Three-year-olds' understanding of desire reports is robust to conflict.

Kaitlyn P. Harrigan, Valentine Hacquard, Jeffrey Lidz

1. Introduction

Human beings explain each other’s behavior in terms of concepts like belief and desire. When we see, say, that Sally opened the cupboard, we infer that it was because she believed there was food inside and she wanted to eat. An important question in the domain of cognitive development centers around the origins of these concepts. When do children understand that other people have beliefs and desires?

In the domain of beliefs, there has been considerable controversy concerning the age at which children can be said to be sensitive to other people's beliefs (de Villiers 1995, 2005, 2007, de Villiers & de Villiers 2000, de Villiers & Pyers 2002, Johnson & Maratsos 1977, Lewis 2013, Perner et al. 2003, Wellman, Cross & Watson 2000, Wimmer & Perner 1983, and others). This question is somewhat vexed by the fact that different researchers set the standard for what counts as good evidence for belief attribution differently. All are agreed that the attribution of false beliefs to someone else is required to demonstrate mastery of the belief concept. But researchers disagree about what counts as evidence of false belief attribution, with some arguing that implicit measures indicate that the concept is in place early (Onishi & Baillargeon 2005, Southgate et al. 2007, Baillargeon et al. 2010, Kovacs et al. 2010), and others arguing that such measures may be reflective of heuristics or concepts with similar extensions, and that only explicit verbal measures count as definitive evidence for belief understanding (Apperly & Butterfill 2009, Butterfill & Apperly 2013, Heyes 2014, Thoermer et al. 2012). But certainly all are in agreement that when children pass explicit, verbal false belief tasks, the relevant concept must be in place. The gold
standard for demonstrating the full richness of belief is the use and comprehension of the words that make reference to that concept.

Likewise, in the domain of desires we can ask what the behavioral entailments associated with an adult-like desire concept are. And again we find the same kinds of issues circling around this question. A desire concept should support an ability to represent desires that are counterfactual, to represent two people as having different and possibly conflicting desires, to represent your own desires as distinct from the desires of others, to have desires about states of affairs, not just about objects, and to support the acquisition and comprehension of words like want, which express this concept.

Prior work using implicit measures suggests that children are sensitive to the desires of human agents from as young as 5 months of age (Woodward 1998, Woodward 1999, Woodward 2003), and that they are sensitive to the desires of others which differ from their own by 18 months of age (Repacholi & Gopnik 1997). However, it is not clear that success on these tasks requires a full-fledged concept of desire: perhaps young children grasp that others have desires that differ from their own, so long as these desires do not conflict with each other; perhaps children have an “objective” notion of desire about what is good, but not yet a desire concept that holds across counterfactual, subjective and conflicting desires. Studies testing children’s understanding of such a “subjective” concept of desire reveal considerable variability in children’s performance (e.g., Moore et al 2005, Rakoczy et al 2007, Rakoczy 2010).

In this paper, we probe young children’s concept of desire through their comprehension of desire reports using the verb want. We ask whether 3-year-olds can understand sentences reporting desires under three conditions: (1) counterfactual desires that relate to states of affairs that differ from reality; (2) 3rd person conflicts in which a desire report describes a situation in
which two people have different desires; and, 3) 1st person conflicts, in which a desire report describes a situation in which the child's desire differs from that described in the sentence. We show that three-year-olds can succeed with such desire reports, thus setting an upper bound on the acquisition of both desire language and the concept that supports it.

This paper investigates the robustness of three year olds’ competence with the verb want, and indirectly their ability to represent situations involving several types of conflicting desires, by testing whether comprehension of sentences with the verb want is adult-like in contexts that involve these conflicts. We present two experiments testing children’s understanding of want in conflicting situations. The first experiment tests children’s interpretation of want when it reports a desire that both conflicts with reality, and with another person’s desire. The second experiment tests children’s interpretation of want when it reports a desire that conflicts with the child’s own. We show that even with such tests, children are able to understand want before their fourth birthday. This suggests that three year olds have a robust adult-like understanding of want, which is not affected by additional conflicts. These findings indirectly show that three year olds have a robust understanding of the desire concept, and can represent even conflicting desires.

2. Previous Research

Many previous studies have explored children’s understanding of the desire concept and the verb want. Children as young as 2 have been shown to expect agents to act in accordance with their desires. Wellman & Woolley (1990) gave 2-year-olds information about characters’ desires in a story, and participants were asked to make predictions about each person’s actions. They found that 2-year-olds were successful at predicting actions related to people’s simple desires. Implicit tasks have been used to test even younger children’s desire representation using eye gaze
measures. Woodward (1998) tests 5- and 9-month olds in a looking-time paradigm. In her task, infants see an agent reach for one of two objects situated next to each other. After children are habituated to this action, the positions of the objects switch, and the agent either reaches for the same toy (now in a new position), or makes the same arm movement (reaching for a new toy). Infants look longer when the agent reaches for a new toy, even though in this case the agent’s arm motion is the same as in familiarization. They did not look longer when a non-human grasping device was used. Woodward finds that this pattern holds for both 9-month olds, and also (although with weaker results) for 5-month olds. This set of studies shows that by as young as 5 months old, infants can encode the actions of agents, attributing goals to them, and expect people to act in accordance with those goals. This could be seen as an early form of “desire representation,” at least with respect to objects, if not to states of affairs.

Other tasks look at slightly older children’s ability to represent conflicting desires. In a classic task, Repacholi & Gopnik (1997) examined whether 14- and 18-month-olds could appropriately represent the desires of an agent that potentially differed from the child’s own. In the task, the child is introduced to two familiar and distinctive foods: goldfish crackers and broccoli. After introducing each food, the researcher produces a salient response—either positive or negative—toward each of the foods, and then requests that the child give her some food. In order to control for the possibility that children assume that everyone has the same desires, or the possibility that the child will give the experimenter their own non-desired food in order to keep the desired food for themselves, experimenters manipulated whether the experimenter expressed preference for the same or the opposite food from the child’s preferred food. The authors found that most of the children preferred the crackers. They found that at 14-months, children were more likely to give the researcher the crackers, regardless of which food she had expressed
preference for. By 18 months, however, the children were more likely to give the experimenter the food that she preferred, regardless of whether it matched the child’s preference or not. This study shows that by 18 months, children seem to understand that different people may have desires that differ from their own. As Rakoczy et al (2007) point out, however, while the experimenter’s desire for broccoli differs from the child’s desire for crackers, the two desires are not incompatible. Perhaps children’s notion of desire that allows them to pass the task is a mere “objective” sense of desire about what’s generally good: broccoli for the experimenter, and crackers for the child. Hence, while the task shows that children understand that others can have desires that differ from their own, it doesn’t show that they understand that desires can conflict.

Note that these previous studies examine desires for an object. To probe children’s understanding of desires about states of affairs, researchers have turned to children’s ability to interpret want sentences. Here the results are mixed. Perner et al. (2003) tested children’s interpretation of want sentences in German-speaking children, and compared their understanding to think sentences. Children (2;5-4;5) saw six stories, each of which were accompanied by a drawing. For example, in one story, Mom and Dad were in one room and their son Andy was watching television in his bedroom. In the want condition, Mom asked Dad to see what Andy was doing. Dad asked Mom what Andy should do, and Mom answered, “Andy should go to bed.” Then the child was asked the want test question, shown in (1).

(1) Was will die Mutter, dass Andreas tut?

what wants the Mom, that Andy does

‘What does Mom want Andy to do?’
In the *think* condition, Dad asked Mom what Andy was doing and she answered, “Andy is going to bed.” Then the child was asked the *think* test question, shown in (2).

(2) Was glaubt die Mutter, dass Andreas tut?

*what thinks the Mom, that Andy does*

‘What does Mom think that Andy does?’

They found that children were much better at answering the questions with *want* than with *think*, and concluded that it is easier for children to remember discrepant desires than discrepant beliefs.

Note that English and German differ in the syntactic properties of the verbs *think* and *want*. In English, *think* takes a tensed complement, while *want* takes an untensed complement. By default, the untensed complement of *want* receives a future-orientation when the verb is eventive (as in (3)). The temporal interpretation of the complement of *think* depends on the tense in the complement: with a present tense, the belief is present-oriented (a future-orientation would require a future tense morpheme, as in ‘Mom thinks that Andy WILL go to bed’). In German, both *think* and *want* take tensed complements. However, with *want* (but not with *think*), it is still possible to get a future-orientation with a present tense in the complement (in fact it is the preferred interpretation). Sentence (2) can thus be interpreted in two ways. It can get the interpretation in (3), which sets up a conflict between desire and reality, but it also allows the (preferred) interpretation in (4), which is future-oriented and thus avoids a conflict with reality.

---

1 *Think* sentences in German can receive a future-oriented interpretation, but this requires an adverb, such as ‘later’. *Want* sentences get a future-oriented interpretation without an adverb by default.
What does Mom want Andy to be doing (right now)?

What does Mom want Andy to do (later)?

If children interpret (3) as meaning (4), Mom’s desire can still be satisfied if Andy’s future actions match her current desire, in which case there is no conflict between her desire and reality. Thus, this task did not require children to interpret want under truly conflicting conditions.

To test whether children’s early successes might be to a reliance on a merely “objective” concept of desirability, Rakoczy et al. (2007) (using a task first introduced by Litchermann 1991) tested whether children fail to correctly interpret want in cases where desires are subjective, with different people having non-compatible desires. They showed children (3;0-3;6) stories in which two characters “quarreled” about which of two either compatible or incompatible outcomes they preferred. In the compatible desires stories, two characters are each in boats. One character wants his boat to go to one location, the other wants her boat to go to another location. The boats then go to one of the two locations. The incompatible desires stories are the same, except that both characters are in a boat together, thus it is impossible for each character’s desire to be satisfied simultaneously. After the story, the children were asked the test questions shown in (5) and (6).

Susi wanted the boat to go where?

And Tom wanted the boat to go where?

Children succeeded on this task, suggesting that they can represent incompatible desires and interpret sentences with want before their fourth birthday. Note however that the future
orientation of (5) and (6) could still prevent a conflict between reality and the desire, and hence not provide a stringent test of children’s ability to represent incompatible desires. The question in (5) describes a past desire about an outcome future to this past desire time. Although the boat did go to one of the two locations at a time future to this desire time (namely, at the end of the story), the future is open, and it is possible that the boat could still subsequently go to a second location, and the desire be satisfied in the near future. To rule out this possibility, it is necessary to make explicit that the desire is about a *concurrent* state of affairs (see experiment 1 below).

A pilot study reported in de Villiers (2005) attempts to better control for the possibility of future-oriented readings with *want*. This study looked at children’s interpretation of *want* sentences with a participial complement in English. This structure eliminates the possibility for a future-oriented reading in the desire sentence, making this a better test of an actual present-oriented conflict. In this study, children saw pictures and heard stories where someone had an unfulfilled desire, and had to assess sentences such as in (7).

(7)  *Mom wants Bella playing on the computer.*

De Villiers found that children were successful in interpreting these sentences with *want*. Although the participial complement does not allow a future orientation, it is not a fully sentential complement. This may make interpretation of sentences like (7) easier, and these results consequently overestimate children’s knowledge of *want*.

A couple of studies probe children’s understanding of desires that conflict with their own. Moore *et al* (1995) looked at three year olds’ understanding of conflicting desires in a task in which they played a game against a puppet, “Fat Cat.” Both the child and Fat Cat had to solve
their own jigsaw puzzle for which they needed parts from a blue or red box. In each round, a card was drawn from a stack, turned around and shown to be either blue or red. Both players could then take a piece from the corresponding box. At first, both players needed pieces from the same box. However, there came a point where their needs diverged, and thus their desires for which color the card should be became incompatible. At this point children were asked three control questions and two test questions, shown in (8) and (9).

(8) *Which color card does Fat Cat want now?*

(9) *Which color card did you want last time?*

Only 7 of 20 children passed both test questions on the conflicting-desire task, leading Moore *et al* to conclude that when children are forced to represent incompatible and conflicting desires, they have difficulty interpreting sentences with *want*.

Rakoczy and colleagues (2007, 2010) were concerned that the methodology used in the Moore *et al* study may have underestimated children’s knowledge. They point out that the task was very complex, and suggest that adjustments in methodology may improve children’s performance. Using a game format modeled on the Moore *et al* study, they tested both conflicting “third person desires”, where two puppets played against each other, and conflicting “first person desires”, where the child played against a puppet. In this task, children (3;0-3;6) worked together with a puppet to make a sticker book, but only one sticker could go inside. A “chance machine,” out of which a marble was dispensed, determined one of two sticker possibilities: one was an exciting sticker, and one was a boring sticker. The child always
preferred the more exciting sticker, and the puppet expressed interest in the other sticker. The children were asked the test questions shown in (10)-(11).

(10) *You want the marble to roll where?*

(11) *Rudi [puppet] wants the marble to roll where?*

Rakoczy and colleges found that improving the methodology did improve performance, compared to the original Moore et al study, with accuracy around 55% overall. There were no differences between first and third person conditions. However, children’s performance was still not adult-like. There may, however, be further methodological concerns in this improved task. As in the Moore et al study, there were only one or two critical trials. Additionally, the game was still fairly complex for preschoolers, and the authors did not report training on the task or having a criterion determining whether children understood the rules of the game. This potentially underestimates children’s understanding of *want*, as failures observed in these cases may have been due to children’s lack of understanding of the rules of the game. To get at children’s understanding of *want* in conflicting situations, it is critical to exclude children who do not understand the rules of the game, to ensure that errors are due to difficulty processing or understanding conflicting desires, and not to confusion about the rules of the game. To sum up, it remains an open question whether children can understand *want* sentences (and the underlying desire concept), in cases of conflicting desires.

The previous tasks looking at children’s interpretation of *want* can be improved in several ways, in order to license better inferences about children’s knowledge. The first is in the temporal orientation of the attitude verb. As discussed above, *want* sentences often get a future-
oriented reading, and thus do not necessarily report a desire that *conflicts* with reality, or with another person’s desire. Experiment 1 controls for the temporal orientation of the complement of *want*, allowing us to test children’s understanding of *want* when there is a conflict between the reported desire and reality.

In experiment 2, we test children’s understanding of *want* sentences used to report desires that conflict with their own. We improve on the methodology of the previous studies by including a thorough training in those aspects of the task that do not have to do with desire, and excluding participants who do not understand the rules of the game. This ensures that we will not underestimate children’s understanding of *want* sentences.

3. Experiment 1: Conflict with Reality

Experiment 1 tests *want* in sentences that force a present-orientation, and thus describe desires that potentially conflict with reality and with another character’s desires.

3.1. Subjects

Participants were 44 children aged 3;0 to 4;0 (mean = 3;8). 16 additional children were excluded from the task, either due to *yes* or *no* biased responses on the task, or parental interference. Children in all three studies were recruited from the College Park, Maryland area, and were reported by their parents to be monolingual speakers of English. Participants were recruited via telephone or email from the University of Maryland Infant Studies Database.
3.2. Design and Materials

Experiment 1 was a *Truth Value Judgment Task* (TVJT), which requires children to correct sentences uttered by a “silly” puppet (Crain & Thornton 1998; Crain & McKee 1985). TVJT tasks gauge whether children at a given age pair certain linguistic stimuli to a given situation in an adult-like way, or whether their interpretation of the stimuli differs in some way from adult judgments. In this task, children listened to stories with pictures. They were told that a puppet who was “very silly and sometimes gets things wrong” was listening to the stories as well, and asked to tell the puppet whether he was right or wrong after every utterance. Each child saw eight stories. After each story the puppet uttered two sentences: a filler sentence and a test sentence. The fillers were intended to ensure that the child was paying attention and had a basic understanding about what happened in the story. Test sentences had a sentential complement which forced a present orientation by using a progressive (‘be ___ING’), and the temporal modifier ‘right now’ (see (12)).

(12) *Mom wants Megan to be sitting in the grocery cart right now.*

There were a total of 8 stories, each with two different versions. Between subjects we manipulated whether the stories contained a desire that conflicted with reality (*CONFLICT* condition) or not (*NO CONFLICT* condition). Each of the stories described a situation in which a child starts out doing a given activity, and then an adult asks the child to either continue doing the same activity (*STAY* condition) or switch to a new activity (*SWITCH* condition). This manipulation was within subjects. Half of the conflict stories were *STAY* stories, and the other
half were switch stories. For the participants in the no conflict condition, the stories that were stay stories in the conflict condition were switch stories, and the stories that were switch stories in the conflict condition were switch stories. This ensured that each story was equally plausible as a conflict or no conflict situation, as well as a switch or stay scenario. Additionally, it ensured that in both the conflict and no conflict conditions, the character did not always start and end doing the same activity. We also manipulated the truth-value of the test sentences within subjects. Table 1 illustrates the within and between-subjects factors in experiment 1.

**Table 1: Within and between subjects factors in Experiment 1**

<table>
<thead>
<tr>
<th>Conflict/No Conflict (between subjects)</th>
<th>Switch/Stay (within subjects)</th>
<th>Truth Value (within subjects)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conflict</td>
<td>Switch</td>
<td>True</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
</tr>
<tr>
<td>Conflict</td>
<td>Stay</td>
<td>True</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
</tr>
<tr>
<td>No conflict</td>
<td>Switch</td>
<td>True</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
</tr>
<tr>
<td>No conflict</td>
<td>Stay</td>
<td>True</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
</tr>
</tbody>
</table>

### 3.2.1. Sample Story

The stories consisted of four pictures each. Each picture represented about one sentence of a story. A sample of the text of one story is laid out in Figure 1.

**Figure 1: Experiment 1 Sample Story**

<table>
<thead>
<tr>
<th>Introduction Phase</th>
<th>Megan is at the grocery store with her mom. She’s sitting in the cart while her mom shops.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stay/Switch Phase</td>
<td>STAY: Megan’s mom says,</td>
</tr>
<tr>
<td></td>
<td>SWITCH: Megan’s mom says,</td>
</tr>
<tr>
<td>Conflict/No Conflict Phase</td>
<td>Stay/Conflict: Mom leaves, and Megan says to herself, “I know my mom said I should stay in the cart, but I’d like to get out and go get some cereal, so I will!”</td>
</tr>
<tr>
<td>Outcome Phase</td>
<td>Stay/Conflict: So she climbs out of the cart to go get some cereal.</td>
</tr>
<tr>
<td>Test Sentence</td>
<td>Stay/Conflict: So she stays right there in the cart.</td>
</tr>
<tr>
<td>True</td>
<td>Mom wants Megan to be sitting in the cart right now!</td>
</tr>
<tr>
<td>False</td>
<td>Mom wants Megan to be getting cereal right now!</td>
</tr>
<tr>
<td>True</td>
<td>Mom wants Megan to be sitting in the cart right now!</td>
</tr>
<tr>
<td>False</td>
<td>Mom wants Megan to be getting cereal right now!</td>
</tr>
</tbody>
</table>

### 3.2.2. Procedure

Each child was tested in a quiet room with two experimenters. One experimenter told the child the stories and showed her the pictures, while a second experimenter controlled the puppet and uttered the filler and test sentences. The second experimenter also coded the child’s responses. Permission was obtained from parents to video record each subject for an additional round of coding off-line.
The experiment began with the child being introduced to a silly puppet, “Froggy.” The experimenters were somewhat flexible with the script, adjusting to each child’s level of attention, but followed the following script fairly closely:

“We're going to be looking at some pictures and hearing some stories that go along with them. And our friend Froggy is going to listen to the stories with us, ok? And after we hear a story, Froggy's gonna try to tell us what happened in the story. But sometimes, he's not a very good listener. And so sometimes when he tells us, he might get it wrong, ok? And you get to tell us whether Froggy was right or wrong. Does that sound like a good plan? OK, so listen carefully, because he says silly stuff sometimes!”

The child then practiced interacting with Froggy. First Froggy named a few items, and the child practiced telling him yes and no. The child was corrected during this practice phase if they did not correctly tell Froggy yes and no. Then the child was told two very simple stories, and practiced responding to sentences Froggy said about the stories. Froggy was correct once and incorrect once. Again, the child was corrected if she did not respond correctly to Froggy’s sentences.

During each test trial, experimenter 1 read the story and showed the child the pictures. After each story, experimenter 1 turned to Froggy and asked “what happened in that story, Froggy?” Then Froggy uttered the filler sentence, after which the experimenter turned to the child and asked “did Froggy get it right?” Then the child either responded yes or no. After the filler, Froggy uttered the test sentence, and the child was again asked whether or not he was right and
given the chance to respond yes or no. Experimenter 1 gave feedback to Froggy that was in accordance with how the child had responded—“good job, Froggy!” when the child said that Froggy was correct, and—“oh, silly Froggy! Try again next time!” when the child said that Froggy was incorrect. The entire experiment took around 8-10 minutes per child.

3.3. Results

3.3.1. Coding

Children’s responses were coded online by the second experimenter. Four out of the sixty videos were coded by a second experimenter offline, because coding did not happen online. Responses were coded as yes, or no. One response (out of 1,080 total responses) was unintelligible. An additional 25% (11 videos) were coded offline by an additional coder. We found 99.4 percentage of agreement between coders (Cohen’s Kappa = .989).

3.3.2. Filler Accuracy

The fillers were designed to ensure that children were listening to the story. They did not rely on understanding want. Children who answered either all yes or all no to 15 out of 16 total items were excluded from analysis. Twelve children were excluded due to yes-biased responses (20%). Three children were excluded due to no-biased responses (5%). One additional child was rejected due to parental interference. The age range that we were testing for this study is quite
young for the TVJT paradigm, which likely contributed to the high number of children with yes- or no-biases.

3.3.3. Truth-Value Judgments

The results for each condition are shown in Table 2.

<table>
<thead>
<tr>
<th>CONFLICT/NO CONFLICT (between subjects)</th>
<th>SWITCH/STAY (within subjects)</th>
<th>Truth Value (within subjects)</th>
<th>% Yes Responses children 3;0-4;0</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONFLICT</td>
<td>Switch</td>
<td>True</td>
<td>93.2%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>13.6%</td>
</tr>
<tr>
<td>CONFLICT</td>
<td>Stay</td>
<td>True</td>
<td>79.1%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>17.8%</td>
</tr>
<tr>
<td>NO CONFLICT</td>
<td>Switch</td>
<td>True</td>
<td>79.5%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>11.4%</td>
</tr>
<tr>
<td>NO CONFLICT</td>
<td>Stay</td>
<td>True</td>
<td>93.2%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>False</td>
<td>15.9%</td>
</tr>
</tbody>
</table>

Because our data is categorical, we used mixed-effect logit models (Bates & Sakar, 2007) to analyze the results (Baayen, 2007; Jaeger, 2008). We ran a mixed-effects logit model (with random intercept) with 'yes' responses as the dependent measure and with fixed effects for Condition, Switch, Target and all interactions between these, and with subject and item as random effects. Using the function glmer and drop1 from R's Design library (R Core Team, 2013), we were able to compare the full model against all of its subset models derived by removing exactly one predictor. The model without Target has considerably lower likelihood ($X^2(1) = 6.41, p<0.001$), but the models that removed Condition ($X^2 (1) = -0.67, p = 0.504$) and Switch ($X^2 (1) = -1.82, p = 0.068$) did not. Thus Target contributes significant information to the model. We also found that the interaction between Condition and Switch contributed significant
information to the model ($X^2 (1) = 1.99, p = 0.047$), but no other interactions did so. The best-fit model was one in which only Target and the interaction between Condition and Switch predicted children’s ‘yes’ responses, with the effect of Target being much greater than that of the Condition/Switch interaction.\(^2\) Children were significantly more likely to respond yes to the yes-target items, and no to the no-target items, regardless of whether the item was a CONFLICT or a NO CONFLICT item, or whether it was a SWITCH or STAY item. Results from Experiment 1 indicate that three years olds understand want correctly, even when it is present-oriented and describes a desire that conflicts with reality.

3.4. Experiment 1 Discussion

In this study, we set up situations in which characters had desires that conflicted both with reality, as well as with another character’s desires. We used linguistic stimuli that disallowed future-oriented readings of want sentences. We find that three year old children are still adult-like in interpreting want sentences. This shows that three year olds have a robust representation for the meaning of want, and can correctly understand want sentences when they ascribe counterfactual desires.

4. Experiment 2

Previous results looking at three-year-olds’ ability to understand reports of desires that conflict with their own are inconsistent, and raise several methodological concerns. Experiment 2

\(^2\) We used the same procedure as described in the text with the full data set (including children with yes/no biases) and that same contributors were present in the best fitting model.
remedies these concerns. We set up a task where the child plays a game with a puppet, in which their desires sometimes conflict, and then the child is asked about those conflicting desires. This task requires children to maintain in memory both their own desire and the puppet’s desire.

4.1. Subjects

Participants were 43 children aged 3;0 to 4;0 (mean = 3;8). 18 additional children were excluded from the task. 16 did not pass the practice, and two did not finish the task.

4.2. Design and Materials

Experiment 2 was set up like a game. The child played with a puppet, Froggy, while another puppet, Booboo, was “learning” and said things about the game. The child’s job was to tell Booboo whether he was right or wrong. The experimenter flipped colored cards, and depending on the color of the card, the outcome was either positive for Froggy, the child, both of them, or neither of them (the positive outcome was that someone got to stamp). This set-up induced desires in the child, which sometimes conflicted with the puppet’s. Booboo then uttered test sentences. The child was told that Booboo was “not very good at colors and sometimes gets things mixed up”, and was asked to tell Booboo whether he was right or wrong after every utterance. Each child participated in four, eight or twelve practice trials and sixteen test trials. The purpose of the practice trials was to teach the child how the game is played, and to have a measure to exclude children who did not understand how the game worked. The practice trials involved Booboo uttering a sentence that the child had to correct, just like the test trials, but the
sentences were about the structure of the game, not a desire. An example of a practice question is shown in (14).

(13) **Oh, I see how the game works! When it’s green, Froggy gets to stamp!**

After the child corrected Booboo, the experimenter flipped the card, and asked the child to tell everyone who got to stamp based on the color of the card. This ensured that participants understood the rules of the game and were comfortable playing before the test trials started. Each child had at least four and at most twelve practice trials. We continued with the practice until the child got four in a row correct, and then we moved on to the test trials. If the child did twelve practice items and did not learn how the game worked, they did not move on to the test trials and were excluded from analysis.

Each test trial consisted of two test sentences (examples in (14) and (15)), one about Froggy’s desire and one about the child’s desire.

(14) **Froggy wants the card to be green!**

(15) **You want the card to be green!**

After Booboo uttered the test sentences and the child said whether he was right or wrong, the experimenter flipped the next card on the pile. The experimenter then asked the child the filler question, which was about the outcome based on color (example in (16)), and then the appropriate player(s) stamped their paper.
Oh! We got green! Who gets to stamp when we get a green card?

The fillers were intended to ensure that the child was paying attention and understood how the game worked. Children were encouraged to try again if they got the fillers incorrect. This happened very rarely during the game.

This study was a 2x2x2 design, and all manipulations were within subjects. We manipulated whether we were asking about a desire with a conflict (CONFLICT condition) or not (NO CONFLICT condition). Additionally, we manipulated whether we were asking about a positive outcome from the child’s perspective (POSITIVE condition) or not (NEGATIVE condition). We also manipulated whose desire we were asking about, the child’s (CHILD DESIRE condition) or Froggy’s (FROGGY DESIRE condition). We counterbalanced order between subjects. Table 3 illustrates the within-subjects factors in experiment 2.

<table>
<thead>
<tr>
<th>CONFLICT/NO CONFLICT</th>
<th>POSITIVE/NEGATIVE</th>
<th>SENTENCE (Froggy v. Child desire)</th>
<th>Truth Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONFLICT</td>
<td>POSITIVE</td>
<td>&quot;FROGGY WANTS&quot;</td>
<td>False</td>
</tr>
<tr>
<td></td>
<td>(child stamps)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>&quot;YOU WANT&quot;</td>
<td></td>
<td>True</td>
</tr>
<tr>
<td>CONFLICT</td>
<td>NEGATIVE</td>
<td>&quot;FROGGY WANTS&quot;</td>
<td>True</td>
</tr>
<tr>
<td></td>
<td>(Froggy stamps)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>&quot;YOU WANT&quot;</td>
<td></td>
<td>False</td>
</tr>
<tr>
<td>NO CONFLICT</td>
<td>POSITIVE</td>
<td>&quot;FROGGY WANTS&quot;</td>
<td>True?</td>
</tr>
<tr>
<td></td>
<td>(Child &amp; Froggy stamp)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>&quot;YOU WANT&quot;</td>
<td></td>
<td>True?</td>
</tr>
<tr>
<td>NO CONFLICT</td>
<td>NEGATIVE</td>
<td>&quot;FROGGY WANTS&quot;</td>
<td>False</td>
</tr>
<tr>
<td></td>
<td>(No one stamps)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>&quot;YOU WANT&quot;</td>
<td></td>
<td>False</td>
</tr>
</tbody>
</table>

Color and outcome were counterbalanced within subjects, so that every color and every outcome occurred an equal amount of times during the game. We also rotated which colors were
paired with which outcomes throughout the game, to ensure that a color bias would not affect the results. We rotated a total of four times during the game, after every four sets of test questions. Within each of the four blocks, each color and each outcome occurred one time. A schematic of a trial is shown in Table 4.

**Table 4: Sample of trial in Experiment 2**

<table>
<thead>
<tr>
<th><strong>Booboo</strong>: Froggy wants the card to be blue!</th>
<th><strong>E1</strong>: Did Booboo get it right?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Child</strong>: yes/no</td>
<td></td>
</tr>
<tr>
<td><strong>E1</strong>: Good job/try again, Booboo!</td>
<td></td>
</tr>
<tr>
<td><strong>Booboo</strong>: You want the card to be green!</td>
<td></td>
</tr>
<tr>
<td><strong>E1</strong>: Did Booboo get it right?</td>
<td></td>
</tr>
<tr>
<td><strong>Child</strong>: yes/no</td>
<td></td>
</tr>
<tr>
<td><strong>E1</strong>: Good job/try again, Booboo! OK, let’s flip! [<strong>E1 flips card</strong>] Oh! we got green! What happens when we get a green card?</td>
<td></td>
</tr>
<tr>
<td><strong>Child</strong>:   gets to stamp!</td>
<td></td>
</tr>
<tr>
<td><strong>E1</strong>: Good job! Let’s stamp! … ok, Booboo, tell us something about the game …</td>
<td></td>
</tr>
</tbody>
</table>

4.2.1. Procedure

The procedure for Experiment 2 was the same as for Experiment 1. Experiment 2 began with the child being introduced to “Froggy,” with whom they would be playing the game. The experimenters were somewhat flexible with the script, adjusting to each child’s level of attention, but the experimenters followed the following script fairly closely:

“We’re going to play a game with Froggy today where we get to flip cards! And every time that we flip a card, someone gets to put a stamp on their paper. Froggy loves stamps… do you like stamps? OK, so every card that we flip has a color,
and we can look at the board [experimenter points] to see who gets to stamp when we flip that color. OK, so when we flip a green card, you and Froggy both get to stamp. When we flip a brown\(^3\) card just Froggy gets to stamp. When we flip a blue card just you get to stamp. And when we flip a pink card no one gets to stamp.”

Then the child was introduced to the silly puppet, Booboo:

“OK, one more thing! Froggy’s friend Booboo the baboon wants to learn how to play the game, so he’s going to watch us play. But he’s not very good at colors, so sometimes he gets things mixed up! Sometimes he’s going to try to tell us something about how the game works, but he might get it wrong, and your job is going to be to help him out and tell him whether he’s right or wrong so he can learn how to play the game. How does that sound?”

The child then practiced interacting with Booboo. First Booboo practiced naming colors, half of which he got right and half of which he got wrong, and the child practiced telling him yes and no. The child was corrected during this color practice phase if they did not correctly tell Booboo yes and no. Then we moved on to the practice phase, where the child saw between four and twelve practice trials. Again, during this phase the child was corrected when they made an error.

\(^3\) The brown card was occasionally named by the child as a different color, e.g., tan. In these cases we adapted the name of the color based on what the child said. Children had no difficulty with any other colors used in the game.
After sufficient practice, we moved on to the test phase. During each test trial, Booboo uttered each test sentence, and experimenter 1 asked the child if Booboo was right. Then the child gave her response. Experimenter 1 gave feedback to Booboo that was in accordance with how the child had responded, as in the previous experiment. After both test sentences, experimenter 1 flipped the next card on the pile and asked the child the filler question. After the child responded, the appropriate player(s) stamped their paper, and we moved on to the next test trial. The entire experiment took around 20 minutes per child.

4.3. Results

4.3.1. Coding

Children’s responses were coded online by the second experimenter. Four out of the 60 videos were coded by a second experimenter offline, because coding did not happen online. Responses were coded as yes, or no. An additional 25% (15 videos) were coded offline by an additional coder. We found 97.9 percent agreement (Cohen’s Kappa = .952.)

4.3.2. Practice and Filler Accuracy

This experiment included an extensive training and practice section. There were four practice items. Children had to get all four in a row right to be included. We went through all the items either once (four items), twice (eight items) or three times (12 items). This means that children had a minimum of four practice items, and a maximum of 12. 16 out of the total of 63 children
tested (25%) did not pass the practice after three rounds and were thus excluded from the rest of the experiment and analysis. Of the 40 included subjects, 20 of them went through the practice items once, 11 went through the practice items twice, and 2 went through the practice items three times. Overall accuracy was not different for the groups of children who did the practice items once or twice. The two children who did the practice items three times had lower accuracy (Table 5). This may be related to an overall difficulty to perform on this task, or it may be due to the fact that for these children the task took a longer amount of time, thus, they may have been more fatigued by the time they got to the test questions, which may have contributed to their overall accuracy. For the additional six children, the practice session was not recorded due to experimenter error.

<table>
<thead>
<tr>
<th>Number of practice items</th>
<th>% Accuracy overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>75%</td>
</tr>
<tr>
<td>8</td>
<td>74.4%</td>
</tr>
<tr>
<td>12</td>
<td>56%</td>
</tr>
<tr>
<td>no data</td>
<td>72.3%</td>
</tr>
</tbody>
</table>

The fillers were designed as a control to ensure that children were paying attention during the game as well as to keep them engaged. Once children were included after the practice phase, they rarely had any difficulty correctly saying who got to stamp after each card flip, and asking the child after each card flip was an extremely natural question during the game. If they incorrectly answered the filler, they were directed to try again.

4.3.3. Truth-Value Judgments
The results for each condition are shown in Table 6.

### Table 6: Percent Accuracy by Condition for Experiment 2

<table>
<thead>
<tr>
<th>Conflict/No Conflict (within subjects)</th>
<th>Outcome (within subjects)</th>
<th>Sentence (Froggy v. Child desire)</th>
<th>Target</th>
<th>% Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Conflict</strong></td>
<td>Positive (child stamps)</td>
<td>“FROGGY WANTS”</td>
<td>No</td>
<td>13%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>“YOU WANT”</td>
<td>Yes</td>
<td>80%</td>
</tr>
<tr>
<td><strong>Conflict</strong></td>
<td>Negative (Froggy stamps)</td>
<td>“FROGGY WANTS”</td>
<td>Yes</td>
<td>74%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>“YOU WANT”</td>
<td>No</td>
<td>15%</td>
</tr>
<tr>
<td><strong>No Conflict</strong></td>
<td>Positive (Child &amp; Froggy stamp)</td>
<td>“FROGGY WANTS”</td>
<td>Yes?</td>
<td>40%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>“YOU WANT”</td>
<td>Yes?</td>
<td>41%</td>
</tr>
<tr>
<td><strong>No Conflict</strong></td>
<td>Negative (No one stamps)</td>
<td>“FROGGY WANTS”</td>
<td>No</td>
<td>9%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>“YOU WANT”</td>
<td>No</td>
<td>6%</td>
</tr>
</tbody>
</table>

As in experiment 1, we used mixed-effect logit models (Bates & Sakar, 2007) to analyze the present results. The reported models have random intercepts. We ran a mixed-effect logit model with ‘yes’ as the dependent measure, with Conflict, Outcome and Sentence type as fixed effects, and item and subject as random effects. We tested for each of the two-way interactions, testing the full model against all subset models that eliminate only one effect. We find a significant two-way interaction between interaction between Sentence type and Outcome \( \chi^2(1) = 105.5, p < 0.0001 \), and a significant, but much smaller, interaction between Sentence type and conflict \( \chi^2(1) = 7.59, p = 0.006 \). Children’s responded different based on whose desire was reported and whether that outcome was positive or negative, and their responses were much less influenced by the presence of a conflict. Results for all conditions are shown in Figure 2.

**Figure 2: Results of Experiment 2**
Children were successful in the conflict cases, whether they were being asked about either positive or negative outcomes. In the **NEGATIVE NO CONFLICT** condition, children were successful in saying that neither they nor Froggy wanted an outcome where no one got to stamp. In the **POSITIVE NO CONFLICT** condition, where both the child and Froggy got to stamp, some children played the game in a competitive way, responding that neither they nor Froggy wanted the outcome where both players got to stamp, while other participants were more charitable, responding that both they and Froggy wanted this outcome. The histogram in Figure 3 shows that children were normally choosing either a strategy of responding *yes* to this condition all the time, or *no* to this condition all the time. This indicates that they were not confused by this condition, but that different children simply differed in how competitive they chose to be with Froggy. This however does not affect whether a conflict with someone else’s desires impacts performance on interpreting *want*. 
Our results indicate that three year olds are adult-like in interpreting *want* even when they are asked to assess a character’s desire that conflicts with their own, as long as they are given adequate training and opportunity to understand the rules of the “game” used to test this ability.

One potential problem with this study is the possibility that the “desire” that we induced in the child in the game is not a real one. During the game, the child does not choose which color is linked to the opportunity for them to stamp, and the color-outcome pairs are constantly changing every four trials. The rotation of colors and outcomes was an important manipulation to ensure that the results were not affected by color biases, but it’s possible that this manipulation had an effect on the strength of participants’ desires. If the participants in this task did not have a real desire about the color of the card, this task could not successfully test conflicting desires, since the child would not have to override their own desire about color. In order to control for this possibility, we ran a small sample (n=8) on a slight modification of experiment 2, designed to induce a more deeply rooted desire in children. In this manipulation, the child chose which color led to which outcome, and the color/outcome pairings did not change during the course of the
game. This allowed the child to have a deeper-rooted desire, because they had a say over which color would mean a positive outcome for them, and this stayed consistent throughout the game. All other aspects of the task and analysis were identical to experiment 2. We find the same pattern of performance in this pilot study, indicating that success in Experiment 2 is not due to the lack of a real desire in the participants.

4.4. Discussion of Experiment 2

Previous tests of want either did not require the child to evaluate desires that directly conflicted with their own, or the methodology was problematic and results were inconclusive. In order to fully understand children’s early understanding of the verb want, it was necessary to further explore children’s interpretations of sentences with want under these conditions. Experiment 2 improved on previous methodology, and showed that three year olds responded in adult-like fashion to want sentences used to report desires that conflicted with their own desires. This suggests that children’s knowledge of the verb want and the underlying concept is robust at this age, and not disrupted by complex situations involving processing conflicting desires.

Although most children learned the rules of the game, given the complexity of the task and the number of children excluded, it is possible that this task was measuring more than just interpreting want. In order to truly test interpretation of want without the confound of interference from errors caused by lack of understanding of the rules of the game, it was critical to include only children who can show, outside the context of interpreting conflicting desires, that they understand the rules of the game. However, it is possible that the children who were included in the test items of our task are children who are more advanced in other cognitive
systems, such as executive function or memory. Given the evidence that children with higher executive function are better at tasks involving processing mental states (Fitzke et al 2014, Rakoczy 2010, German & Leslie 2000, Leslie and Polizzi 1998, Leslie and Thaiss, 1992, Leslie & Roth 1993, Roth & Leslie 1998, Leslie, German & Polizzi 2005, Devine & Hughes 2014), it is possible that the sample of children who succeeded in learning the rules of our game were children who may have an independent advantage in understanding the test sentences. Additionally, some evidence suggests that the ability to participate in a competitive game is correlated with children’s performance on the traditional false belief task (Priewasser et al. 2013). If so, the sample of children who were able to pass the training on this task may also have been more advanced in attributing mental states. Experiment 2 does demonstrate, however, that as long as children are given adequate training and are able to understand the rules of this “game”, they are successful in interpreting want sentences which report conflicting desires.

5. General Discussion

The results from our experiments show that three year olds understand want sentences in an adult-like way, even when they report counterfactual desire and conflicting desires, including desires that conflict with their own. This suggests that at least by age three, children have the underlying concept of desire necessary to understand such subjective desires.

These results also have implications for the vast literature looking at children’s understanding of the belief concept and the verb think. At the same age as we find success in interpreting want, children have notorious difficulty understanding think sentences when they report a false belief, that is, a belief that conflicts both with reality, and with their own beliefs.
Our results suggest that the source of children’s difficulty with think cannot be explained solely as difficulty processing a report of a mental state which conflicts with reality or with their own mental state. This suggests that the asymmetry observed in the acquisition of these two verbs is thus not due differences in the ways in which these verbs have been tested. Given that these conflicts cannot be the source of difficulty with think, more work needs to be done in order to explore the source of the difficulty.

One possible explanation for the asymmetry observed between the acquisition of want and think identifies this asymmetry with the parallel asymmetry in the development of the concepts that these verbs express (cf. Perner 1981, 1988, Perner et al. 2003, Perner et al. 2005, Perner & Ruffman 2005, Tardif & Wellman 2000; for an overview see Steglich-Petersen and Michael, to appear). According to this Conceptual Asymmetry, the desire concept appears earlier than the belief concept, which awaits the development of a full Theory of Mind, around age 4, as evidenced by children’s consistent failure at explicit false belief tasks before then (Wellman et al. 2001). If children, in fact, have developed, or have access to, the concept of desire before the concept of belief, it follows that children would acquire words that refer to the former earlier than the latter.

Our studies show that three year olds also have a robust knowledge of the verb want—even in cases where it reports counterfactual or conflicting desires, and so are consistent with an explanation attributing ease with want to an early conceptual advantage. However, the results of studies investigating the concept of belief have been mixed. While many tasks have shown difficulty with interpreting sentences with think, a number of recent studies with young infants suggest that children show understanding of belief very early when tested through implicit measures (Onishi & Baillargeon 2005, Southgate 2007, Kovács et al. 2010, Senju et al 2011,
Woodward 1998, 1999, 2003). While these results cast doubt on the conceptual hypothesis, some researchers (Apperly & Butterfill 2009, Butterfill & Apperly 2013, Heyes 2014, Thoermer et al. 2012) have argued that infants “pass” these implicit tasks using behavioral heuristics, and that these tasks do not actually tap into the adult-like concept of belief. This is related to the claim that the concepts of belief and desire differ in complexity (see Schwitzgebel 1999, Flavell 1988, Fergusson & Gopnik 1988, and others).

Our findings are also consistent with a view that attributes the asymmetry between think and want to linguistic factors. De Villiers (2005), for instance, argues for a SYNTACTIC ASYMMETRY in the acquisition of the syntax of tensed vs. untensed complements. This can explain why we see adult-like performance with want earlier than has been observed with think. Potentially problematic for this type of explanation, however, are languages like German. In German, want can also take a tensed sentential complement, and is still acquired before think, which makes it doubtful that this specific piece of syntactic knowledge can entirely explain the asymmetry (cf. Perner et al 2003). A more recent proposal from de Villiers has assumed that the important feature of false complements must be semantic, having to do with how the different perspectives on a proposition are represented (de Villiers & de Villiers, 2009).

A third possible source for the think/want asymmetry may lie in the kinds of pragmatic enrichments that these verbs trigger (Hacquard 2014, Lewis, Hacquard & Lidz 2012, Harrigan 2015). Think sentences, unlike want sentences, can be used to make indirect assertions. Lewis and colleagues (Lewis et al 2012, Lewis 2013) provide evidence that children may assume that think is being used in this way, even in contexts where adults do not. If this is right, then such pragmatic uses might be responsible for children’s tendency to reject think sentences when they report a false belief. Given that want sentences report preferences and, as such, are not routinely
used to make indirect assertions; they thus do not lead children down the same pragmatic garden path as *think* sentences.

No matter what the right explanation for the *think/want* asymmetry, the two studies presented in this paper can be taken as evidence that under the strictest conditions, 3-year-olds show adultlike comprehension of sentences containing *want*, and hence that their DESIRE concept is in place. Because we find children succeeding with want even when it is used to report desires that conflict with reality or with other desires, our results cast some doubt on analyses that take such conflicts as a major part of the explanation of children's difficulty with *think*.

References


Rakoczy, H., Warneken, F., & Tomasello. M. (2007). ‘This way!’, ‘No! That way!’—3 year olds know that two people can have mutually incompatible desires. Cognitive Development, 22, 47-68.


Appendix

Experiment 1 Materials

Practice Items:

Practice 1:

Story: *Stacey is at home, making a sandwich for her mom.*

Test Sentence: *Stacey is making a sandwich.*

Practice 2:

Story: *It is so hot outside, Alan is swimming at the pool.*

Test sentence: *Alan is playing on the playground.*

Story 1: Bedtime

Conflict/Switch:

Amy is in her room playing with toys. Amy’s mom comes in and says “Amy, come downstairs, we have company!” Amy says “OK, mom!” Amy’s mom leaves, and Amy says to herself, “I know my
mom asked me to go downstairs, but I’m having too much fun playing, I’m going to stay here!”
And she keeps playing with toys in her room.

No Conflict/Stay

Amy is in her room playing with toys. Amy’s mom comes in and says “Amy, come downstairs, we have company!” Amy says “OK, mom!” Amy’s mom leaves, and Amy says to herself, “I’d like to keep playing with my toys, but my mom asked me to come downstairs, so I’d better do it!” And she starts to go downstairs.

False filler: Amy’s dad came up to her room to talk to her.
Test sentence: Amy’s mom wants her to be going downstairs right now.

Story 2: Picking up from School

Conflict/Switch

Jimmy is at school reading with his friend. Dad comes to pick Jimmy up from school. Dad says, “It’s time to go home! I’m going to go get your brother from his classroom, go outside and play and wait for me.” And Jimmy says “OK, I will.” Dad leaves, and Jimmy says to himself “I know my dad said to go outside and play and wait, but I’m having too much fun reading!” And he stays with his friend.

No Conflict/Stay

Jimmy is at school reading with his friend. Dad comes to pick Jimmy up from school. Dad says, “It’s time to go home! I’m going to go get your brother from his classroom, go outside and play and wait for me.” And Jimmy says “OK, I will.” Dad leaves, and Jimmy says to himself “I’m
having so much fun with my friend, I’d like to keep reading, but my dad said I have to go outside and wait, so I’d better do that” And Jimmy goes outside to play and wait.

True filler: Jimmy’s dad came to pick him up from school.
False filler: Jimmy’s mom came to pick him up from school.
Test sentence: Jimmy’s dad wants him to be outside playing right now.

Story 3: At the doctor’s office

Conflict/Switch

Jeffrey is in the waiting room at the doctor’s office. While he waits for the nurse to come in, he plays with a stethoscope that he sees in the waiting room. The nurse comes in. She says, “Hi, Jeffrey, I’m going to come and take your temperature, but I have to go get my thermometer. Please put down that stethoscope and play with these blocks until I come back” and Jeffrey says “OK!” While the nurse is gone, Jeffrey says to himself, “I know the nurse said I should play with the blocks instead, but I’m having too much fun playing with the stethoscope!” And he keeps playing with the stethoscope.

No Conflict/Stay

Jeffrey is in the waiting room at the doctor’s office. While he waits for the nurse to come in, he plays with a stethoscope that he sees in the waiting room. The nurse comes in. She says, “Hi, Jeffrey, I’m going to come and take your temperature, but I have to go get my thermometer. Please put down that stethoscope and play with these blocks until I come back” and Jeffrey says “OK!” While the nurse is gone, Jeffrey says to himself, “I’d like to keep playing with this stethoscope, but the nurse said I should play with these blocks instead, so I’d better do that.” And he starts to play with the blocks.
True filler: *The nurse came into the waiting room to talk to Jeffrey.*
False filler: *The doctor came into the waiting room to talk to Jeffrey.*
Test sentence: *The nurse wants Jeffrey to be playing with the stethoscope right now.*

**Story 4: In the kitchen**

**Conflict/Switch**

*Maggie is helping her sister bake cookies in the kitchen. Maggie is stirring a pot on the stove.*

*Maggie’s sister realizes she has to run to the store to get chocolate chips because they are out.*

*She says, “Maggie, that’s enough stirring, could you go upstairs and play while I run to the store?” and Maggie says “OK!” When her sister leaves, Maggie says to herself “I know my sister told me to go upstairs and play, but I really like stirring this pot!” And she keeps stirring the pot anyway.*

**No Conflict/Stay**

*Maggie is helping her sister bake cookies in the kitchen. Maggie is stirring a pot on the stove.*

*Maggie’s sister realizes she has to run to the store to get chocolate chips because they are out.*

*She says, “Maggie, that’s enough stirring, could you go upstairs and play while I run to the store?” and Maggie says “OK!” When her sister leaves, Maggie says to herself “I really like stirring this pot, but my sister said I should go upstairs and play, so I’m going to!” And she goes upstairs to play.*

True filler: *Maggie is baking cookies with her sister.*
False filler: *Maggie is baking cookies with her brother.*
Test sentence: *Maggie’s sister wants her to be stirring the pot right now.*
Story 5: At School

**Conflict/Stay**

Alex is at school, playing dress-up. His teacher calls out to the class, “OK, everyone! I have to run next door, stay at the play stations you’re in until I get back!” Alex says “OK!” Alex says to himself, “I know my teacher said I have to stay in the dress-up corner, but I’d like to go color, so I will!” And he goes over to color.

**No Conflict/Switch**

Alex is at school, playing dress-up. His teacher calls out to the class, “OK, everyone! I have to run next door, stay at the play stations you’re in until I get back!” Alex says “OK!” Alex says to himself, “I’d like to go color, but my teacher said I have to stay here in the dress-up corner, so I will!” And he stays in the dress-up corner.

True filler: Alex is playing at school.
False filler: Alex is playing at home.
Test sentence: The teacher wants Alex to be playing dress-up right now.

Story 6: At the office

**Conflict/Stay**

Jeremy’s teacher sent him to the office to pick something up for her from the principal. Jeremy is waiting for the principal and looking at a book. The secretary comes over to Jeremy. She says, “the principal is talking to someone right now, sit here and keep looking at that book until he’s ready for you.” And Jeremy says “OK!” When the secretary turns away, Jeremy says to himself
“I know the secretary said that I should keep looking at this book, but I’d like to see what’s happening in the principal’s office.” And he goes over and peeks into the principal’s office.

No Conflict/Switch

Jeremy’s teacher sent him to the office to pick something up for her from the principal. Jeremy is waiting for the principal and looking at a book. The secretary comes over to Jeremy. She says, “the principal is talking to someone right now, sit here and keep looking at that book until he’s ready for you.” And Jeremy says “OK!” When the secretary turns away, Jeremy says to himself “I’d really like to peek in the principal’s office right now, but the secretary said I should keep looking at this book, so I will.” And he keeps looking at the book.

True filler: Jeremy is waiting in the principal’s office.
False filler: Jeremy is waiting in the doctor’s office.
Test sentence: The secretary wants Jeremy to be looking at a book right now.

Story 7: Grocery shopping

Conflict/Stay

Megan is at the grocery store with her mom, and she is sitting in the cart while her mom shops. Megan’s mom says, “Megan, I have to run and get something in the next aisle, stay right there in the cart until I get back,” and Megan says “No problem, mom!” Megan says to herself, “I know my mom said to stay in the cart, but I’d like to get out and go get some cereal, so I will.” And she climbs out of the cart to go get some cereal.

No Conflict/Switch
Megan is at the grocery store with her mom, and she is sitting in the cart while her mom shops. Megan’s mom says, “Megan, I have to run and get something in the next aisle, stay right there in the cart until I get back,” and Megan says “No problem, mom!” Megan says to herself, “I’d like to get out of the cart and go get some cereal, but my mom said to stay in the cart, so I will.” And she stays in the cart.

True filler: Megan is at the grocery store with her mom.
False filler: Megan is at the grocery store with her dad.
Test sentence: Megan’s mom wants her to be getting cereal right now.

**Story 8: After dinner**

**Conflict/Stay**

After dinner, Doug is playing a computer game while his parents clean up from dinner. His dad comes in and says, “Doug, it’s going to be time for bed soon, just play your game for awhile while we clean up from dinner.” and Doug says “OK, dad!” Doug says to himself, “I know my dad said I should keep playing my game, but I would really like to go outside and play, so I will.” And he goes outside to play.

**No Conflict/Switch**

After dinner, Doug is playing a computer game while his parents clean up from dinner. His dad comes in and says, “Doug, it’s going to be time for bed soon, just play your game for awhile while we clean up from dinner.” and Doug says “OK, dad!” Doug says to himself, “I’d really like to go outside and play right now, but my dad said I should keep playing my computer game because it’s almost time for bed.” And he keeps playing his computer game.
True filler: Doug just finished eating dinner.

False filler: Doug just finished eating breakfast.

Test sentence: Doug’s dad wants him to be outside playing right now.