistemic interpretations.

Alternative accounts for children’s difficulties with epistemic modality evolve along a non-syntactic line. Papafragou (1998) proposes that epistemic modalities
involve greater metacognitive (theory of mind) representations than deontic modalities, since they require awareness of speaker beliefs. Theory of mind (TOM) is the ability to be aware of the mental states of other people and to differentiate others’ mental states from one’s own, which comes in around 4-5 years of age.

In a typical TOM task, one character (Mary) hides an object (e.g. chocolate), leaves the room, and another character displaces the object. The question is where will Mary look for the chocolate? Those who have TOM will correctly say that she will look for it where she thinks it is. Those who do not have TOM will incorrectly say that Mary will look for chocolate where the chocolate actually is, thereby indicating that these children cannot represent Mary’s false belief.

Thus it is possible that children cannot get the epistemic meaning not for syntactic, but for conceptual reasons: they may be incapable of conceptualizing that someone would be making a statement based on their mental state.

Notably, many TOM tasks involve modal “will.” The question in these tasks usually is, “Where will the character look for the object?” If children do not know epistemic modality (for reasons independent of theory of mind), and children think that “will” has only a deontic interpretation, they will answer the question of where Mary ought to look for her chocolate, given the state of the world. Children will not answer the intended question that involves the epistemic “will”: where Mary will actually look for her chocolate, given her knowledge. While it has been shown that this level of theory of mind comes in around 4-5 years of age in nonverbal tasks (Call & Tomasello, 1999; Colle, Baron-Cohen, & Hill, 2007), it is an interesting idea that children’s incorrect, deontic interpretation of the modal “will” might influence their answers. Thus, we should investigate children’s knowledge of the possible interpretations of “will” and other modals, in a task that tries to minimize the presence of false belief, and to understand what children know.

Our goal in this study is to investigate whether children are sensitive to the contextually driven differences in modal interpretations. We also investigate a new method for testing modal understanding. Ultimately, our goal is to test interpretations of modals in parallel with theory of mind and syntactic ability, but this is beyond the scope of the present study.

**Method**

We investigated children’s and adults’ sensitivity to the context-driven ambiguity in the modal verbs “will,” “can,” and “may.” We presented participants with video scenarios that created deontic or epistemic conversational backgrounds for the modals. Their predictions of the outcomes were recorded. After an expected (fulfilled modal force) or an unexpected (unfulfilled) outcome, we asked if they were surprised by the outcome, and whether the character saying the modal lied.

Participants viewed four introduction scenarios (straightforward scenes which
tested participants’ sensitivity to fulfilled/unfulfilled outcomes), and eight test scenarios (two each for “will,” “may,” “can,” and “must”) that differed in their conversational backgrounds (epistemic or deontic) and their outcomes (fulfilled or unfulfilled). Results for “must” are excluded from the present analysis for reasons of incongruity of scenarios. The storylines and conditions did not repeat within participants (since we found in piloting that once young children see a storyline with an outcome, they cannot change their mind if the same storyline is presented with a different outcome), so a complete data set consisted of 4 participants.

Let us summarize the reasoning for our method. In a deontic context, fulfillment of action is expected. In case of unfulfillment, the statement was “wrong,” hence participants should be surprised and should indicate “lying.” In an epistemic context, fulfillment is possible, but not necessary, because the statement was about the speaker’s belief rather than a fact about the world. In case of unfulfillment, participant may or may not be surprised, but should not indicate “lying.” Key conditions are EU (epistemic unfulfilled) vs DU (deontic unfulfilled).

Here is an example of a DU test condition (deontic context, unfulfilled outcome). A mom and child are in a park. The child says, “Look, Mom, a bicycle!” The mom answers, “You can ride the bicycle, just be careful.” This is the key modal sentence: here the mom is giving the child permission to ride the bike. Then a black screen comes up, and we ask the participant, “What happens next?” Under the deontic interpretation, a fulfilled outcome of the statement would be that the child rides the bicycle (because he is allowed to), and this is what most participants predict. In the next screen we see UNfulfillment of outcome: the mom says, “Actually, that bicycle looks dangerous, and you shouldn’t be riding without a helmet.” The parent does NOT let the kid ride the bike: she reverses her earlier promise. Now we ask participants whether they are surprised—in this condition, they should be, because the mom allowed the kid to ride the bike, then took the permission back. Then we ask participants whether they think the mom—the person uttering the modal sentence—lied (in this case, they might be likely to say “yes”).
Here is the corresponding EU test condition (epistemic context, unfulfilled outcome). A mom and a child are in a park. The child says, “Look, Mom, a bicycle!” The mom answers, “You can ride the bicycle, remember? Dad taught you how last summer.” Here the mom is stating that the child has the ability to ride the bike. Then a black screen comes up, and we ask the participants, “What happens next?” Under the epistemic interpretation, a fulfilled outcome of the statement would be that the child rides the bicycle (because he is able to), and this is what most participants predict. In the next screen we see UNfulfillment, but this time the child falls from the bike. Now we ask participants whether they are surprised—here they should be, because the kid was expected to ride the bike successfully, but he fell down. Then we ask participants if they think the mom—the person uttering the modal sentence—lied (in this case, they might be likely to say “no,” because mommy was merely estimating the kid’s ability, not giving permission).

To summarize our method, in both the deontic and epistemic scenarios the fulfilled outcomes should elicit little indication of surprise or deception from participants. In deontic scenarios, unfulfilled outcomes should elicit indications of surprise and of deception: a promise or a permission is made, but then taken back. In epistemic scenarios with unfulfilled outcome, participants may or may not be surprised, but they should rarely indicate deception: there is merely an estimate of ability or possibility which is inherently flexible.

<table>
<thead>
<tr>
<th>Group</th>
<th>Number</th>
<th>Mean age</th>
<th>Min. age</th>
<th>Max. age</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-4s</td>
<td>21</td>
<td>4.36</td>
<td>3.39</td>
<td>4.9</td>
</tr>
<tr>
<td>6-7s</td>
<td>13</td>
<td>6.86</td>
<td>6.0</td>
<td>7.8</td>
</tr>
<tr>
<td>8-11s</td>
<td>17</td>
<td>9.59</td>
<td>8.0</td>
<td>11.3</td>
</tr>
<tr>
<td>adults</td>
<td>39</td>
<td>22.25</td>
<td>19.10</td>
<td>33.49</td>
</tr>
</tbody>
</table>

Table 1. Details of participants recruited from Cambridge, MA

Results

MANOVA (with participants’ prediction, surprise, and deception (lie) indications as dependent variables, and modal (“can,” “may,” “will”), deontic/epistemic context, fulfilled/unfulfilled outcome and the four age groups as between subject factors) showed a significant effect of age group on all dependent variables (F(3,492)=5.5, p=.001 for prediction, F=26.5, p<.001 for surprise, F=4.1, p=.007 for lie), indicating a maturational trend. There was also a significant effect of un/fulfilled outcome on lie response (F(1,492)=130, p<.001); an effect approaching significance of deontic/epistemic context on lie (F(1,492)=3.3, p=.068); and a significant effect of the modal verb on prediction (F(2,492)=6.5, p=.002, with post hoc analysis indicating that is due to difference between “can”
and “may” (p=.024) and difference between “can” and “will” (p<.001)). The only interactions that reached significance were that between context and outcome for lie response (F(1,492)=6.6, p=.011); that between outcome and age group for lie response (F(3,492)=3.6, p=.014). The three-way interaction between age group, context and outcome was also significant for lie (F(3,492)=2.7, p=.043).

Consider the adults’ data first (Table 2). The expectations of fulfilled outcome were at 60%-80% for all contexts and modals, but for “will” such expectation occurred only in deontic contexts, with epistemic contexts eliciting expectations of unfulfillment at similar rates. Adults were surprised for “can” unfulfilled outcomes at 50% for both contexts and were not surprised in fulfilled outcomes. With “may” adults were more surprised with deontic unfulfilled outcomes (67%), and much less surprised with epistemic unfulfilled. For “will” adults showed more surprise in fulfilled outcomes (60%) than in unfulfilled outcomes. Notably adults showed more surprise in unfulfilled deontic “will” (56%) than in unfulfilled epistemic “will” (10%), indicating that for adults “will” is strongly associated with failure of follow-through. Overall, adults’ surprise indications showed their awareness of the contextually encoded differences between the epistemic and the deontic modality.

The youngest children (Table 2) had similar high rates of positive expectations as adults, but the 6-7s and the 8-11s were more skeptical than 3-4s and adults,

<table>
<thead>
<tr>
<th>Age group</th>
<th>“Expectation”</th>
<th>“Surprise”</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3-4s</td>
<td>6-7s</td>
</tr>
<tr>
<td>can - DF</td>
<td>0.78</td>
<td>0.29</td>
</tr>
<tr>
<td>can - DU</td>
<td>0.75</td>
<td>0.33</td>
</tr>
<tr>
<td>can - EF</td>
<td>0.67</td>
<td>0.67</td>
</tr>
<tr>
<td>can - EU</td>
<td>0.67</td>
<td>0.57</td>
</tr>
<tr>
<td>may - DF</td>
<td>0.44</td>
<td>0.43</td>
</tr>
<tr>
<td>may - DU</td>
<td>0.83</td>
<td>0.33</td>
</tr>
<tr>
<td>may - EF</td>
<td>0.75</td>
<td>0.33</td>
</tr>
<tr>
<td>may - EU</td>
<td>0.33</td>
<td>0.14</td>
</tr>
<tr>
<td>will - DF</td>
<td>0.44</td>
<td>0.29</td>
</tr>
<tr>
<td>will - DU</td>
<td>0.83</td>
<td>0.50</td>
</tr>
<tr>
<td>will - EF</td>
<td>0.67</td>
<td>0.33</td>
</tr>
<tr>
<td>will - EU</td>
<td>0.44</td>
<td>0.71</td>
</tr>
<tr>
<td>average</td>
<td>0.63</td>
<td>0.41</td>
</tr>
</tbody>
</table>

Table 2. Details of participants’ rates of expecting a fulfilled outcome and of indicating surprise at the outcome
primarily expecting unfulfilled outcomes. The 3-4s however were very often surprised by all outcomes, even in fulfilled conditions, as were the older children, indicating that estimation of surprise may not be the best metric for children.

Figures 2-4 indicate participants’ rates of whether they think the characters lied. All participants’ indications of lying, unlike indications of being surprised, were very informative as to the participants’ knowledge.

In all fulfilled outcomes, adults and 8-11s indicated deception only around 10%. In unfulfilled outcomes, adults and 8-11s indicated deception for deontic “can,” “may,” and “will” at significantly higher rates than for epistemic contexts, thereby indicating adults’ and oldest children’s sensitivity to the contextually determined epistemic or deontic modality. While both surprise and lie metrics indicate that adults are sensitive to the deontic vs. epistemic contexts with “can,” “may,” and “will,” the lie metric seems to provide a more clear indication of participants’ awareness of the modalities as set by the conversational background. Notably, the adult participants expected negative outcomes with “will” and were surprised at positive outcomes.

In all fulfilled outcomes with all modals, the 3-4s and 6-7s indicated deception at rates of 10-30%. In unfulfilled outcomes with “may” and “will,” these children showed deception around 40-65%, indicating that young children do pay attention to the factual story context, and the unfulfilled–fulfilled distinction. Unlike adults and older children, the 3-4s and 6-7s did not show differences between deontic unfulfilled and epistemic unfulfilled conditions for “may” and “will,” indicating that they are not yet aware of the epistemic vs. deontic distinction for these verbs. For “can” however, the 6-7s were able to indicate deception in the adult-like manner, unlike for “may” or “will.”

Figure 2. Details of participants’ indications of lying for “can”
Discussion

The present study illustrates a new method for testing children’s and adults’ modal comprehension and their intuitions about contextually determined modal interpretations. We find that young children seem unaware of contextually encoded differences between the epistemic and the deontic modality. Our data indicate that children can be aware of the epistemic–deontic distinctions as set by the context as early as 6 years with “can,” and by age 8 with “may” and “will.”
Our data also indicate that while young children do not differentiate the epistemic from the deontic modalities, their indications of lying with “may” and “will” on EU and DU conditions seem to track with older children’s and adults’ levels of indication of lying in EU condition. This is a puzzle, and it goes against the prediction that young children’s modal interpretations are all deontic by default. In other words, in our data the younger children do not seem to show the levels of deontic knowledge observed in the older children and adults.

Our overall results are consistent with previous findings in the literature. E.g. “can” is the first modal to be fluently used in production by 30 month olds (e.g. Wells, 1979), and in our study children are sensitive to its deontic vs. epistemic meanings before other verbs. However, Perkins (1983) found “can” still used in deontic modality by six year-olds, and only eight year-olds used “can” with “circumstantial possibility” reading, i.e. an epistemic reading that included a given situation and not the world laws. Similarly, Bascelli and Barbieri (2002) found that full understanding of modal verb strength of “must” vs. “may” in Italian is not achieved until around age eight, with five year-olds still interpreting “may” very much like “must.” Thus explicit awareness of modality and its interpretation relative to a situation is something that takes a while for children to acquire. This delay in acquisition is consistent with the idea that children’s deficit with the epistemic modality is mediated by their deficit in raising syntax. However this does not explain the difference between a relatively earlier acquisition of the epistemic interpretation of “can” vs. a relatively later acquisition of the epistemic interpretation of “may” and “will” in our data. It is possible that epistemic knowledge also relies on explicit awareness and recognition of possibilities and others’ states of mind. Unfortunately, our study did not include five year-olds who are likely to have theory of mind but lack raising syntax, and present a good age for disambiguating theories.

In further investigations we hope to fine-tune our task, include other modal verbs, and investigate children’s knowledge of (nonverbal) theory of mind in parallel to this task in 3-6 year-old children.

A final explanation for why children acquire epistemic modals after deontic modals can only be provided by studies of theory of mind, raising syntax, and development of modal verbs within individual children.

References


