Research Report

Thinking While Talking

Adults Fail Nonverbal False-Belief Reasoning

Ashley M. Newton and Jill G. de Villiers

Smith College

ABSTRACT—*This experiment tested the ability of 81 adult subjects to make a decision on a simple nonverbal false-belief reasoning task while concurrently either shadowing prerecorded spoken dialogue or tapping along with a rhythmic shadowing track. Our results showed that the verbal task, but not tapping, significantly disrupted false-belief reasoning, suggesting that language plays a key role in working theory of mind in adults, even when the false-belief reasoning is nonverbal.*

How is language involved in thought? According to a weak account (Bloom & Keil, 2001), language plays the role of conveying the content of areas of human reasoning that are inaccessible to behavioral observation: for example, history, abstract scientific theories, cultural understandings, and religion. A shared language is indispensable for conveying such knowledge. According to a strong account, language plays a crucial role in thinking itself. There are four clear contemporary versions of this latter account. First, language could be involved in all conceptual thinking. Evidence of the conceptual abilities of nonverbal infants and chimpanzees is inconsistent with this position, however (Carruthers, 2002). Second, language may be a tool employed whenever there are long inference chains (Clark, 1998). Third, language could serve central cognition as a cross-modular bridge whenever the output from two separate cognitive modules must be combined (Carruthers, 2002; Spelke, 2003). Fourth, linguistic representation may be uniquely able to represent truth and falsity (Fodor, 1975; Jackendoff, 1996). In particular, it has been claimed that language is the medium most suited for representing concepts of other individuals' false beliefs (J.G. de Villiers & de Villiers, 2000; Segal, 1998).

Being able to represent other individuals' false beliefs has long been regarded as the acid test of a theory of mind. Many developmental researchers argue that language is an essential precursor of false-belief representation (Astington & Baird,

2005). The weak position is that language is necessary as a means of conveying the cultural theory about minds (Harris, 2005; Nelson, 2005). Researchers testing deaf children who are not exposed to early sign language and have severe language delays concur that language is a crucial aid to theory-of-mind development (P.A. de Villiers, 2005; Peterson & Siegal, 1999; Woolfe, Want, & Siegal, 2002). However, the findings about deafness show only that language is needed for normal acculturation. A strong version of linguistic determinism contends that certain critical forms in language provide the representational tools for later theory of mind, especially the understanding of other individuals' false beliefs (J.G. de Villiers & de Villiers, 2000; Tager-Flusberg, 1997). In this view, a child without language is not just deprived of access to the contents of the theory, but lacks the means to hold in mind for further processing the content of other individuals' beliefs when that content differs from reality. The researchers who hold this view tie *specific* linguistic achievements, such as sentence complementation, to success on false-belief tasks (J.G. de Villiers, 2005).

If language is needed just for acculturation, then it is quite possible that adult reasoning is independent of language. Case studies of adult aphasia call into question the necessity of language for false-belief reasoning (Varley & Siegal, 2000; Varley, Siegal, & Want, 2001).

Dual-task studies have been used successfully to explore the contribution of different skills to cognitive operations in intact adults. Hermer-Vasquez, Spelke, and Katsnelson (1999) used a dual-task study to investigate whether adults' ability to locate objects could be disrupted by having them simultaneously engage in a linguistic task. By hypothesis, the location task involved integrating information from two separate cognitive domains, namely, color and geometry, into a structure such as "to the left of the blue wall." Spatially disoriented adult subjects had to locate an object while simultaneously shadowing, or immediately echoing, unrelated verbal material. A group of control subjects instead shadowed a rhythmic tapping. Verbal shadowing disrupted the adults' ability to locate the object, whereas rhythmic shadowing did not. Hermer-Vasquez et al. proved via an unrelated visual search task that verbal shadowing

Address correspondence to Jill de Villiers, Psychology Department, Smith College, Northampton, MA 01063, e-mail: jdevil@smith.edu.

and rhythmic shadowing posed equal difficulty on that simultaneous search task. Because the location task entailed the combination of information from two different modules, it required language mediation. Verbal but not rhythmic shadowing differentially disrupted that mediation.

Carruthers (2002) suggested that the dual-task design is useful to investigate whether or not tying up language resources disrupts how adults reason about others' false beliefs. In the present study, we tested the following hypotheses:

- Hypothesis 1: Adult subjects will find it difficult to respond in a reasoning task that requires understanding another individual's false beliefs if their language resources are tied up in verbal shadowing.
- Hypothesis 2: Adult subjects will not find it difficult to respond in an equivalent task that entails only true beliefs, not false beliefs.
- Hypothesis 3: Adult subjects will find it easy to take into account another individual's false belief if their attention is taken up by a nonverbal rhythmic tapping task that makes attentional demands equal to those of the verbal shadowing task.

METHOD

Subjects

Subjects were 81 female adults (ages 18–35) who possessed native English fluency.

Tasks and Procedure

Participants viewed two videos, one of two false-belief skits (rabbit or mouse) and one of two true-belief skits (rabbit or mouse). Participants were divided into two groups defined by which interference task was paired with which video category. Group A received the verbal shadowing task with a false-belief video and the rhythmic shadowing task with a true-belief video; Group B received the reverse pairing. There were two task orders within each group.

Subjects participated in two trials, one with a false-belief video and one with a true-belief video. In both skits (with costumed actors), a character placed a food item under one of three buckets (A, B, or C), which were distributed evenly on the stage. In the false-belief video, after the character left, a cat moved the item to a different bucket (see Table 1 and Fig. 1). In the true-belief video, the character watched a cat move the food to another bucket. Half the subjects saw videos in which the mouse had the true belief and the rabbit had the false belief, and half saw videos in which the mouse had the false belief and the rabbit had the true belief.

For both videos, the decision point was the same: The screen flashed, and two alternate endings played simultaneously on the same screen. One ending had a solid red background, and the other had a solid green background (counterbalanced across subjects). One ending showed the main character returning to the original location of the food item; the other showed the character approaching the new location of the food item. The subject responded using a wooden block, tapping a red section on the block if the scene with the red background showed the correct ending and a green section on the block if the scene with the green background showed the correct ending (Fig. 2). The key difference between the two movie trials is that one character did not see the food moved (false belief), whereas the other one did (true belief).

Before each trial, subjects practiced the interference task until competent. Next, they watched a practice video while performing the interference task. The content of this video was a causal event: A glass of milk was knocked over. The two ending choices showed the milk spilling or the milk remaining solid inside the glass (fake milk). This practice allowed subjects to become comfortable performing the interference task while watching a video and then responding.

One interference task entailed constant verbal shadowing of English sentences; the other required the subject to repeat short rhythmic patterns by tapping them out with a stick on a piece of wood. The subject listened to the stimuli for both tasks through headphones. Although the verbal shadowing was performed constantly, the rhythmic tapping was broken up into short "call

TABLE 1
Examples of the Sequence of Events in the False-Belief and True-Belief Videos (see Fig. 1)

| Sample false-belief video	Sample true-belief video
Mouse nibbles cheese	Rabbit nibbles carrots
Mouse puts cheese under C	Rabbit puts carrots under A
Mouse yawns and stretches	Rabbit yawns and stretches
Mouse goes behind curtain in central door	Rabbit stands and watches cat go to A
Cat goes to C	Cat moves carrots to B and then hides them at C
Cat moves cheese to B and then hides it at A	Cat leaves
Cat leaves	Rabbit goes behind curtain in central door
Scene is bare, light flashes	Scene is bare, light flashes
Mouse emerges at curtain	Rabbit emerges at curtain
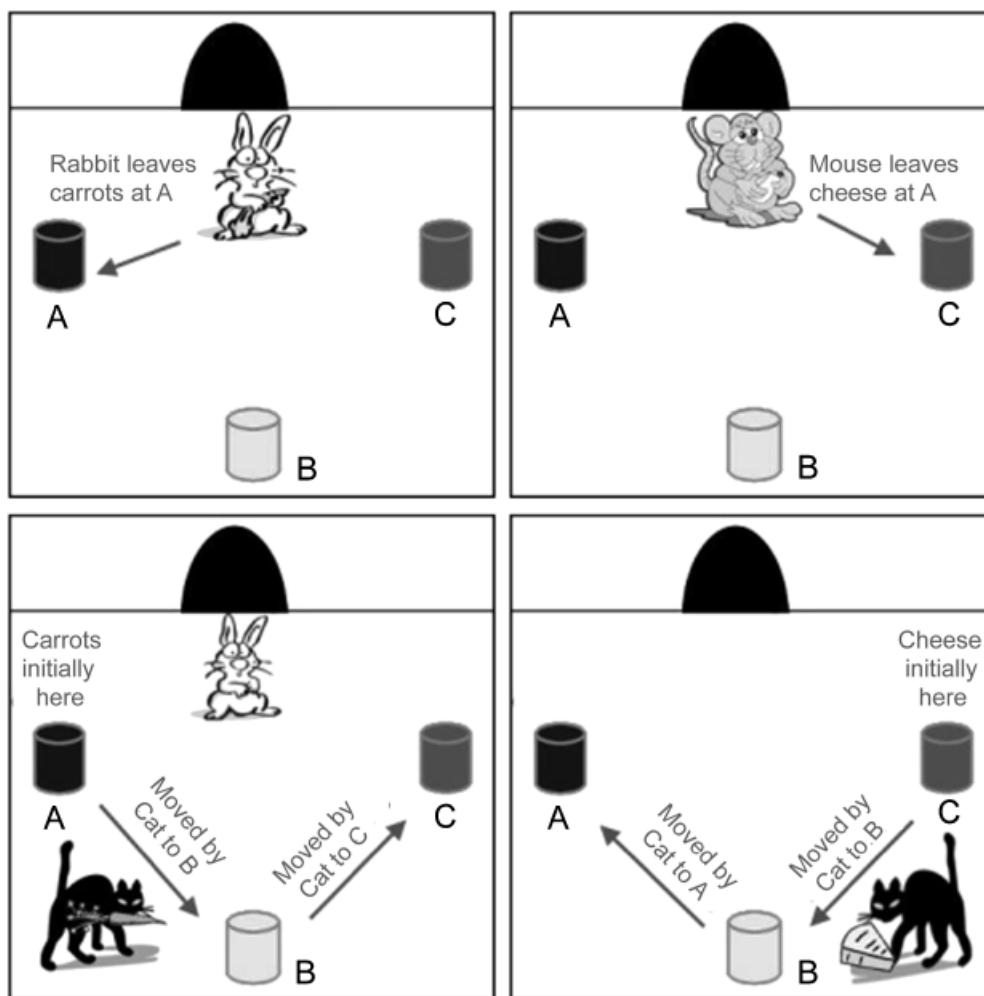
Fig. 1. Cartoon illustration of the object's placement and movement in the true-belief (left) and false-belief (right) videos. Half the subjects viewed the events illustrated here, and the other half viewed a pair of videos in which the mouse had the true belief and the rabbit had the false belief. In the actual videos, real actors portrayed the events. The videos were carefully matched for length.

and response" segments: The subject listened to a 4/4 measure of beats (averaging around 5–6 notes) and then tapped the rhythm out during a 4/4 measure of silence, after which a new rhythmic measure played, and the subject had to listen attentively.[1] All the subjects were video-recorded for later analysis.

To ensure the equivalence of the two interference tasks, we conducted a pilot test. Various interference tasks were imposed on a different type of reasoning task involving visual search of an array of faces. Each subject watched a computer monitor showing a series of randomized images that contained various stylized faces in varying layouts and orientations. The subject had to decide quickly whether the faces in each image were homogeneously or heterogeneously oriented on the screen. The subject's reaction time and answer were recorded for each image shown. This task, though taxing, does not involve any kind of belief reasoning. During this task, the subject was simultaneously engaged in five interference tasks, one at a time: an English verbal shadowing task, a Swahili verbal shadowing task, the same rhythmic tapping task used in the main study, a rhythmic babbling task (similar to rhythmic tapping), and a humming task (shadowing music similar to continuous verbal speech). There was no statistically significant difference in reaction time or error rate on the visual reasoning task between trials on which the rhythmic tapping task was performed and trials on which the English verbal shadowing task was performed. Therefore, these were the interference tasks selected for the main study.

RESULTS

Data from 15 subjects were discarded because of their failure to maintain the shadowing task throughout either video. Careful analysis of the recorded video of each subject's two trials was

---

[1]We tried first to have subjects shadow the rhythm continuously, as in verbal shadowing, but our subjects proved unable to learn the task.
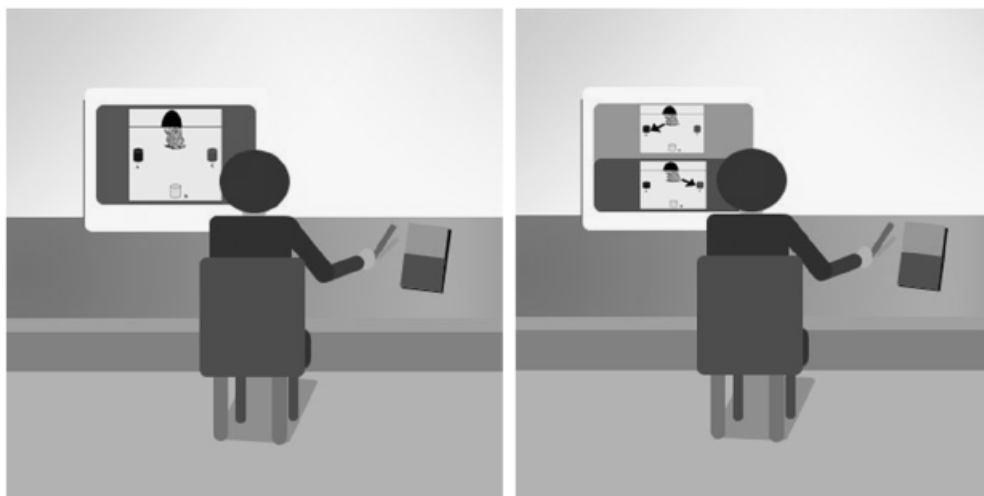
Fig. 2. Illustration of a subject watching the test video (left) and then responding to the alternative endings (right).

used to determine whether the subject paused for more than 2 s in the shadowing task or failed to maintain shadowing at the decision point of the reasoning task. Each decision was checked by a second coder.

The subjects' success or failure on the reasoning task was calculated for each interference task and belief condition (see Table 2). A Fisher's exact test examining the interaction between interference task and belief condition was highly significant ($p < .0006$). Subjects fared significantly worse in the false-belief condition when they simultaneously performed the verbal task (13 of 31 succeeded) than when they performed the nonverbal, rhythmic tapping task (29 of 35 succeeded). However, subjects appeared to fare about the same in the true-belief condition regardless of which interference task they performed (verbal: 32 of 35 succeeded, rhythmic: 31 of 31 succeeded; $p < .14$). The proportion of subjects who succeeded in the false-belief condition during verbal shadowing (.4) was not significantly different from chance performance (.5). However, the proportion of subjects who succeeded during the rhythmic tapping task was high in each belief condition ($M = .83$ and 1.0 for the false-belief and true-belief conditions, respectively). These data indicate that a subject who is unable to recruit the use of relevant language in the performance of a false-belief task has difficulty successfully recognizing and reasoning about a false belief.

TABLE 2

Number of Subjects Who Passed and Failed the False-Belief and True-Belief Tasks

Performance	False-belief video		True-belief video	
	Verbal shadowing	Rhythmic shadowing	Verbal shadowing	Rhythmic shadowing
Pass	13	29	32	31
Fail	18	6	3	0
Total	31	35	35	31

We also ran a mixed-design analysis of variance with group (A or B), order of interference task, and false-belief character as between-subjects variables and interference task (verbal vs. rhythmic) as the repeated measure. The results showed no significant effect of the order of interference task or which character held the false belief. The type of interference task was highly significant, $F(1, 58) = 15.68$, $p < .001$, $\eta^2 = .213$, and group was also highly significant, $F(1, 58) = 7.30$, $p < .01$, $\eta^2 = .112$. However, the Task × Group interaction explains these effects, $F(1, 58) = 29.44$, $p < .001$, $\eta^2 = .337$, $p_{rep} = .99$ (see Fig. 3). Those subjects who were engaged in verbal shadowing could not choose the correct outcome in the false-belief video.

DISCUSSION

The results reveal that verbal shadowing significantly disrupts a subject's ability to monitor a character's belief in a nonverbal false-belief task. The subjects who shadowed rhythm during the same false-belief task could easily monitor the character's belief and respond properly to predict the event's outcome. Furthermore, during both shadowing tasks, when no false belief was involved, subjects successfully tracked the desired object across three locations, predicting where a character would go to retrieve it. Therefore, neither shadowing task was so difficult that subjects could not simultaneously focus on, and decide about, a complex physical event. In summary, only verbal shadowing disrupted only false-belief reasoning.

However, tying up the language module could mean that (a) executive skills relying on verbal inference were disrupted, (b) language could not serve as the bridge between two different modules, or (c) no medium was available to represent false belief. The data do suggest that language is necessary in complex theory-of-mind tasks, rather than merely serving as the means by which children learn about the culture's theory of mind. The language faculty continues to be called upon in adults' false-belief reasoning.
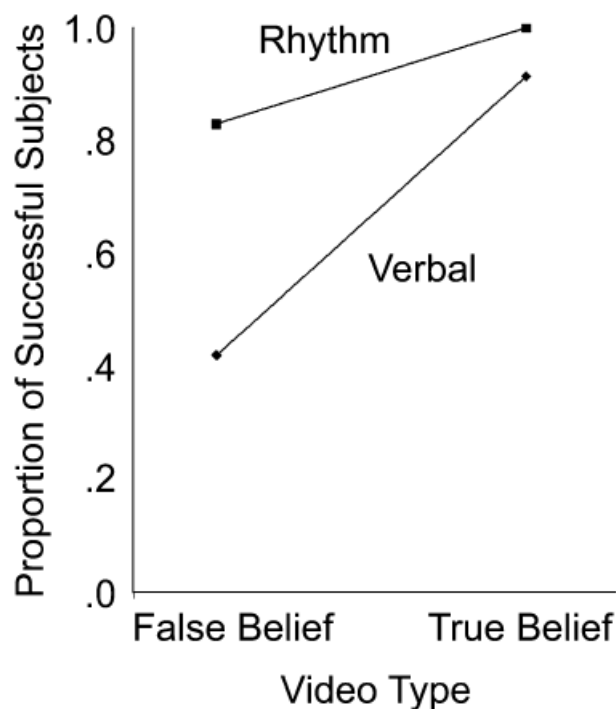
Fig. 3. Proportion of subjects who responded correctly as a function of video (false-belief vs. true-belief) and interference task (verbal vs. rhythmic shadowing).

The results are counter to what might be predicted given the results of Varley and Siegal (2000) and Varley et al. (2001), whose 2 aphasic subjects could still reason about false-belief tasks though they apparently could not produce or comprehend syntax. Possibly these global aphasic patients still have intact logical form (Carruthers, 2002) despite their performance deficits. Alternatively, perhaps adults' theory-of-mind reasoning requires only enough language resources for the executive functioning that supports this kind of reasoning, and verbal shadowing wipes those resources out.

Why is language needed specifically to support executive function in the case of a false-belief event? Most theories posit that the phonological loop is used in executive function (e.g., maintained rehearsal of a command such as "go left" or "not the red"). Suppose that reasoning about beliefs takes place without the language faculty, but that the outcome of this process is a decision to act that must be maintained in the phonological loop. Such an account fails to explain the delays in false-belief reasoning found in children who have language deficits but show good executive function on other tasks (P.A. de Villiers, 2005), but ruling out this account is more difficult in the case of adults. Could one tie up the phonological loop without engaging the language module itself? In our dual-task pilot test, subjects' performance on the reasoning task showed no disruption either in accuracy or in speed when the interference task was shadowing rhythmic babbling (e.g., *bababa*). We plan to test performance on the present false-belief task when the interference task involves more variegated babbling. However, an alternative reasoning task in which adults do not have to make a decision that depends on phonological rehearsal may hold more promise for unpacking the alternative possible roles for language.

Consider Onishi and Baillargeon's (2005) claim that 16-month-olds demonstrate sensitivity to the false beliefs of another character when looking time is used as an index of surprise. Perner and Ruffman (2005) offered some interesting alternative explanations for how gaze time might be affected by the stimuli without invoking false-belief understanding. However, Perner and Ruffman argued that implicit false-belief reasoning as indexed by preferential looking may be manifest in children as young as 2 years 11 months (Clements & Perner, 1994).

If it is only the decision-making stage that calls upon the phonological loop in adults, then perhaps they could show by some other index that they had processed the appropriate information about a character's belief during verbal shadowing, but were blocked at the decision point by the unavailability of their language faculty. We are currently exploring the possibility that adults have access to implicit understanding of false belief, indexed by eye gaze, during verbal shadowing.

REFERENCES

Astington, J.W., & Baird, J.A. (Eds.). (2005). *Why language matters for theory of mind*. New York: Oxford University Press.

Bloom, P., & Keil, F.C. (2001). Thinking through language. *Mind and Language*, 16, 351–367.

Carruthers, P. (2002). The cognitive functions of language. *Behavioral and Brain Sciences*, 25, 657–674.

Clark, A. (1998). Magic words: How language augments human computation. In P. Carruthers & J. Boucher (Eds.), *Language and thought* (pp. 162–183). New York: Cambridge University Press.

Clements, W.A., & Perner, J. (1994). Implicit understanding of belief. *Cognitive Development*, 9, 377–395.

de Villiers, J.G. (2005). Can language acquisition give children a point of view? In J.W. Astington & J.A. Baird (Eds.), *Why language matters for theory of mind* (pp. 186–219). New York: Oxford University Press.

de Villiers, J.G., & de Villiers, P.A. (2000). Linguistic determinism and the understanding of false beliefs. In P. Mitchell & K. Riggs (Eds.), *Children's reasoning and the mind* (pp. 189–226). Hove, England: Psychology Press.

de Villiers, P.A. (2005). The role of language in theory-of-mind development: What deaf children tell us. In J.W. Astington & J.A. Baird (Eds.), *Why language matters for theory of mind* (pp. 266–297). New York: Oxford University Press.

Fodor, J.A. (1975). *The language of thought*. New York: Thomas Y. Crowell.

Harris, P. (2005). Conversation, pretence and theory of mind. In J.W. Astington & J.A. Baird (Eds.), *Why language matters for theory of mind* (pp. 70–83). New York: Oxford University Press.

Hermer-Vasquez, L., Spelke, E.S., & Katsnelson, A.S. (1999). Sources of flexibility in human cognition: Dual-task studies of space and language. *Cognitive Psychology*, 39, 3–36.

Jackendoff, R. (1996). How language helps us think. *Pragmatics and Cognition*, 4, 1–34.

Nelson, K. (2005). Language pathways into the community of minds. In J.W. Astington & J.A. Baird (Eds.), *Why language matters for theory of mind* (pp. 26–49). New York: Oxford University Press.

Onishi, K.H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, *308*, 255–258.

Perner, J., & Ruffman, T. (2005). Infants' insight into the mind: How deep? *Science*, *308*, 214–216.

Peterson, C.C., & Siegal, M. (1999). Representing inner worlds: Theory of mind in autistic, deaf, and normal hearing children. *Psychological Science*, *10*, 126–129.

Segal, G. (1998). Representing representations. In P. Carruthers & J. Boucher (Eds.), *Language and thought* (pp. 146–161). New York: Cambridge University Press.

Spelke, E. (2003). What makes humans smart? In D. Gentner & S. Goldin-Meadow (Eds.), *Advances in the investigation of language and thought* (pp. 277–311). Cambridge, MA: MIT Press.

Tager-Flusberg, H. (1997). The role of theory of mind in language acquisition: Contributions from the study of autism. In L. Adamson & M.A. Romski (Eds.), *Research on communication and language disorders: Contributions to theories of language development* (pp. 133–158). Baltimore: Paul Brookes.

Varley, R., & Siegal, M. (2000). Evidence for cognition without grammar from causal reasoning and theory of mind in an agrammatic aphasic patient. *Current Biology*, *10*, 723–726.

Varley, R., Siegal, M., & Want, S.C. (2001). Severe impairment in grammar does not preclude theory of mind. *Neurocase*, *7*, 489–493.

Woolfe, T., Want, S.C., & Siegal, M. (2002). Signposts to development: Theory of mind in deaf children. *Child Development*, *73*, 768–778.