

The role of executive function in perspective taking during online language comprehension

SARAH BROWN-SCHMIDT

University of Illinois at Urbana-Champaign, Urbana, Illinois

During conversation, interlocutors build on the set of shared beliefs known as *common ground*. Although there is general agreement that interlocutors maintain representations of common ground, there is no consensus regarding whether common-ground representations constrain initial language interpretation processes. Here, I propose that executive functioning—specifically, failures in inhibition control—can account for some occasional insensitivities to common-ground information. The present article presents the results of an experiment that demonstrates that individual differences in inhibition control determine the degree to which addressees successfully inhibit perspective-inappropriate interpretations of temporary referential ambiguities in their partner's speech. Whether mentioned information was grounded or not also played a role, suggesting that addressees may show sensitivity to common ground only when it is established collaboratively. The results suggest that, in conversation, perspective information routinely guides online language processing and that occasional insensitivities to perspective can be attributed partly to difficulties in inhibiting perspective-inappropriate interpretations.

Interlocutors bring separate perspectives to a conversation. For any pair of individuals, some aspects of their perspectives are shared, known as *common ground* (Stalnaker, 1978), whereas other aspects are private, known as *privileged ground*. As a conversation progresses, common ground grows when interlocutors introduce new information into the discourse. Common ground is thought to include shared community and cultural experiences, the physically copresent environment, and shared linguistic exchanges (Clark & Marshall, 1981).

Although there is general agreement that interlocutors maintain representations of common ground, the results of studies examining whether perspective guides online interpretation processes have been equivocal, with some results pointing to early use of perspective in online processing (Brown-Schmidt, Gunlogson, & Tanenhaus, 2008; Hanna & Tanenhaus, 2004; Hanna, Tanenhaus, & Trueswell, 2003; Heller, Grodner, & Tanenhaus, 2008; Nadig & Sedivy, 2002) and other results pointing to unreliable or late use of perspective (Keysar, Barr, Balin, & Brauner, 2000; Keysar, Lin, & Barr, 2003). Thus, the challenge for models of the role of perspective in language processing is to account for why perspective sometimes constrains online processing and sometimes does not.

According to perspective-adjustment models, listeners generate perspective-inappropriate interpretations during online interpretation either due to strategic egocentric processing strategies (Keysar et al., 2000) or because perspective-inappropriate interpretations are autonomously activated (Barr, 2008). On the strategic account,

online perspective taking occurs only in special circumstances (i.e., when global ambiguities require perspective use). On the autonomous account, a combination of anticipatory perspective taking (e.g., anticipating reference to a shared object) and a failure to use perspective information as words are interpreted online yields results that misleadingly appear to show online use of perspective (Barr, 2008).

Alternatively, constraint-based models maintain that perspective is one of multiple probabilistic, partial constraints that guide online interpretation; perspective-inappropriate interpretations are likely to occur when evidence supports the perspective-inappropriate interpretation, or when perspective information is unreliable or uncertain (Heller et al., 2008).

Each of these accounts provides partial explanations for the competing results in the literature. However, little to no work experimentally manipulates or predicts when online perspective-taking failures are and are not likely to occur. In the present article, I explore the possibility that subtle individual differences in inhibitory control can account, in part, for adults' occasional failure to use perspective information to inhibit perspective-inappropriate interpretations. In principle, such a relationship could be theory neutral; thus, the goal of this endeavor is to inform current models by exploring the factors that contribute to perspective errors.

To consider another person's perspective when it conflicts with your own requires inhibiting the response prompted by your own perspective (Friedman & Leslie,

S. Brown-Schmidt, brownsch@illinois.edu

2005). Thus, individuals with low inhibition control may be unable to inhibit egocentric interpretations. Adults with low comprehension skills are less likely to suppress inappropriate meanings of homonyms (Gernsbacher, 1997), suggesting that language processing abilities correlate with the ability to suppress alternative interpretations. Furthermore, 2- to 5-year-olds with low inhibition control are more likely to consider egocentric interpretations of a partner's requests (Nilsen & Graham, 2009). However, it is not clear whether inhibition control would guide adults' online perspective taking, and the results with children may reflect their still-developing conceptual and executive functioning (Ramscar & Gitcho, 2007; Wellman, Cross, & Watson, 2001) and may not be representative of adults' abilities.

The experiment presented here used an interactive dialogue task in which participants' eye movements were monitored as they interpreted questions that contained *temporary* ambiguities that could be resolved early if they took perspective into account. Because questions typically ask about information known only to the addressee, perspective use during question interpretation requires interpreting the question as asking about privileged- rather than common-ground information. If the degree to which participants use perspective is predicted by their inhibition control, this would offer a partial explanation for the equivocal findings: When inhibition control is lacking, addressees are unable to inhibit perspective-inappropriate interpretations, despite having access to perspective information.

This experiment also manipulated whether critical perspective information was grounded in common ground. On one account, perspective-taking failures can occur when interlocutors have not interactively established which referents are in common ground through the grounding process (see Clark & Brennan, 1991; Hanna et al., 2003). Grounding can take the form of, among other things, verbal acknowledgment (e.g., "Got it, OK") or maintained attention. A lack of grounding can be indicated by distraction or denial (e.g., "Sorry, I didn't hear you."). Thus, whether addressees use common-ground representations may depend also on whether the critical information in question was in fact grounded and accepted into common ground.

METHOD

Participants

Forty-eight native English-speaking participants from the University of Illinois at Urbana-Champaign participated. Three additional participants were excluded from analysis due to equipment malfunction ($n = 1$), not being a native English speaker ($n = 1$), and fatigue ($n = 1$). Participants were compensated \$8 or given partial course credit.

Materials and Procedure

The participant and an experimenter sat at separate computers. The participant's gaze was monitored using a desktop-mounted Eye-Link 1000 eyetracking system, which recorded gaze monocularly at 1000 Hz. The experimenter wore a headset microphone. Audio and eyetracking data were recorded to disk. Following the main experi-

ment, participants completed two inhibitory-control tasks, which were used to predict performance in the main task.

Across 72 trials, the participant and experimenter examined a series of displays, each of which showed a 5×3 grid of animals (Figures 1A–1B). Each square on the grid contained an animal (pig, cow, horse, bear, caterpillar, or turtle). Each animal wore an accessory (glasses, shoes, or lipstick). On each trial, one third of the animals were seen by both partners and thus were in visual common ground; these animals were distinguished by a white background. One third were seen only by the participant and were in the participant's privileged ground. The remaining one third were seen only by the experimenter and were in the experimenter's privileged ground. Visually privileged animals had a gray background. Animals in the other person's privileged ground were presented as black squares.

Participants were instructed that they would play an interactive game with an experimenter who worked in the lab. On each trial, the game was to check whether the animals were arranged in a pattern in accordance with the following rules: Animals in adjacent squares had to be different types of animals and had to be wearing different accessories. If adjacent animals were of the same type or had the same accessory, this was considered an error. Participants were told that the task required sharing information because neither participant nor experimenter saw all the animals. So, in order to identify errors, the experimenter would ask questions about the participant's images. Nine filler trials contained errors. Critical trials never contained errors. The task lasted approximately 35 min.

On each trial, the experimenter used a scripted set of three utterances that formed the conditions of interest (participants were unaware of the scripting). The first utterance was a setup question that either brought the competitor or the noncompetitor into linguistic common ground. The second utterance asked about an unrelated animal. A point-of-disambiguation manipulation (Eberhard, Spivey-Knowlton, Sedivy, & Tanenhaus, 1995) was used to examine the temporary, online commitments that addressees made as they interpreted the final utterance, which was the critical question. For example, given the scenes in Figures 1A and 1B, if the experimenter asks "What's above the cow that's wearing shoes?" at the onset of the word *cow*, the question is temporarily ambiguous between asking about the animal above the cow with glasses and the animal above the cow with shoes. This ambiguity is resolved linguistically at the point of disambiguation: the word *shoes*.¹ If the animal above one of the cows were in common ground, addressees should be unlikely to interpret the question as asking about that animal and should direct attention away from it before the point of disambiguation. Critically, in this study, taking perspective into account involves interpreting the critical question as asking about privileged- rather than common-ground information. The remainder of the dialogue was unscripted.

Two factors were manipulated (Examples 1–4). The first was whether the competitor was in linguistic common ground at the start of the critical question. This was determined by whether the experimenter asked about the competitor or noncompetitor during the set-up question. The second was whether the participant's response to the setup question was grounded by the experimenter. In order to motivate why the experimenter would sometimes not ground the participant's response, the participant was told that, in order to make the task more challenging, the experimenter's pictures would sometimes disappear from her computer screen and, whenever that occurred, the experimenter could not check whether any of her animals matched with the animals that the participant had just mentioned. The experimenter notified the participant when her pictures disappeared and again when they reappeared (~2 sec later). On half of the trials, the experimenter's pictures did not disappear, and she grounded the participant's response to the set-up question. On the remaining trials, the experimenter's pictures did disappear, and the experimenter notified the participant of this just as he or she responded to the setup question. Importantly, in this condition, the experimenter did not ground the participant's response. Thus,

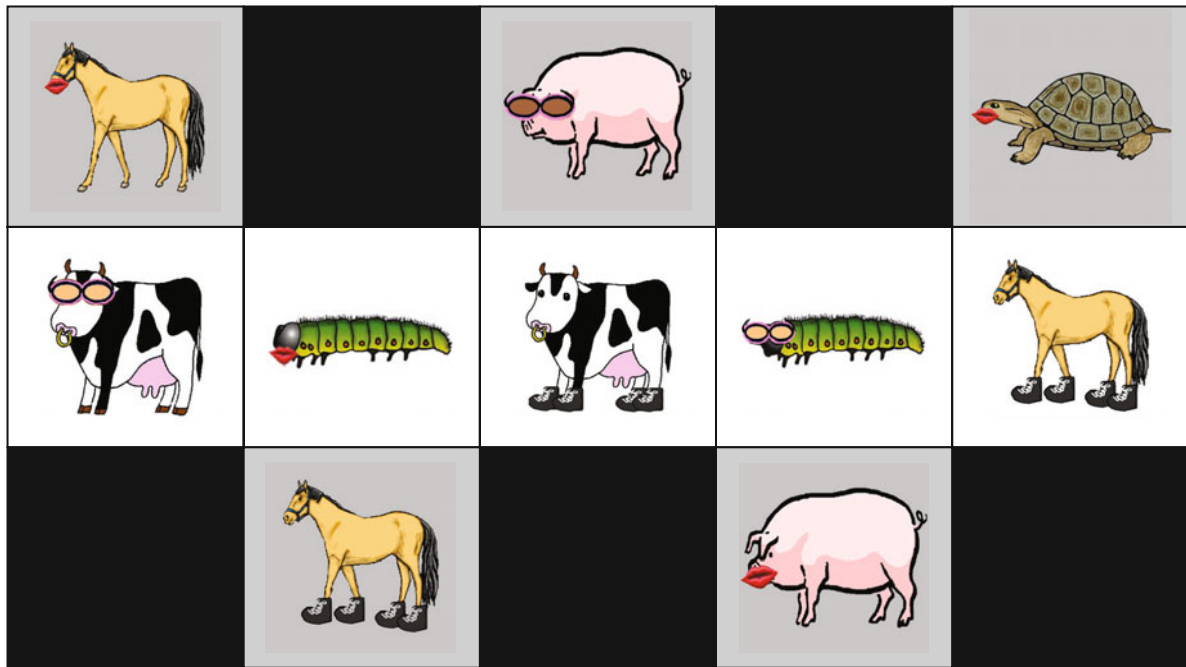
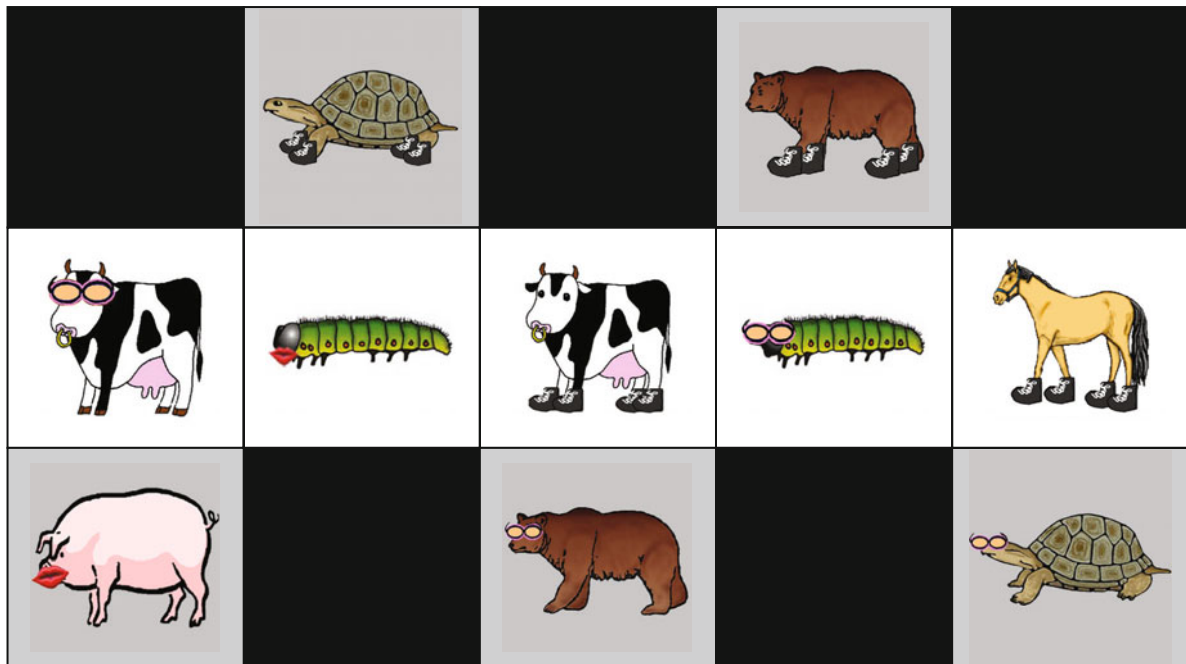
A**B**

Figure 1. (A) Example scene from the participant's perspective. Critical questions were of the form "What's above/below the *x* that's wearing *y*?" Given the question "What's above the cow that's wearing shoes?" the target was the cow with shoes (and the pig with glasses above it), and the competitor was the cow with glasses (and horse with lipstick). When the critical question asked about an animal "below the . . .," the arrangement of shared and privileged squares was flipped, such that there were three addressee-privileged animals on the bottom row. (B) Example scene from the experimenter's perspective.

according to the grounding model, the mentioned competitor (or noncompetitor) did not enter linguistic common ground.

1. Competitor-mentioned, grounded condition

Experimenter: "What's in the top left corner?"

Participant: "A horse wearing lipstick."

Experimenter: "OK. What's below the caterpillar that's wearing glasses?"

Participant: "A pig that's wearing lipstick."

Experimenter: "OK. What's above the cow that's wearing shoes?"

2. Competitor-mentioned, not grounded condition

Experimenter: "What's in the top left corner?"

Participant: "A horse wearing lipstick."

Experimenter: "Whoops, my pictures disappeared. Sorry, I didn't get that. My pictures are back. Uh, what's below the caterpillar that's wearing glasses?"

Participant: "A pig that's wearing lipstick."

Experimenter: "OK. What's above the cow that's wearing shoes?"

3. Non-competitor-mentioned, grounded condition

Experimenter: "What's in the top right corner?"

Participant: "A turtle wearing lipstick."

Experimenter: "OK. What's below the caterpillar that's wearing glasses?"

Participant: "A pig that's wearing lipstick."

Experimenter: "OK. What's above the cow that's wearing shoes?"

4. Non-competitor-mentioned, not grounded condition

Experimenter: "What's in the top right corner?"

Participant: "A turtle wearing lipstick."

Experimenter: "Whoops, my pictures disappeared. Sorry, I didn't get that. My pictures are back. Uh, what's below the caterpillar that's wearing glasses?"

Participant: "A pig that's wearing lipstick."

Experimenter: "OK. What's above the cow that's wearing shoes?"

Each participant was presented with 9 trials in each of the four conditions. The 36 critical trials were combined with 36 filler trials and presented in a single random order. Fillers were identical in structure to critical trials up until the third question, at which point the experimenter asked about a nontarget animal (e.g., "What's below the caterpillar with lipstick?").

Eight lists counterbalanced target and competitor animals and rotated critical trials through the four conditions. Each participant was presented with one list. A prosodic analysis on a random subset of trials (21% of all critical trials) revealed no differences between the conditions in any of the words in the critical questions in pitch, intensity, or duration.

Inhibitory Control Tasks

Two tasks, a verbal task and a nonverbal task, were used to assess participants' inhibitory control.

Nonverbal task. The nonverbal task was modeled after Eigsti et al. (2006). Participants viewed one shape on the screen at a time (circle, heart, square, triangle, x, or star) and were instructed to quickly click each shape except for the star. Each shape was on the

screen for ~1,100 msec, followed by a blank screen for ~70 msec and then the next shape. There were a total of 30 no-go trials, preceded by between 0 and 10 go trials. Responses (whether the participant clicked) were recorded by the computer.

Verbal task. Following the nonverbal task, participants completed a modified version of the Stroop task (Stroop, 1935). During the first phase, participants saw a series of 30 color words on the screen (*red, orange, yellow, green, blue, and purple*). The color of the text was always inconsistent with the color word. Each word was on the screen for ~700 msec, followed by a ~100-msec blank screen. Participants were instructed to read each word aloud. During the second phase, participants were presented with the same 30 color words but were instructed to say the color of the text. The purpose of the first phase was to establish a response (i.e., reading) that would then have to be inhibited in the second phase. Verbal responses during the second phase were recorded as either correct or incorrect (either a miss or a disfluency; e.g., "purp . . . , uh, green").

RESULTS

Eyetracking Analyses

Eye movements associated with interpretation of the critical noun were analyzed in terms of *target advantage* (Arnold, Eisenband, Brown-Schmidt, & Trueswell, 2000), calculated as the proportion of fixations to the target minus the proportion of fixations to the competitor. Following the critical noun, the typical pattern of fixations, given Example 1 and the scene shown in Figure 1A, was a fixation to one of the cows (i.e., animal matching the critical noun), followed by either a fixation to the animal above that cow or a fixation to the other cow and then a fixation to the animal above it. On the basis of this pattern, a target fixation was defined as a fixation either to the animal mentioned in the critical noun phrase or to the animal above or below it (e.g., the cow with shoes or the pig with glasses above it in Figure 1A). A competitor fixation was defined as a fixation either to the animal that temporarily matched the critical noun phrase or to the animal above or below it (e.g., the cow with glasses or the horse with lipstick).

One challenge for visual-world eyetracking analyses is that saccades are state dependent (i.e., one cannot make a saccade to what one is already looking at, and the current fixation can affect future fixations). There is some controversy over how to handle starting state. The approach taken here is to control for state by excluding trials in which the participant was already fixating the target or competitor at critical noun onset (the Appendix includes a supplemental analysis on all trials; see also Heller et al., 2008; Tanenhaus, Frank, Jaeger, Salverda, & Masharov, 2008). The drawback of this approach is data loss (because of the large area encompassed by the target and competitor, 39.6% of trials were excluded). The advantage is that it equates state across conditions and focuses on trials that are likely to be most informative (i.e., those with an equal chance of a saccade to the target or competitor in response to the linguistic input; see Barr, 2008, for an alternative approach).

Target advantage scores were analyzed in three consecutive 400-msec regions. It takes approximately 200 msec to program and launch an eye movement (Hallett, 1986); thus, the analysis regions are offset by 200 msec. Region 1

(200–600 msec poststimulus) was designed to capture the first eye movements made in response to interpreting the critical noun. Region 2 (600–1,000 msec) captured later predisambiguation effects. The disambiguating word (e.g., *shoes*) began, on average, 834 msec following critical noun onset, so interpretation of the disambiguating word should begin to affect fixations around 1,034 msec postnoun. Thus, Region 3 (1,000–1,400 msec) captured primarily postdisambiguation effects.

Target advantage scores (Figure 2) were analyzed in an ANOVA with mention (competitor vs. noncompetitor mentioned), grounding (grounded vs. not grounded), and time region as factors. The ANOVA revealed main effects of region [$F_1(2,94) = 168.27, p < .0001$; $F_2(2,34) = 94.69, p < .0001$] and mention [$F_1(1,47) = 12.25, p < .01$; $F_2(1,17) = 11.04, p < .01$]. The mention effect was due to higher target advantage scores when the competitor was mentioned than when the noncompetitor was mentioned. These main effects were qualified by a significant region \times mention interaction [$F_1(2,94) = 4.80, p < .05$; $F_2(2,34) = 6.16, p < .01$] and a significant grounding \times mention interaction [$F_1(1,47) = 4.75, p < .05$; $F_2(1,17) = 6.87, p < .05$].

Planned comparisons determined that, when the participant's response was grounded, there was a significant effect of mention [$F_1(1,47) = 12.31, p < .01$; $F_2(1,17) = 27.60, p < .0001$] and a mention \times region interaction was significant by participants [$F_1(2,94) = 3.49, p < .05$; $F_2(2,34) = 2.91, p = .07$]. At each region, target advantage scores were significantly higher for the competitor-mentioned than for the non-competitor-mentioned condition ($ts > 2.3$). The mention \times region interaction resulted from the fact that the mention effect grew stronger at Region 2 ($F_s > 6.8$) but did not change between Regions 2

and 3 ($F_s < 3.7$). These results demonstrate that, when the competitor was in common ground, participants used this information to interpret the temporarily ambiguous question as asking about a privileged-ground animal.

In contrast, when the participant's response was not grounded, the main effect of mention was not significant ($F_s < 1.6$), nor was the mention \times region interaction ($F_s < 2.25$).

Inhibitory Control Analyses

In the nonverbal inhibitory-control task, participants incorrectly clicked the picture on 10.1% ($SE = 1.2\%$) of no-go (star) trials. In the verbal (Stroop) task, participants correctly named the color of the text on 77.2% ($SE = 2.1\%$) of trials.² Linear regressions were used to predict performance in the online perspective-taking task on the basis of performance in the two inhibitory-control tasks. Two dependent measures were used. The *perspective-use score* was defined as the difference in target advantage scores between the competitor-mentioned, grounded condition and the non-competitor-mentioned, grounded condition, averaged across Regions 1 and 2. The better that participants were at using perspective information to suppress the perspective-inappropriate response (i.e., interpreting the question as asking about the mentioned competitor), the higher this score should be. The *grounding-use score* was defined as the difference in perspective-use scores between the grounded and the not grounded conditions (calculated in the same way as it was for the grounded conditions). The more sensitive that participants were to whether mentioned information was grounded, the higher this score should be.

Accuracy in the Stroop task accounted for a significant proportion of the variance in online perspective use

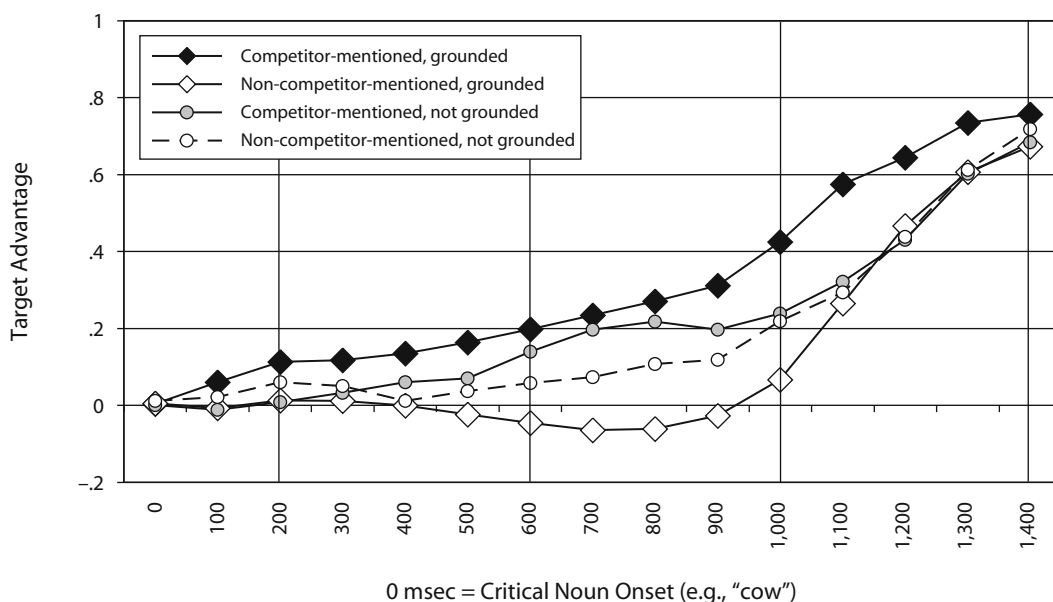


Figure 2. Target advantage scores over time; data are centered at the critical noun onset (e.g., *cow*). Vertical bars indicate analysis regions.

$[F(1,45) = 6.06, p < .05, r^2 = .12]$. The better that participants were at the Stroop task, the better their performance was in perspective use (unstandardized $\beta = 1.05$, $SE = 0.43$; standardized $\beta = 0.35$). Supplemental analyses revealed that this effect was caused primarily by a suppression of competitor fixations ($r^2 = .16, p < .01$) rather than facilitation of target fixations ($r^2 = .06, p = .11$). There was also a significant, positive relationship between Stroop performance and grounding use [$F(1,45) = 5.26, p < .05, r^2 = .11$; unstandardized $\beta = 1.19$, $SE = 0.52$; standardized $\beta = 0.32$]. In contrast, the relationship between error rate in the nonverbal task and perspective or grounding use was not significant ($ps > .7$).

DISCUSSION

Replicating previous findings (Brown-Schmidt et al., 2008), addressees maintained grounded discourse referents in common ground and used this information to resolve temporary ambiguities: As they interpreted a temporarily ambiguous question (e.g., "What's above the cow that's wearing shoes?"), if the object above one of the cows had been mentioned and grounded, addressees directed attention away from it and its associated cow and toward the unmentioned target and its associated cow. This effect was attenuated when the mentioned competitor was not grounded.

Crucially, a verbal measure of inhibitory control significantly predicted individual participants' ability to use perspective representations. A lack of a significant effect for the nonverbal measure of inhibitory control may be due to near-ceiling performance or because it did not involve *conflict* inhibition (in which a new response replaces a dominant one; see Carlson & Moses, 2001).

The starting point for the present article was that models of the use of perspective information in language processing must explain why perspective information sometimes does and sometimes does not appear to constrain online interpretation. The present article suggests that difficulties in inhibiting the perspective-inappropriate interpretation can account for some perspective-taking failures; grounding also appears to play a role. How can these results inform models of online perspective taking?

According to constraint-based accounts, perspective information is one of many partial, probabilistic constraints that inform language-processing decisions (Brown-Schmidt et al., 2008; Hanna & Tanenhaus, 2004; Hanna et al., 2003; Heller et al., 2008; Nadig & Sedivy, 2002). In the present experiment, addressees interpreted noun phrases that were temporarily ambiguous between two potential referents. Goodness of fit between an expression and potential referents is a highly reliable cue to speaker meaning; thus, the spoken input should have strongly activated both target and competitor referents. However, perspective information ruled out the competitor when it was mentioned and grounded. The role of inhibition-control processes, then, was to use perspective to inhibit or reduce activation of the competitor. Addressees with poor inhibition control were less successful at inhibiting the competitor and thus showed

more evidence of perspective-inappropriate interpretations. When perspective information was less certain (i.e., the not grounded conditions), it was weighted less strongly, and the target and competitor became similarly activated. This also resulted in more perspective-inappropriate interpretations. A challenge and a goal for future research will be incorporating inhibition-control processes into implemented constraint-based models in order to better understand the relationship between inhibition and the use of different constraints.

According to autonomous activation (Barr, 2008), addressees are unable to eliminate interference from perspective-inappropriate referents; thus, the observation that inhibition-control plays a key role in perspective use is generally consistent with this model. A challenge for this approach, however, will be explaining why many participants did use perspective. Finally, the basic results are inconsistent with strategic egocentrism (Keysar et al., 2000), because addressees used perspective online in contexts that lacked global ambiguity (see Brown-Schmidt et al., 2008; Heller et al., 2008).

In conclusion, when interpreting language in conversation, addressees use interactively established common-ground representations to select perspective-appropriate interpretations of ambiguities and to inhibit perspective-inappropriate ones. Whether shared or privileged discourse referents are facilitated or inhibited depends on the syntactic position of a given referring expression (for example), as well as utterance form (e.g., question vs. statement). The degree to which addressees are successful is determined partly by individual differences in inhibition control and by the strength of perspective information. Memory for perspective and the distinctiveness of perspectives may also play a role (see Horton & Gerrig, 2005). Finally, an open question is whether the proposed role of inhibition control extends to ambiguity resolution processes in general. If it does, this would predict that, during the interpretation of ambiguities, the lower an individual's inhibition control, the more likely he or she is to consider unintended interpretations, because of difficulty in using various information sources (e.g., semantic, pragmatic) to rule out alternative interpretations.

AUTHOR NOTE

Thank you to Michael Tanenhaus and Duane Watson for helpful comments and discussions. Direct correspondence to S. Brown-Schmidt, Beckman Institute, 405 N. Mathews Ave., Urbana, IL 61801 (e-mail: brownsch@illinois.edu).

REFERENCES

- ARNOLD, J. E., EISENBAND, J. G., BROWN-SCHMIDT, S., & TRUESWELL, J. C. (2000). The immediate use of gender information: Eyetracking evidence of the time-course of pronoun resolution. *Cognition*, *76*, B13-B26.
- BARR, D. J. (2008). Pragmatic expectations at linguistic evidence: Listeners anticipate but do not integrate common ground. *Cognition*, *109*, 18-40.
- BROWN-SCHMIDT, S., GUNLOGSON, C., & TANENHAUS, M. K. (2008). Addressees distinguish shared from private information when interpreting questions during interactive conversation. *Cognition*, *107*, 1122-1134.
- CARLSON, S. M., & MOSES, L. J. (2001). Individual differences in inhibi-

- tory control and children's theory of mind. *Child Development*, **72**, 1032-1053.
- CLARK, H. H., & BRENNAN, S. E. (1991). Grounding in communication. In L. B. Resnick, J. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127-149). Washington, DC: American Psychological Association.
- CLARK, H. H., & MARSHALL, C. R. (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. L. Webber, & I. A. Sag (Eds.), *Elements of discourse understanding* (pp. 10-63). Cambridge: Cambridge University Press.
- EBERHARD, K. M., SPIVEY-KNOWLTON, M. J., SEDIVY, J. C., & TANENHAUS, M. K. (1995). Eye movements as a window into real-time spoken language comprehension in natural contexts. *Journal of Psycholinguistic Research*, **24**, 409-436.
- EIGSTI, I.-M., ZAYAS, V., MISCHEL, W., SHODA, Y., AYDUK, O., DADLANI, M. B., ET AL. (2006). Predicting cognitive control from preschool to late adolescence and young adulthood. *Psychological Science*, **17**, 478-484.
- FRIEDMAN, O., & LESLIE, A. M. (2005). Processing demands in belief-desire reasoning: Inhibition or general difficulty? *Developmental Science*, **8**, 218-225.
- GERNSBACHER, M. A. (1997). Group differences in suppression skill. *Aging, Neuropsychology, & Cognition*, **4**, 175-184.
- HALLETT, P. E. (1986). Eye movements. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (Vol. 1, pp. 10.1-10.112). New York: Wiley.
- HANNA, J. E., & TANENHAUS, M. K. (2004). Pragmatic effects on reference resolution in a collaborative task: Evidence from eye movements. *Cognitive Science*, **28**, 105-115.
- HANNA, J. E., TANENHAUS, M. K., & TRUESWELL, J. C. (2003). The effects of common ground and perspective on domains of referential interpretation. *Journal of Memory & Language*, **49**, 43-61.
- HELLER, D., GRODNER, D., & TANENHAUS, M. K. (2008). The role of perspective in identifying domains of reference. *Cognition*, **108**, 831-836.
- HORTON, W. S., & GERRIG, R. J. (2005). Conversational common ground and memory processes in language production. *Discourse Processes*, **40**, 1-35.
- KEYSAR, B., BARR, D. J., BALIN, J. A., & BRAUNER, J. S. (2000). Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science*, **11**, 32-38.
- KEYSAR, B., LIN, S., & BARR, D. J. (2003). Limits on theory of mind use in adults. *Cognition*, **89**, 25-41.
- NADIG, A. S., & SEDIVY, J. C. (2002). Evidence of perspective-taking constraints in children's on-line reference resolution. *Psychological Science*, **13**, 329-336.
- NILSEN, E. S., & GRAHAM, S. A. (2009). The relations between children's communicative perspective-taking and executive functioning. *Cognitive Psychology*, **58**, 220-249.
- RAMSCAR, M., & GITCHO, N. (2007). Developmental change and the nature of learning in childhood. *Trends in Cognitive Sciences*, **11**, 274-279.
- STALNAKER, R. C. (1978). Assertion. In P. Cole (Ed.), *Pragmatics* (pp. 315-332). New York: Academic Press.
- STROOP, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, **18**, 643-662.
- TANENHAUS, M. K., FRANK, A. F., JAEGER, T. F., SALVERDA, A. P., & MASHAROV, M. (2008, March). *The art of the state: Mixed effects regression modeling in the visual world*. Paper presented at the 21st Annual CUNY Conference on Human Sentence Processing, Chapel Hill, NC.
- WELLMAN, H. M., CROSS, D., & WATSON, J. (2001). Meta-analysis of theory mind development: The truth about false belief. *Child Development*, **72**, 702-707.

NOTES

1. Note that the noncompetitor pair (the horse with shoes and the turtle with lipstick above it) prevents the ambiguity from being resolved at above.
2. The data from 1 participant were dropped because of equipment error.

APPENDIX

Supplementary analyses on all trials (i.e., including those with an initial target or competitor fixation) included a baseline region (−200 to 200 msec) and confirmed the basic pattern observed in the state-dependent analysis, superimposed on an early mention effect (see Figure A1). Significant effects of region ($F_s > 117$) and mention ($F_s > 36$) were qualified by a region \times mention interaction [$F_1(3,141) = 5.92, p < .01$; $F_2(3,51) = 5.87, p < .01$] and a marginal region \times grounding \times mention interaction [$F_1(3,141) = 2.29, p = .08$; $F_2(3,51) = 2.31, p = .09$]. The mention effect emerged well before the onset of the word “What” (regardless of feedback), consistent with a general bias to look at unmentioned entities. However, perspective modulated interpretation of the critical noun: Following *cow*, target advantage scores rose significantly for the competitor-mentioned, grounded condition (i.e., from baseline to Region 1) [$t_1(47) = 3.03, p < .01$; $t_2(17) = 2.45, p < .05$] and marginally for the mentioned, not grounded condition [$t_1(47) = 2.00, p = .051$; $t_2(17) = 1.61, p = .13$]; there was no change when a noncompetitor was mentioned ($t_s < .7$).

Analyses at Regions 1–3 revealed that the mention \times grounding interaction did not become fully significant until Region 3 ($F_s > 5.5$); however, inhibition control modulated the effect (shown in the Figure A1 insets). Individuals with high Stroop scores showed the early mention \times grounding interaction at Region 1 [$F_1(1,24) = 4.91, p < .05$; $F_2(1,17) = 3.65, p = .07$] and Region 2 [$F_1(1,24) = 7.18, p < .05$; $F_2(1,17) = 9.81, p < .01$]; individuals with low inhibition control did not ($F_s < 2.8$).

Finally, regression analyses on the full data set confirmed a significant, positive relationship between Stroop performance and both perspective use [$F(1,45) = 6.50, p < .05, r^2 = .13$; unstandardized $\beta = 0.61, SE = 0.24$; standardized $\beta = 0.36$] and grounding use [$F(1,45) = 7.56, p < .01, r^2 = .14$; unstandardized $\beta = 1.25, SE = 0.45$; standardized $\beta = 0.38$].

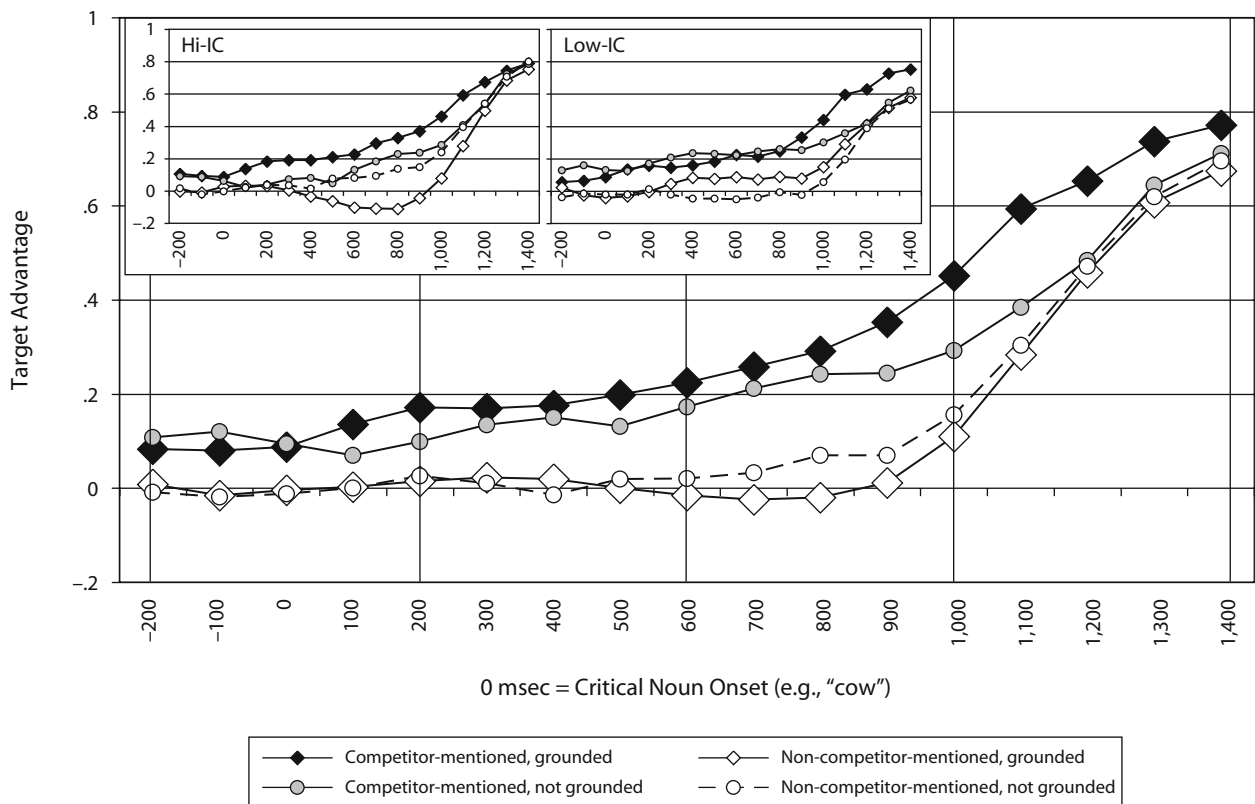


Figure A1. Target advantage scores over time (all trials); data are centered at the critical noun onset (e.g., *cow*). Vertical bars indicate analysis regions. Insets show data for participants grouped by performance on the Stroop task: High-inhibition control (Hi-IC) mean Stroop score = 88% ($SD = 7\%$, $n = 25$) and low-IC mean = 65% ($SD = 10\%$, $n = 22$).