Developmental Psychology
2002, Vol. 38, No. 3, 352–362

0012-1649/02/$5.00 DOI: 10.1037//0012-1649.38.3.352

Conditions Under Which Young Children Can Hold Two Rules in Mind and Inhibit a Prepotent Response

Adele Diamond, Natasha Kirkham, and Dima Amso
Eunice Kennedy Shriver Center, University of Massachusetts Medical School

The day–night task requires saying "night" to a picture of the sun and "day" to a picture of the moon. In this investigation of why young children fail at this task, systematic variations of the task were administered to 96 children, half 4 years old and half 4½ years old. Training children on the strategy of chunking the 2 rules into 1 ("say the opposite"), thus reducing memory load, did not help their performance. What helped was reducing the inhibitory demand by instructing them to say "dog" and "pig" (not "night" and "day") even though memory of 2 rules and inhibiting saying what the pictures represented were still required. Here the response to be activated and the response to be inhibited were unrelated. When the correct response was semantically related to, and the direct opposite of, the to-be-inhibited response, children performed poorly. Inserting a delay between stimulus and response helped even though that delay was filled with distraction. Young children apparently need several seconds to compute the answer on this task. Often they do not take the needed time; when forced to do so, they do well.

It is a well-replicated finding that children 3–4 years of age have great difficulty guiding their actions by rules held in mind that require acting contrary to their inclinations. For example, they perform poorly on the day–night task (Gerstadt, Hong, & Diamond, 1994), which requires that they hold two rules in mind ("Say 'day' to black/moon cards. Say 'night' to white/sun cards.") and resist the temptation to say what the stimuli really represent. Children 3–4 years of age also perform poorly on the tapping task (Diamond & Taylor, 1996; Luria, 1966), which requires that they remember two rules ("Tap once when the experimenter taps twice. Tap twice when the experimenter taps once.") and inhibit the tendency to mimic the experimenter's actions. They perform poorly on the three pegs task (Balamore & Wozniak, 1984), which requires that they remember the sequence in which they should tap three pegs (red-green-yellow) and inhibit the tendency to tap the pegs in the order in which they have been placed in the pegboard (red-yellow-green). Children 3–4 years of age also have difficulty in delay-of-gratification paradigms (Mischel & Mischel, 1983), in which they must (a) remember that if they wait they will receive a better reward and (b) inhibit the temptation to not wait and to reach immediately for the available, but lesser, reward. Most 3-year-olds, and many 4-year-olds, are unable to switch sorting dimensions in the standard condition of Zelazo's card sorting task (Kirkham, Cruess, & Diamond, 2000; Zelazo, Frye, & Rapus, 1996), which requires that they remember the currently relevant dimension and its associated set of rules plus inhibit the pull to attend to the previously correct dimension.

A major difference between a child in Piaget's "preoperational" stage (3–4 years) and one in the "concrete operations" stage (5–7 years) is that the older child can simultaneously hold more than one thing in mind and inhibit the strongest response of the moment. For example, children 4 years of age fail tests of liquid conservation. They do not attend to both height and width. They attend only to the most perceptually salient of the two dimensions. They also fail tests of perspective taking, in which they must hold more than one perspective in mind and inhibit the tendency to give the most salient response (the one corresponding to their own perspective).

Furthermore, theory-of-mind and false-belief tasks (e.g., Perner, Leekham, & Wimmer, 1987) can also be thought of as tests of working memory and inhibitory control: The child must hold two ideas in mind about the same situation (the true state of affairs and the false belief of another person) and inhibit giving the answer he or she knows to be true in favor of reporting the knowledge of another (Carlson, Moses, & Hix, 1998). In other words, 3–4-year-old children respond to questions about another person's perspective by stating their own perspectives. By 5 years of age, they can answer those questions correctly.

Why is it that preschoolers have difficulty with such tasks? Is it because they have trouble remembering what they are supposed to do, or is it because they have trouble stopping themselves from making the most salient response? Under what conditions can

Adele Diamond, Natasha Kirkham, and Dima Amso, Center for Developmental Cognitive Neuroscience, Eunice Kennedy Shriver Center, University of Massachusetts Medical School.

Natasha Kirkham and Dima Amso are now at the Department of Psychology, Cornell University.

This research was supported by National Institutes of Health Grants R01 MH41842 and R01 HD35453.

We would particularly like to thank Susan Power, who came up with the idea for the ditty condition and who began piloting the ditty and dog–pig conditions until family obligations required that she move away. We would also like to thank Lorrie Gehlbach for her help with data analyses and all the children, without whose participation this research would not have been possible.

Correspondence concerning this article should be addressed to Adele Diamond, Center for Developmental Cognitive Neuroscience, Eunice Kennedy Shriver Center Campus, University of Massachusetts Medical School, 200 Trapelo Road, Waltham, Massachusetts 02452. E-mail: adele.diamond@umassmed.edu

4-year-old children succeed in holding rules in mind and inhibiting a dominant response? In this study we investigated these questions, focusing on one of the experimental paradigms mentioned above, the day–night task.

Background on the Day–Night Task

The day–night task was designed to investigate children's ability to act according to remembered instructions that require inhibiting a prepotent response tendency. Gerstadt and colleagues (1994) originally called the task the "Stroop-like day–night task" because, like the Stroop task (MacLeod, 1991; Stroop, 1935), it requires that one act according to remembered instructions and inhibit a prepotent response. In the Stroop task, the response that must be inhibited is the strong tendency to attend to the meaning of words and to ignore surface characteristics, such as the color in which the words are printed. The Stroop task requires that one do the opposite and say the color of the ink rather than read the word (i.e., the correct response to the word *blue* printed in red ink would be "red").

Children 3½–4½ years old find the day–night task extremely difficult; for children 6–7 years old, it is trivially easy (Gerstadt et al., 1994). Passler, Isaac, and Hynd (1985) administered a similar, though slightly easier, variant of the task that required children to recognize the correct answer instead of recalling it. They found that children 6 years of age performed at ceiling, consistent with the excellent performance of children that age on the day–night task.

The reason that younger children fail the day–night task is not that they have difficulty holding two rules in mind, because they perform splendidly if instructed to say "day" to one abstract design and "night" to another (Gerstadt et al., 1994). This abstract-designs condition still requires learning and remembering two rules, but it does not also require inhibiting the tendency to say what the stimuli really represent because the abstract designs do not represent anything in particular.

Predictions

If age-related improvements in performance on the day–night task are charting the development of the ability to exercise inhibitory control (i.e., if immature inhibitory control is at least part of the reason why children 4 or 4½ years old fail the task), then reducing the inhibitory requirement should help 4- and 4½-year-old children succeed. The abstract design variant of the task provided evidence consistent with that prediction (Gerstadt et al., 1994). The present investigation provides another test of that prediction. Unlike the abstract design variant, the black/moon and white/sun cards were retained here, but the response options were changed. Instead of saying "day" to the moon and "night" to the sun, children were instructed, for example, to say "pig" to the moon and "dog" to the sun. This reduced the demand on inhibitory control because the correct responses ("dog" and "pig") were not semantically related to the responses to be inhibited ("day" and "night"). We predicted that more children would succeed. The memory demand, of course, remained the same.

Gerstadt and colleagues (1994) observed that some children 5–7 years old spontaneously reported collapsing the two lower order rules in the day–night task into a single higher order rule ("say the opposite"). No younger children ever reported doing that. It is possible that older children performed better on the day–night task than younger children because older children simplified the task for themselves, reducing the number of rules that had to be held in mind from two to one. In the present investigation we explored that possible explanation by instructing all children in one condition to "say the opposite" to each card. Could younger children perform better if we trained them on the strategy that at least some older children appeared to use? If remembering two rules is the problem, younger children should perform well here. If inhibition is the problem, younger children should perform at the same level here as when given two rules to remember.

Gerstadt et al. (1994) also noticed that younger children performed better on those trials on which they took longer to respond. This finding suggested that when younger children took the time they needed, they could succeed. Thus, those younger children who had longer response latencies performed better than other children of the same age who answered more quickly. In addition, most 4-year-olds started out responding correctly and taking a long time to generate each response. Over the course of the 16-trial session, they answered progressively more quickly and progressively got more answers wrong. Thus, within the same child, when the child took longer to respond, the child was more likely to respond correctly. This finding led to our next prediction: Because formulating the response appears to be sufficiently difficult for 4-year-old children that they need several seconds to execute that computation, providing them with more time to respond should improve their performance.

The younger children in the Gerstadt et al. (1994) study increased their speed of responding over the course of the testing session even though they were never rushed, did not know they were being timed, and were given as long as they wanted to respond. It was not obvious, therefore, how we could help children in the present study sustain a slow rate of responding over the 16 test trials. Asking 4-year-olds to wait before responding is a fruitless exercise. So, in order to give them more time to formulate their answers, on each trial after the stimulus was revealed the experimenter sang a little ditty ("Think about the answer, don't tell me"), after which the child could respond. This increased the time between stimulus presentation and the child's response, and although it filled that time with a verbal chant, we predicted that more children would succeed because they would have more time to formulate their responses.

The "ditty" condition slowed down testing and increased the duration of the session. To control for the possibility that children might perform better because of the slower pace of testing, we presented another condition, one in which the experimenter sang the same ditty but sang it between trials, just before the stimulus was revealed for the next trial. Here the pace of testing and duration of the testing session were identical to those in the ditty condition, but the extra time was not provided while the children were computing their responses. Of course, the extra time before a trial might allow children to gather their thoughts and remind themselves of the rules of the "game." However, because the extra time came before they knew which stimulus would be presented, and so could not be used for computing the answer, we predicted that children would not find this condition significantly easier than the standard condition.

In the ditty condition, the stimulus was visible for a longer time than in the other conditions. To keep the stimulus presentation time more equal across conditions, we administered another ditty condition in which the stimulus card was immediately turned over after it was presented and remained facedown while the experimenter sang the ditty. Note that this imposed a delay during which the child had to hold in mind what the stimulus on that trial had been. Here, not only was "reference memory" (Olton & Samuelson, 1976) of the rules pertaining to all trials required, but updating the contents of working memory was also necessary (on each trial the child had to ask himself or herself, "Which stimulus did I see most recently?"). In addition, children had to inhibit the tendency to say what the stimulus really represented. If the extra computation time helped, even in the face of the additional demand on working memory, that would indeed be impressive.

Thus, we administered six variations of the day–night task (the standard condition, dog–pig, say the opposite, ditty, ditty between trials, and ditty + memory) to investigate the effect on task performance of reducing the memory load (say the opposite), reducing the inhibitory demand (dog–pig), and giving children more time to access the right answer and inhibit the wrong answer (ditty, plus its control conditions: ditty between trials and ditty + memory). Each child participated in only one condition.

Method

Participants

Results for 96 children are reported here: 48 children were 4 years old (24 girls and 24 boys; mean age = 48.3 months), and 48 children were 4½ years old (24 girls and 24 boys; mean age = 55.6 months). Table 1 provides a breakdown of age means and ranges by experimental condition. Past research with the day–night task has shown that the time of the greatest improvement on the task is the period between 4 and 5 years of age (e.g., Gerstadt et al., 1994). Thus, to look at what might help children having difficulty with the task, as was our goal here, we chose to study children between 4 and 4½ years old. Many studies lump together 4-year-olds and 4½-year-olds, 5-year-olds and 5½-year-olds, and so forth. It has consistently been found, however, that there are important differences between these age groups (e.g., Gerstadt et al., 1994), just as important differences have been found between infants of 7 and 7½ months, 8 and 8½ months, and so forth (e.g., Diamond, 1985). We expected that 4-year-olds would have significantly more difficulty with the day–night task than would 4½-year-olds and that the conditions we most expected to help (the dog–pig and ditty conditions) would show a more prominent effect with 4-year-olds because they would have more room for improvement than 4½-year-olds.

Children were recruited from local Boston-area daycare centers, through word of mouth and through a database of parents who had expressed interest in participating in research. Informed consent was obtained from parents of all children, the rights of all study participants were protected, and all children received a present for their participation. All children had been full-term and were healthy. Most were from middle-class homes and of European Caucasian descent. In addition to the 96 children mentioned above, another 6 children who were 4 years of age (3 girls and 3 boys; mean age = 3.9 years) were tested on an earlier version of the dog–pig condition in which the labels the children were to learn were "dog" and "cat." In addition to the children included in the data analyses, another 7 children (five 4-year-olds and two 4½-year-olds) were tested, but their data were not usable. Three failed to show that they understood what was being asked of them, 3 did not pay attention during testing, and 1 consistently responded before the stimulus was presented and so gave no valid responses. After testing, we discovered that 1 participant had attention-deficit/hyperactivity disorder; that child's data were excluded from the analyses as well.

Table 1
Mean Ages and Age Ranges, in Months, of Children in Each Condition

Condition	*N*	Age *M*	Age *SD*	Age Range
4-year-olds				
Standard	8	4.0	0.15	3.7–4.2
Dog–pig	8	4.1	0.01	4.0–4.2
Say the opposite	8	4.0	0.12	3.8–4.2
Ditty	8	4.1	0.11	3.9–4.2
Ditty between trials	8	4.0	0.17	3.7–4.2
Ditty + memory	8	4.0	0.13	3.7–4.2
4½-year-olds				
Standard	8	4.5	0.01	4.4–4.6
Dog–pig	8	4.6	0.12	4.5–4.8
Say the opposite	8	4.7	0.18	4.5–4.9
Ditty	8	4.6	0.17	4.4–4.9
Ditty between trials	8	4.7	0.11	4.5–4.8
Ditty + memory	8	4.7	0.18	4.5–4.9

Materials

The same 16 cards were used for all conditions. The cards were cardboard rectangles (14 cm × 10 cm). The front side of half the cards was black with a yellow moon and several silver stars. The front side of the other eight cards was white with a large bright-yellow sun. The backs of all of the cards were cream-colored.

Procedure

Training phase. During the training phase preceding testing, the experimenter turned over a black card depicting the moon and stars and instructed the child to say "day" to that card, and then turned over a white card depicting the sun and instructed the child to say "night" when shown that card. The child was also quizzed on what response to give to each card. To pass training, the child had to respond correctly to each card on two consecutive trials within the six allotted training trials. During each training and testing trial for all conditions (except ditty + memory), the stimulus card remained visible directly in front of the child until the child responded. The only responses considered correct were "night" and "day." All other responses, such as "day" to the white/sun card,"morning," "sun," or "suppertime," were corrected. If a child gave an incorrect response to either card, the experimenter repeated the instructions beginning with the "fragile" rule first (i.e., the rule on which the child had erred).

Testing phase. As soon as a child responded correctly to two training cards in a row, testing began without further repetition of the rules. If a child was incorrect on either of the first two test trials, the experimenter reiterated both rules, the fragile rule first. There was never any feedback, correction, or reiteration of the rules after Trials 1 and 2 regardless of the child's performance.

If the child had not passed training after six trials, the experimenter began the test trials with a final explanation of both rules before the presentation of the first test card. If the child answered incorrectly on either of the first two test trials after having failed training, the child was considered not to have understood the task, and testing was discontinued. For those children who failed to show understanding of the rules during training but demonstrated that understanding during the first two test trials, testing was continued without further repetition of the rules or feedback. This applied to only 2 children (one 4-year-old in the standard condition and one 4-year-old in the ditty + memory condition).

The test phase for all conditions consisted of 16 cards and lasted 4–5 min. A trial consisted of the presentation of one card and the child's response. Cards were presented so that the child could not see anything on the front of the card until the last moment, because seeing the color of the card was sufficient to determine whether it was a white/sun or black/moon card. The experimenter kept the card facedown as he or she moved it toward the child, and only then flipped it over. The intertrial interval (ITI, the time between the end of the child's response and the beginning of the next trial) was kept constant across trials and conditions except in the ditty-between-trials condition, in which a longer ITI was intended. Eight white/sun cards (S) and 8 black/moon cards (M) were presented in the same pseudorandom order in all conditions (S, M, M, S, M, S, S, M, M, S, M, S, S, M, S, M). Never was a child rushed, told he or she was being timed, or asked to respond as quickly as possible. The procedure described so far is the same as the procedure used by Gerstadt et al. (1994).

Procedures specific to the dog–pig condition. The dog–pig condition differed from the standard procedure only in the rules that were taught. Children were instructed to say "dog" to one of the cards and "pig" to the other. Half the children were instructed to say "dog" to the black/moon card and half to say "dog" to the white/sun card.

Procedures specific to the dog–cat condition. The dog–cat condition differed from the standard procedure only in the rules that were taught. Children were instructed to say "dog" to one of the cards and "cat" to the other. Half the children were instructed to say "dog" to the black/moon card and half to say "dog" to the white/sun card. This was an earlier variant of the dog–pig condition that we decided to discontinue for two reasons. One was the concern that the word *cat* might be associated in some children's minds with the black/moon scene. The other reason that the term *pig* was substituted for the word *cat* is that *dog* and *cat* are semantically related, and so we thought it might be easier to retrieve the second rule if only one of the rules was clearly remembered. We thought that two unrelated words such as *dog* and *pig* would more clearly require the memory of two separate rules and thus that success here would more definitively show that memory of two separate associations was not the problem.

However, an alternative perspective was suggested to us: that not having a semantic relation between the two words to be remembered might make the task easier and that the dog–pig condition had both a lack of semantic relation between the to-be-remembered items and a lack of relation between those items and the stimuli with which they were paired. In the dog–cat condition, those two features were not confounded, and so to address that concern, we present the results for the dog–cat condition here as well.

Procedures specific to the say-the-opposite condition. This condition differed from the standard condition only in how the instructions were worded. Instead of telling children to say "day" to one card and "night" to the other, children were told to "say the opposite" of what each card represented. It was explained that the black/moon card represented night so that the correct answer upon seeing that card was "day," and that the white/sun card represented day so that the correct response to that card was "night," because "this is the opposite game."

Procedures specific to the ditty condition. The ditty condition differed from the standard condition in that after the training phase was finished, the experimenter said to the child, "We're going to try some more cards, but now, before you tell me the answer I want you to wait. While you're waiting, I'm going to sing a little song that goes like this, 'Think about the answer; don't tell me.' Wait to tell me the answer until after I have finished my little song." If the child answered before the experimenter finished the song, the experimenter said, "Shhhh—remember to wait until I finish the song before you give an answer," and went on to the next card. This rarely happened (3 times over 256 trials [16 trials × 16 children]).

The ditty condition differed from the standard condition in three ways: time to compute the answer, pace of testing / length of a testing session, and amount of time the stimulus was presented. What we were interested in was the effect of giving children more time to compute the answer. Therefore, we tested two control conditions: The ditty-between-trials condition controlled for the pace of testing / length of the testing session, and the ditty + memory condition controlled for stimulus presentation time but introduced an additional memory requirement.

Procedures specific to the ditty-between-trials condition. This condition was administered in the same way as the standard condition except that the ditty was sung before each trial. After the child responded, the experimenter sang the ditty and then the next trial began. The child had no more time to respond to a given card in this condition than in the standard condition, but the duration of the 16-trial session was the same as in the ditty condition.

Procedures specific to the ditty + memory condition. This condition was administered in the same way as the ditty condition except that after the child was shown the card, it was flipped over. Thus, during the time that the experimenter sang the ditty, the face of the card was not visible. The stimulus was visible for as long on each trial here as in the standard condition, but this condition imposed an additional memory demand. The child was not allowed to respond when the face of the card was still visible and so had to remember what he or she had seen until the ditty was over.

Results

A 2 (gender) × 2 (age group: 4 years and 4½ years) × 6 (condition: standard, dog–pig, say the opposite, ditty, ditty between trials, and ditty + memory) analysis of variance was conducted. The dependent variable was the percentage of correct responses. There was a main effect of age; 4½-year-olds responded correctly significantly more often than did 4-year-olds (4 years, $M = 70\%$; 4½ years, $M = 87\%$), $F(1, 72) = 10.99$, $p < .001$. There was also a main effect of condition, $F(5, 72) = 5.70$, $p < .001$.

Planned comparisons showed, as predicted, that children performed significantly better in the dog–pig condition than in the standard condition (94% vs. 65%), $t(19) = 5.68$, $p < .001$ (this t test is corrected for unequal variances, as are others below where appropriate). Better performance in the dog–pig condition than in the standard condition was seen both in 4-year-olds (92% vs. 53%), $t(11) = 7.87$, $p < .001$, and in 4½-year-olds (97% vs. 78%), $t(8) = 3.07$, $p < .02$. See Figure 1.

The dog–cat condition was administered only to children 4 years old. Performance in the dog–cat condition was comparable to that in the dog–pig condition (96% vs. 92%) and significantly better than that in the standard condition (96% vs. 65%), $t(12) = 8.64$, $p < .0001$.

There was a slight tendency for the 4-year-olds tested in the standard condition to be younger than the 4-year-olds tested in the dog–pig condition. To make sure that the observed difference in performance between the dog–pig and standard conditions was not an artifact of this slight age difference, we compared performance in these two conditions (a) omitting the 2 children in the standard condition who were younger than 4.0 years (92% vs. 49%), $t(9) = 8.48$, $p < .001$, (b) omitting the 2 oldest children in the dog–pig condition (92% vs. 53%), $t(9) = 6.83$, $p < .001$, and (c) omitting all 4 of those children (92% vs. 49%), $t(7) = 7.34$, $p < .001$. Clearly, the robust difference in performance between the

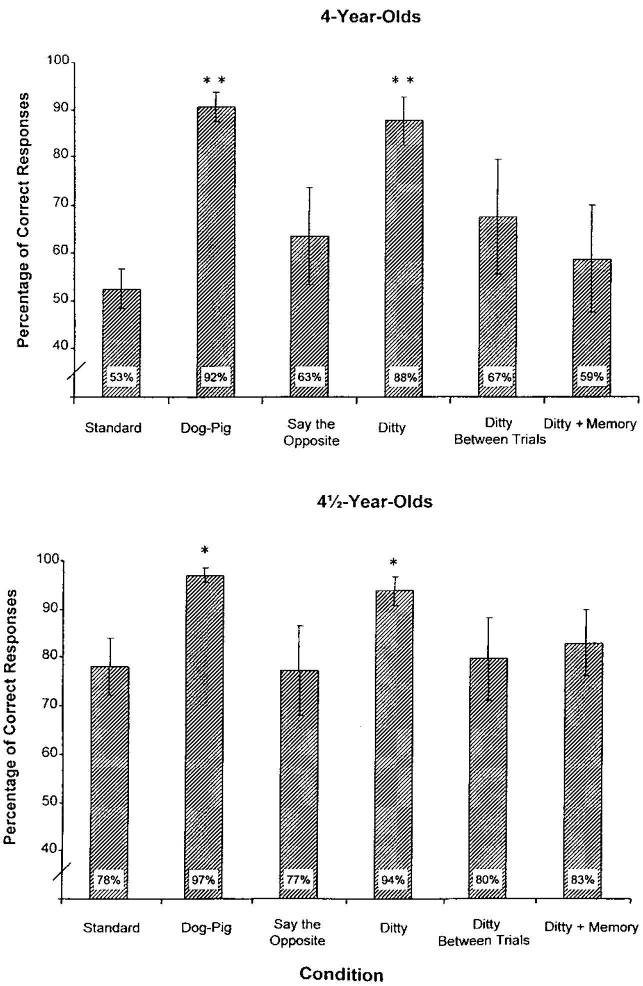

Figure 1. Performance of children in each condition of the day–night task. Clearly, for 4-year-olds, the dog–pig and ditty conditions were easy, and all the other conditions were much more difficult. Already by 4½ years, children were able to perform better on the task, even in the more difficult conditions. Vertical lines depict standard errors. *Performance significantly better than performance in the standard condition, $p < .05$. **Performance significantly better than performance in the standard condition, $p < .001$.

dog–pig and standard conditions at 4 years of age is not in any way due to the slight age difference in the children tested in those two conditions.

Also as predicted, children performed significantly better in the ditty condition than in the standard condition (91% vs. 65%): planned comparison, $t(25) = 4.56$, $p < .001$ (corrected for unequal variances). Better performance in the ditty condition than in the standard condition was seen among the 4-year-olds (88% vs. 53%), $t(13) = 5.42$, $p < .001$, and among the 4½-year-olds (94% vs. 78%), $t(10) = 2.36$, $p < .04$. This better performance cannot be attributed to the slower rate of testing in the ditty condition, because the rate of testing was just as slow in the ditty-between-trials condition and yet performance was not significantly better there than in the standard condition (74% vs. 65%, *ns*), and performance was significantly better in the ditty condition than in the ditty-between-trials condition (91% vs. 74%), $t(20) = 2.18$, $p < .05$.

Consistent with our hypothesis that it is the inhibitory demand of the task, rather than the memory demand, that causes children difficulty, when we reduced the memory demand from two rules (say "day" to the black/moon card and "night" to the white/sun card) to one rule (say the opposite of what each card represents), children performed no better. Thus, performance in the say-the-opposite condition was not significantly better than performance in the standard condition overall (70% vs. 65%, *ns*), among 4-year-olds (63% vs. 53%, *ns*), or among 4½-year-olds (77% vs. 78%, *ns*).

Adding a delay over which children had to remember the stimulus they had seen last wiped out much of the advantage conferred by the ditty condition. Thus, performance did not differ significantly in the ditty + memory and standard conditions overall (71% vs. 65%, *ns*), among 4-year-olds (59% vs. 53%, *ns*), or among 4½-year-olds (83% vs. 78%, *ns*). The performance of 4-year-olds was significantly better in the ditty condition than in the ditty + memory condition (88% vs. 59%), $t(10) = 2.35$, $p < .05$. However, among 4½-year-olds, performance did not differ significantly in the ditty + memory and ditty conditions (83% vs. 94%, *ns*). Thus, for 4½-year-olds, memory was a more minimal part of the reason they had difficulty with the task, because adding the additional memory demand resulted in their performance being between that in the standard and ditty conditions, significantly different from neither.

There was no significant main effect for gender. However, there was a significant interaction between gender and age, $F(1, 72) = 10.11$, $p < .002$. Girls started out performing worse than boys: boys vs. girls, 4 years, $F(1, 92) = 6.64$, $p < .01$. This gender difference at 4 years was driven solely by the ditty-between-trials condition. Four-year-old boys and girls performed comparably in every condition except the ditty-between-trials condition, which helped boys a great deal (their percentage of correct responses in that condition was 92%) and which helped girls not at all (their percentage of correct responses there was 44%).

By 4½ years, the girls had more than caught up. Indeed, at 4½ years, there was a nonsignificant trend for girls to outperform boys, $F(1, 92) = 3.24$, $p < .08$. By 4½ years, the girls had essentially mastered our task, performing well in all conditions. Boys, on the other hand, showed no improvement on the task between 4 and 4½ years. Thus, the difference in performance between 4 and 4½ years was significant for girls (means = 62% and 90%, respectively), $F(1, 92) = 19.39$, $p < .001$, but not for boys (means = 79% and 79%; $F < 1$).

The gender difference at 4 years in the ditty-between-trials condition may have been an accident based on the particular children who happened to have participated in this study. We have never encountered a gender difference before on this task or on any related tasks (such as the tapping or three pegs tasks). However, it may be that young boys, who have a tendency to be more impulsive than girls, can benefit from a manipulation that causes them to slow down whether it occurs after they have seen the stimulus or before. Why the 4-year-old boys could benefit from extra time between trials during which they could gather their thoughts and remind themselves of the rules of the game but the 4-year-old girls could not remains a question that perhaps future studies will be able to answer.

From the videotape records we were able to code response latency for those children for whom we received permission to videotape (just over half the sessions). We coded the reaction times for the standard, dog–pig, and say-the-opposite conditions. For each of those conditions, we coded no fewer than half the sessions (two sessions of boys and two sessions of girls at each age) and no more than three of the boys' or three of the girls' sessions in any condition at any age. Omitted were the two conditions in which a delay had intentionally been inserted between presentation of the stimulus and when the child could respond (ditty and ditty + memory). In addition, too few of the ditty-between-trials sessions were filmed for us to be able to include that condition. These data provide an indication of whether children responded faster in the condition they found easier (the dog–pig condition, in which they responded correctly on more of the trials than in the standard or say-the-opposite conditions) or whether perhaps the children were able to perform better in the dog–pig condition because they took longer to compute their answers.

The results show that 4-year-old children responded slightly more slowly in the dog–pig condition than in the standard condition (reaction times [RTs] = 1.23 and 1.49 s, respectively). However, 4-year-olds responded even more slowly in the say-the-opposite condition (RT = 1.60 s); here their performance was significantly worse than in the dog–pig condition and almost as poor as in the standard condition (see Figure 2). Children 4½ years old, who also performed significantly better in the dog–pig condition than they did in the standard condition, responded faster in the dog–pig condition than in the standard condition (RTs = 1.03 and 1.47 s, respectively) and, like the 4-year-olds, showed the longest response latency in the say-the-opposite condition (RT = 1.58 s; see Figure 2). None of these RTs is anywhere near as long as the roughly 5–6-s RTs seen in the ditty condition, in which the experimenter's chanting was interposed before the child could respond. The time taken to respond cannot account for the differences in performance observed among the standard, dog–pig, and say-the-opposite conditions.

Replicating Gerstadt et al.'s (1994) finding, we found that 4-year-olds and 4½-year-olds started out responding at a slow rate but speeded up their responses over the course of the 16 trials. This can be seen in the mean RTs over the first four trials and the last four trials (see Figure 3). This is true of all three conditions examined, at both ages, except for 4½-year-olds in the say-the-opposite condition.

Discussion

Children 4 and 4½ years of age have difficulty consistently saying "night" to a white/sun card and "day" to a black/moon card. Earlier work from our laboratory had shown that if you change the stimuli to abstract designs (that have no particular words associated with them and hence present no prepotent response to inhibit) but keep the responses the same, even 4-year-olds have no difficulty consistently saying "day" to one stimulus and "night" to the other (Gerstadt et al., 1994). Here we have shown that if you keep the stimuli the same (the white/sun and black/moon cards) but change the responses so that the correct response is not semantically related to the to-be-inhibited response (i.e., saying "dog" or "pig" to a picture of the sun or moon), even children 4 years old succeed easily.[1] What we did not know before from the abstract-design manipulation, but can conclude now from the dog–pig manipulation, is that it is not that children of 4 or 4½ years are unable to inhibit saying what a stimulus represents even if it is a highly representational stimulus with obvious word associations. The problem for children of 4 or 4½ years lies in inhibiting a word that is semantically related to the word they are trying to say.

This finding is consistent with the large literature on associative networks and spreading activation (e.g., J. R. Anderson, 1983a, 1983b; M. C. Anderson & Spellman, 1995; Balota & Duchek, 1989; Neely, 1977). For example, activating the word *night* also activates the semantically related word *day*, which is counterproductive for the standard day–night task. In contrast, activating the word *dog* might also activate words like *cat* or *bone*, but those are not words closely associated with a picture of the sun or moon. Similarly, it is not a problem if in trying to activate the word *night* the word *day* is also activated *if* the stimulus is an abstract design that carries no association with the word *day*. These findings are consistent with those from studies using the directed-forgetting paradigm which demonstrate that children are more likely to recall words they were instructed to forget if those words are semantically related to words they were instructed to remember than if they are unrelated (Harnishfeger & Pope, 1996; Lehman, Srokowski, Hall, Renkey, & Cruz, 2000).

The present findings are also consistent with those from studies of the classic Stroop task with adults which show that interference decreases and performance improves as the strength of the semantic relation between the to-be-ignored word and the correct response (the color of the ink) decreases (e.g., Dalrymple-Alford, 1972; Klein, 1964; Klopfer, 1996; Stirling, 1979). For example, the possible responses (say, e.g., green, blue, and yellow) are more primed or prepotent than the names of other colors not in the set of eligible responses (say, e.g., red, orange, and violet). One might think of green, blue, and yellow as forming an associative network in this situation over and above the associative links between any color and any other color. Consistent with that reasoning, Klein (1964) found the usual Stroop interference effect when the written word named an eligible color (e.g., green), but the interference effect was much reduced on response-ineligible trials in which the word named a color not in the response set (e.g., red). Similarly, Dalrymple-Alford (1972) showed that the Stroop interference effect was significantly greater for the word *blood* (which is strongly associated with red) written in green ink and the word *grass* (strongly associated with green) written in red ink than for the words *joy* or *hand* (neither of which have strong color associations) written in green or red ink. The joy–hand condition is perhaps analogous to our dog–pig condition; the correct response

[1] In the standard condition, the responses were both semantically related to the stimuli and semantic opposites of one another. In the dog–pig condition, the responses were not semantically related to the stimuli and not semantic opposites of one another. We know from the abstract-designs condition of Gerstadt et al. (1994) that if the responses are not semantically related to the stimuli (indeed, if the stimuli do not call forth any particular semantic associations) but the stimuli are semantic opposites of one another, children 4 years of age succeed beautifully. They are able to consistently say "day" to one abstract design and "night" to the other. Further, we know from the dog–cat condition administered in the early stages of the present study that even if the stimuli remain the white/sun cards and the black/moon cards, and even if the responses are semantic opposites of one another, as long as the responses are not semantically related to the stimuli, children 4 years of age succeed beautifully.

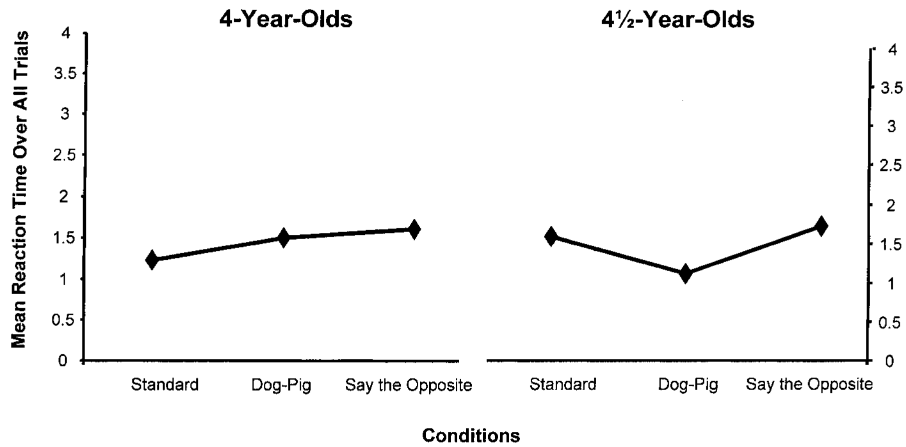

Figure 2. Mean reaction times (in seconds) of the children in the standard, dog–pig, and say-the-opposite conditions of the day–night task for whom permission was granted to videotape their sessions (just over half of the children tested in these three conditions).

was not semantically related to the response to be inhibited, although the stimuli do invoke obvious word associations.

The day–night task is sufficiently difficult for young children that it takes them several seconds to compute the answer. Often, they do not take the time they need. The ditty condition shows that when children at least as young as 4 years are forced to take extra time, even if that time is filled with verbal distraction, they can perform well. Similarly, Heberle, Clune, and Kelly (1999) reported that if children 3–4 years old are made to wait before responding on an appearance–reality task, they perform significantly better. Our finding is also consistent with results from a study with a directional Stroop task, in which it was found that given enough time (stimulus presentation times of 2,500 ms), even 4-year-olds could perform the task (Davidson, Cruess, Diamond, O'Craven, & Savoy, 1999). Gerstadt et al. (1994) found that when children 3½–4½ years old took longer to respond on the day–night task, they were more likely to respond correctly. This association between faster responding and poorer performance decreased with age; older children could still respond correctly when they responded rapidly. There is certainly a substantial body of work showing that younger children have slower processing speeds (e.g., Hale, 1990; Kail, 1988, 1991).

It is not the case, however, that children succeeded in the dog–pig condition because they took more time in that condition to compose their responses. Children of 4½ years responded more quickly in the easy dog–pig condition than they did in the harder standard and say-the-opposite conditions. Children of 4 years responded slightly more slowly in the dog–pig condition than in the standard condition, but they responded still more slowly in the say-the-opposite condition, in which they performed as poorly as they did in the standard condition. All of these response times were one half to one third of the response times seen in the ditty condition, in which the experimenter's little chant was interposed between presentation of the stimulus and the child's response.

It should be noted that the ditty was not only a potential distraction, it was also a reminder to comply with the rules ("Think about the answer; don't tell me."). It is interesting that providing this same reminder before presenting the stimulus had little effect

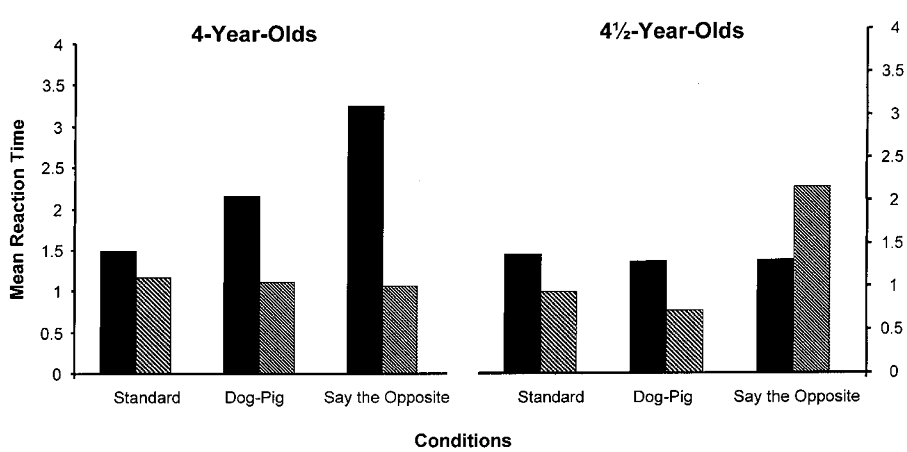

Figure 3. Mean reaction times (in seconds) of children on the first 4 and last 4 trials (out of 16 total trials) in the standard, dog–pig, and say-the-opposite conditions of the day–night task.

(as Simpson & Riggs [2000] also found). In future studies it might be of interest to determine if any ditty (even with words unrelated to the task at hand) would be as effective as the ditty condition used here.

What Neural System Might Underlie the Ability to Succeed on the Day–Night Task?

There is considerable evidence that tasks that require holding information in mind while inhibiting a dominant response (as the day–night task does) involve dorso- and ventro-lateral prefrontal cortex (e.g., Bunge, Ochsner, Desmond, Glover, & Gabrieli, 2001; Cohen, Braver, & O'Reilly, 1996; Diamond, O'Craven, & Savoy, 1998; Roberts & Pennington, 1996). Indeed, patients with prefrontal cortex damage perform poorly on the Stroop task (Perret, 1974). Performance on a task closely analogous to the day–night task, the tapping task ("Tap once when the experimenter taps twice, and tap twice when the experimenter taps once."), increases activation in dorsolateral prefrontal cortex in normal adults in comparison with "Tap once when the experimenter taps once, and tap twice when the experimenter taps twice" (Brass, Zysset, & von Cramon, 2001). On tasks in which participants must (a) continuously monitor a series of stimuli, (b) remember a rule that specifies they should press the response button every time except if a particular condition is met, and (c) inhibit pressing the button when that condition is met, activation of dorsolateral prefrontal cortex increases immediately before successful inhibition (Garavan, Ross, & Stein, 2001).

Moreover, aspects of the response patterns across tasks of patients with damage to prefrontal cortex are similar to the response patterns of 4- and 4½-year-old children in the present study. For instance, although it is well established that frontal lobe patients perform poorly on many cognitive tasks, it has also been repeatedly observed that when given more time to respond, they can perform well on those tasks. For example, Corkin (1964) found that frontal lobe patients had longer response times than other patients or controls of the same age and that when given more time to respond, frontal lobe patients were much more likely to respond correctly. Similarly, Swaab, Brown, and Hagoort (1998) reported that patients with Broca's aphasia could not use context to select the appropriate meaning of a word at short interstimulus intervals but that when these patients were given more time, they did fine. That is, given enough time, they could use context to select the contextually appropriate meaning of ambiguous words. This effect was not due simply to age, because elderly control participants were much more successful at the short interstimulus intervals than were the patients with Broca's aphasia. Thus, we offer as a hypothesis that the developmental improvement in the ability to perform well on the day–night task even when responding rapidly may be made possible, in part, by maturational changes in prefrontal cortex. Indeed, the *generalized* improvement in speed of processing with development (Kail, 1991), besides reflecting increased myelination and other changes, may also reflect maturational changes in prefrontal cortex. Better prefrontal functioning makes possible faster and more efficient domain-general processing by reducing signal:noise ratios throughout diverse neural regions (Knight, Hillyard, Woods, & Neville, 1981; Swick & Knight, 1996).

Other neural regions that appear to figure prominently in the circuit that enables people to hold information in mind and inhibit distracting associations or action tendencies are the anterior cingulate cortex (e.g., Banich et al., 2000; Bench et al., 1993; Bush, Luu, & Posner, 2000; Carter, Botvinick, & Cohen, 1999; Casey, Cohen, Noll, Forman, & Rapoport, 1993; Pardo, Pardo, Janer, & Raichle, 1990), posterior parietal cortex (e.g., Honey, Bullmore, & Sharma, 2000; Jansma, Ramsey, Coppola, & Kahn, 2000; Smith, Jonides, Marshuetz, & Koeppe, 1998), and the cerebellum (e.g., Awh, Jonides, Smith, Schumacher, Koeppe, & Katz, 1996; Berman et al., 1995; Diamond, 2000; Nagahama et al., 1996).

What Are the Demands of the Day–Night Task That Cause Such Difficulty for Children Roughly 4 Years of Age?

The abstract-designs condition of Gerstadt et al. (1994) demonstrated that the problem for children is not simply holding two rules in mind, because that condition required holding two rules in mind and even the youngest children performed splendidly. Is the problem suppressing a dominant response when it is related to the response that needs to be activated (inhibiting saying what the stimuli really represent when the correct response is to say the opposite), or is the problem the conjunction of having to hold two rules in mind and inhibiting a dominant response that is related to the correct response? That is, is the problem inhibition or memory plus inhibition? From the data at hand one cannot conclusively determine which of these possibilities is correct, but there are indications that even with little or no demand on memory, 4-year-olds have difficulty inhibiting a prepotent response.

Reducing the number of rules from two lower level rules to one higher order rule ("say the opposite") did not aid children's performance, which is consistent with the inhibitory requirement's being the main source of difficulty. It is also consistent with interpretations that children younger than 4 years cannot fully grasp an abstract hierarchical rule (e.g., Halford, Wilson, & Phillips, 1998; Zelazo & Frye, 1997), so that although such a rule was taught to our children, the 4-year-olds might not have understood it. If they did not understand the abstract rule instruction, however, worse performance might have been expected in the say-the-opposite condition than in any other condition, and that was not found. Conclusions about the lack of effect of the say-the-opposite condition should be tempered by the possibility that had the children been given more extensive training on the concept of saying the opposite, perhaps they might have been able to perform better in that condition.

Simpson and Riggs (2000) administered the day–night task to children but changed the procedure in one respect—they gave feedback that reminded the children of the rules on every trial. They found comparably poor performance in the young children. Those results provide further evidence that minimizing the memory requirement does not help children perform better on the task and that the inhibitory demand is the critical element.

The Simpson and Riggs (2000) results are consistent with the many examples showing that while children younger than 4 years of age remember what the correct response is, and can say it, they are unable to inhibit a contrary prepotent response. For example, Tikhomirov (1978) told children to squeeze a bulb when a red light came on and not to squeeze the bulb when a green light came on.

Children 3–4 years old could repeat back the instructions correctly, indicating that they remembered them, but they failed the task because they could not inhibit responding on the no-go trials. Livesey and Morgan (1991) obtained exactly comparable results with a different go/no-go task. By the time children are about 5 years old, they succeed on such go/no-go tasks (Livesey, n.d.; Livesey & Morgan, 1991; Miller, Shelton, & Flavell, 1970).

Similarly, on Zelazo's card sort task (Zelazo et al., 1996; Zelazo, Frye, Reznick, Schuster, & Argitis, 1995, cited in Zelazo & Jacques, 1997), in which children are to sort a deck of cards first by color or shape and then by the other dimension, the problem appears to be one of inhibition. Children 3–4 years old sort correctly according to the first dimension (whether it is color or shape), but when the dimension changes, they cannot inhibit continuing to regard the stimuli in terms of the first dimension and continuing to sort by that dimension (a) even though they can correctly tell you what the new dimension is and what the rules for sorting by that dimension are and (b) even if they are reminded of the new dimension and its rules on each trial before they respond (Kirkham et al., 2000; Zelazo et al., 1996).

Besides these results showing inhibitory errors in conditions that reduced the memory demand or that tried to make sure the children's memory was accurate, we found that increasing the memory demand (in the ditty + memory condition) did not impair the performance of 4½-year-olds. However, it did significantly mitigate the beneficial effect of having more time to respond for 4-year-olds. Thus, holding two rules in mind plus remembering which stimulus one saw last plus inhibiting a dominant response appears to be too difficult for 4-year-olds even if they are given more time.

Inhibitory abilities undergo an extremely slow, protracted developmental progression, not reaching full maturity until early adulthood (Diamond, 2002). Thus, for example, in the directed forgetting paradigm (in which participants are directed to forget some words they are shown and to remember others), even 11-year-olds show more intrusions of the to-be-forgotten words than do adults (e.g., Harnishfeger & Pope, 1996; Lehman et al., 2000). Similarly, on the "anti-saccade" task (in which participants are instructed to look away in the opposite direction from a visual stimulus, suppressing the tendency to reflexively look at [saccade to] the stimulus), performance improves continuously from early childhood through 20–25 years of age (Fischer, Biscaldi, & Gezeck, 1997; Luna et al., 2001; Munoz, Broughton, Goldring, & Armstrong, 1998). With memory demand held constant, children show a long developmental progression from 4 to 14 years of age in the ability to inhibit the prepotent tendency to respond on the side where a stimulus appears, rather than the opposite side, and even 14-year-olds are not yet at adult levels (Davidson et al., 1999).

In conclusion, the dog–pig condition refines our understanding of the inhibitory demands of the day–night task and of the inhibitory abilities of young children. Children at least as young as 4 years can inhibit saying what a stimulus represents even when they also have to hold two rules in mind (e.g., they can say "dog" to the white/sun card and "pig" to the black/moon card, and they can keep that up accurately over 16 trials). The relation between the response to be activated and the response to be suppressed is key. What children of 4, or even 4½, years are unable to do consistently is inhibit saying what a stimulus represents if the correct response is semantically related to, and directly opposite of, the to-be-inhibited response. This inability to inhibit occurs even if the children need hold only one rule in mind (as in the say-the-opposite condition). To a large extent, even very young children know what the correct responses are on the day–night task: Their problem lies in getting the responses they make to reflect the knowledge they have; when forced to take more time before they answer, they can answer correctly.

References

Anderson, J. R. (1983a). Retrieval of information from long-term memory. *Science*, *220*, 25–30.

Anderson, J. R. (1983b). A spreading activation theory of memory. *Journal of Verbal Learning & Verbal Behavior*, *22*, 261–295.

Anderson, M. C., & Spellman, B. A. (1995). On the status of inhibitory mechanisms in cognition: Memory retrieval as a model case. *Psychological Review*, *102*, 68–100.

Awh, E., Jonides, J., Smith, E. E., Schumacher, E. H., Koeppe, R. A., & Katz, S. (1996). Dissociation of storage and rehearsal in verbal working memory: Evidence from positron emission tomography. *Psychological Science*, *7*, 25–31.

Balamore, U., & Wozniak, R. H. (1984). Speech–action coordination in young children. *Developmental Psychology*, *20*, 850–858.

Balota, D. A., & Duchek, J. M. (1989). Spreading activation in episodic memory: Further evidence for age independence. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *41*, 849–876.

Banich, M., Milham, M., Cohen, N. J., Wszalek, T., Kramer, A., Liang, Z. P., et al. (2000). fMRI studies of Stroop tasks reveal unique roles of the anterior and posterior brain system in attentional selection. *Journal of Cognitive Neuroscience*, *12*, 988–1000.

Bench, C. J., Frith, C. D., Grasby, P. M., Friston, K. J., Paulesu, E., Frackowiak, R. S., & Dolan R. J. (1993). Investigations of the functional anatomy of attention using the Stroop test. *Neuropsychologia*, *31*, 907–922.

Berman, K. F., Ostrem, J. L., Randolph, C., Gold, J., Goldberg, T. E., Coppola, R., et al. (1995). Physiological activation of a cortical network during performance of the Wisconsin Card Sorting Test: A positron emission tomography study. *Neuropsychologia*, *33*, 1027–1046.

Brass, M., Zysset, S., & von Cramon, D. Y. (2001, March). *The inhibition of imitative response tendencies: A functional MRI study.* Poster session presented at the annual meeting of the Cognitive Neuroscience Society, New York, NY.

Bunge, S. A., Ochsner, K. N., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. (2001). Prefrontal regions involved in keeping information in and out of mind. *Brain*, *124*, 2074–2086.

Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences*, *4*, 215–222.

Carlson, S. M., Moses, L. J., & Hix, H. R. (1998). The role of inhibitory processes in young children's difficulties with deception and false belief. *Child Development*, *69*, 672–691.

Carter, C. S., Botvinick, M. M., & Cohen, J. D. (1999). The contribution of the anterior cingulate cortex to executive processes in cognition. *Reviews in the Neurosciences*, *10*, 49–57.

Casey, B. J., Cohen, J. D., Noll, D. C., Forman, S., & Rapoport, J. L. (1993). Activation of the anterior cingulate during the Stroop conflict paradigm using functional MRI [Abstract]. *Society for Neuroscience Abstracts*, *19*, 1285.

Cohen, J. D., Braver, T. S., & O'Reilly, R. C. (1996). A computational approach to prefrontal cortex, cognitive control and schizophrenia: Recent developments and current challenges. *Philosophical Transactions of the Royal Society (London) Series B*, *351*, 1515–1527.

Corkin, S. (1964). *Somesthetic function after focal cerebral damage in man*. Unpublished doctoral dissertation, McGill University, Montreal, Quebec, Canada.

Dalrymple-Alford, E. C. (1972). Associative facilitation and interference in the Stroop color–word task. *Perception & Psychophysics, 11*, 274–276.

Davidson, M., Cruess, L., Diamond, A., O'Craven, K. M., & Savoy, R. L. (1999, April). *Comparison of executive functions in children and adults using directional Stroop tasks.* Paper presented at the biennial meeting of the Society for Research in Child Development, Albuquerque, NM.

Diamond, A. (1985). Development of the ability to use recall to guide action, as indicated by infants' performance on A-not-B. *Child Development, 56*, 868–883.

Diamond, A. (2000). Close interrelation of motor development and cognitive development and of the cerebellum and prefrontal cortex. *Child Development, 71,* 44–56.

Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 466–503). London: Oxford University Press.

Diamond, A., O'Craven, K. M., & Savoy, R. L. (1998). Dorsolateral prefrontal cortex contributions to working memory and inhibition as revealed by fMRI [Abstract]. *Society for Neuroscience Abstracts, 24,* 1251.

Diamond, A., & Taylor, C. (1996). Development of an aspect of executive control: Development of the abilities to remember what I said and to "Do as I say, not as I do." *Developmental Psychobiology, 29*, 315–334.

Fischer, B., Biscaldi, M., & Gezeck, S. (1997). On the development of voluntary and reflexive components in human saccade generation. *Brain Research, 754,* 285–297.

Garavan, H., Ross, T. J., & Stein, E. A. (2001, March). *Dissociating executive functions: The roles of right prefrontal cortex and anterior cingulate in inhibitory control and error detection.* Poster session presented at the annual meeting of the Cognitive Neuroscience Society, New York, NY.

Gerstadt, C., Hong, Y., & Diamond, A. (1994). The relationship between cognition and action: Performance of 3½-7 year old children on a Stroop-like day–night test. *Cognition, 53,* 129–153.

Hale, S. (1990). A global development trend in cognitive processing speed. *Child Development, 61*, 653–663.

Halford, G. S., Wilson, W. H., & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behavioral and Brain Sciences, 21,* 803–864.

Harnishfeger, K. K., & Pope, R. S. (1996). Intending to forget: The development of cognitive inhibition in directed forgetting. *Journal of Experimental Child Psychology, 62*, 292–315.

Heberle, J., Clune, M., & Kelly, K. (1999, April). *Development of young children's understanding of the appearance–reality distinction.* Paper presented at the meeting of the Society for Research in Child Development, Albuquerque, NM.

Honey, G. D., Bullmore, E. T., & Sharma, T. (2000). Prolonged reaction time to a verbal working memory task predicts increased power of posterior parietal cortical activation. *Neuroimage, 12,* 495–503.

Jansma, J. M., Ramsey, N. F., Coppola, R., & Kahn, R. S. (2000). Specific versus nonspecific brain activity in a parametric N-back task. *Neuroimage, 12*, 688–697.

Kail, R. (1988). Developmental functions for speeds of cognitive processes. *Journal of Experimental Child Psychology, 45*, 339–364.

Kail, R. (1991). Developmental change in speed of processing during childhood and adolescence. *Psychological Bulletin, 109*, 490–501.

Kirkham, N., Cruess, L., & Diamond, A. (2000). *Helping children apply their knowledge to their behavior on a dimension-switching task.* Manuscript submitted for publication.

Klein, G. S. (1964). Semantic power measured through the interference of words with color-naming. *Americal Journal of Psychology, 77*, 576–588.

Klopfer, D. S. (1996). Stroop interference and color–word similarity. *Psychological Science, 7,* 150–157.

Knight, R. T., Hillyard, S. A., Woods, D. L., & Neville, H. J. (1981). The effects of frontal cortex lesions on event-related potentials during auditory selective attention. *Electroencephalography and Clinical Neurophysiology, 52,* 571–582.

Lehman, E., Srokowski, S. A., Hall, L. C., Renkey, M. E., & Cruz, C. A. (2000). *Directed forgetting of related words: Evidence for the inefficient inhibition hypothesis*. Manuscript submitted for publication.

Livesey, D. J. (n.d.). *The development of response inhibition in four and five year old children.* Unpublished manuscript, University of Sydney, Sydney, Australia.

Livesey, D. J., & Morgan, G. A. (1991). The development of response inhibition in 4- and 5-year-old children. *Australian Journal of Psychology, 43*, 133–137.

Luna, B., Thulborn, K. R., Munoz, D. P., Merriam, E. P., Garver, K. E., Minshew, N. J., et al. (2001). Maturation of widely distributed brain function subserves cognitive development. *Neuroimage, 13*, 786–793.

Luria, A. R. (1966). *The higher cortical functions in man*. New York: Basic Books.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109*, 163–203.

Miller, S. A., Shelton, J., & Flavell, J. H. (1970). A test of Luria's hypotheses concerning the development of verbal self-regulation. *Child Development, 41*, 651–665.

Mischel, H. N., & Mischel, W. (1983). The development of children's knowledge of self-control strategies. *Child Development, 54*, 603–619.

Munoz, D., Broughton, J., Goldring, J., & Armstrong, I. (1998). Age-related performance of human subjects on saccadic eye movement tasks. *Experimental Brain Research, 217,* 10.

Nagahama, Y., Fukuyama, H., Yamauchi, H., Matsuzaki, S., Konishi, J., Shibaski, H., & Kimura, J. (1996). Cerebral activation during performance of a card sorting test. *Brain, 119*, 1667–1675.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General, 106*, 226–254.

Olton, D. S., & Samuelson, R. J. (1976). Remembrance of places passed: Spatial memory in rats. *Journal of Experimental Psychology: Animal Behavior Processes, 2,* 97–115.

Pardo, J. V., Pardo, P. J., Janer, K. W., & Raichle, M. E. (1990). The anterior cingulate cortex mediates processing selection in the Stroop attentional conflict paradigm. *Proceedings of the National Academy of Sciences, USA, 87*, 256–259.

Passler, P. A., Isaac, W., & Hynd, G. W. (1985). Neuropsychological development of behavior attributed to frontal lobe functioning in children. *Developmental Neuropsychology, 4,* 349–370.

Perner, J., Leekham, S. R., & Wimmer, H. (1987). Three-year-olds' difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology, 5,* 125–137.

Perret, E. (1974). The left frontal lobe of man and the suppression of habitual responses in verbal categorical behaviour. *Neuropsychologia, 12,* 527–537.

Roberts, R. J., Jr., & Pennington, B. F. (1996). An interactive framework for examining prefrontal cognitive processes. *Developmental Neuropsychology, 12,* 105–126.

Simpson, A., & Riggs, K. J. (2000). *The non-veridical Stroop task as a tool to investigate inhibitory and memory development.* Manuscript submitted for publication.

Smith, E. E., Jonides, J., Marshuetz, C., & Koeppe, R. A. (1998). Components of verbal working memory: Evidence from neuroimaging. *Proceedings of the National Academy of Sciences, USA, 95,* 876–882.

Stirling, N. (1979). Stroop interference: An input and an output phenomenon. *Quarterly Journal of Experimental Psychology, 31,* 121–132.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643–662.

Swaab, T. Y., Brown, C., & Hagoort, P. (1998). Understanding ambiguous words in sentence contexts: Electrophysiological evidence for delayed contextual selection in Broca's aphasia. *Neuropsychologia, 36,* 737–761.

Swick, D., & Knight, R. T. (1966). Dorsolateral prefrontal cortex modulates visual processing in extrastriate cortex [Abstract]. *Society for Neuroscience Abstracts, 22,* 1107.

Tikhomirov, O. K. (1978). The formation of voluntary movements in children of preschool age. In M. Cole (Ed.), *The selected writings of A. R. Luria* (pp. 229–269). White Plains, NY: M. E. Sharpe.

Zelazo, P. D., & Frye, D. (1997). Cognitive complexity and control: A theory of the development of deliberate reasoning and intentional action. In M. Stamenov (Ed.), *Language structure, discourse, and the access to consciousness* (pp. 113–153). Philadelphia: John Benjamins.

Zelazo, P. D., Frye, D., & Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cognitive Development, 11,* 37–63.

Zelazo, P. D., & Jacques, S. (1997). Children's rule use: Representation, reflection, and cognitive control. *Annals of Child Development, 12,* 119–126.

Received April 2, 2001
Revision received November 15, 2001
Accepted November 15, 2001 ■