Title
Is perseveration caused by inhibition failure? Evidence from preschool children's inferences about word meanings

Journal
Journal of Experimental Child Psychology, 86(3)

ISSN
0022-0965

Authors
Deák, GO
Narasimham, G

Publication Date
2003-11-01

DOI
10.1016/j.jecp.2003.08.001

Peer reviewed

Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
J. Experimental Child Psychology 86 (2003) 194–222
Journal of Experimental Child Psychology
www.elsevier.com/locate/jecp

Is perseveration caused by inhibition failure? Evidence from preschool children's inferences about word meanings

Gedeon O. Deák[a,*] and Gayathri Narasimham[b]

[a] *Department of Cognitive Science, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92037-0515, USA*
[b] *Department of Psychology and Human Development, Vanderbilt University, USA*

Received 13 November 2001; revised 13 August 2003

Abstract

Four studies examined the relation between children's cognitive inhibition and flexibility in a lexical inference task. Children's linguistic flexibility was assessed by the Flexible Induction of Meaning (FIM) test (Deák, 2000a), which requires that children shift inferences about the meanings of several words for novel objects. In Study 1, 54 3-year-olds either were trained between blocks of problems, for a delay of 3 min, or received no training or delay. Training delays did not influence perseveration. In Study 2 ($N = 72$ 3- and 4-year-olds') novel word problems were grouped either to increase the frequency of cue switches (i.e., reduce response "set") or minimize the interval between problems about the same objects. Again, no effect was found. In Study 3, 48 3- and 4-year-olds completed 6 preliminary trials; in a high interference group these trials generated a response set to be inhibited upon the first switch to a new cue context. This group did not perseverate more than a control group. There was no association between FIM perseveration and a Stroop-like test of verbal inhibition though both were marginally related to receptive vocabulary. In study 4 (48 3- and 4-year-olds), FIM was again unrelated to Stroop performance, but was related to the ability to tell whether a situation or problem is indeterminate. Thus, flexibility across semantic inferences is not influenced by timing, order, and number of pre-switch problems and is not predicted by individual differences in a test of verbal inhibition. However previously reported age and individual differences in flexible induction of word meanings are robust and related to vocabulary and logical ability.

* Corresponding author. Fax: 1-858-534-1128.
E-mail address: deak@cogsci.ucsd.edu (G.O. Deák).

0022-0965/$ - see front matter
doi:10.1016/j.jecp.2003.08.001

Keywords: Flexible cognition; Inference; Inhibition; Interference; Stroop; Vocabulary; Word learning

Introduction

Cognitive flexibility is the capacity to shift inferences or strategies in response to changing cues and task demands. This capacity plays a critical role in everyday language processing, because referential and conceptual shifts are ubiquitous events during discourse. For example, in conversation we must track concepts as the locutions that refer to them change (e.g., "deficit" and "California" might change to "it" and "there"). Similarly, to follow a story, listeners must rapidly activate and flexibly modify mappings between conceptual entities and terms in the teller's utterances (see Fauconnier, 1997). This requires flexible shifting of associations between words and phrases and conceptual representations, which in turn requires selective attention to linguistic and paralinguistic cues that signal changes in reference and meaning.

Fluent adults often can shift their referential mappings and underlying meanings with little difficulty (assuming that, e.g., speakers are clear and informative). Young children, by contrast, have difficulty in using dynamically changing linguistic, paralinguistic, and non-linguistic cues to adapt their representations of meaning. Limitations of vocabulary and understanding of linguistic cues certainly contribute to these problems. In addition, basic processes of cognitive flexibility seem to be immature in young children (Deák, 2003). Preschool children make many *perseverative* errors: inappropriate repetitions responses over a series of questions or problems (Luria, 1959). Children perseverate in tests of rule use, word learning, and naming, and in non-verbal tasks (Gerstadt, Hong, & Diamond, 1994; Johnson, 1994; Sophian & Wellman, 1983; Wertlieb & Rose, 1979). For example, infants tend to search for a toy in the last place they found it, not where it was last hidden (Piaget, 1954). Two-year-olds perseverate when sorting a series of items, even if they know enough to sort them correctly (Zelazo & Reznick, 1991).

The current study investigated how preschool children's tendency to perseverate interacts with the demands for flexible cognition that arise in language processing. We used the *Flexible Induction of Meaning*, or *FIM*, test (Deák, 2000a), which assesses how children infer meanings of different words for an array, based on changing linguistic cues to meaning.

In the FIM test children hear three words for each of six sets of novel objects. Each set includes a standard and four comparison objects: one with the same body shape as the standard, one made of the same material, one with an identical part, and one dissimilar foil (see Fig. 1 for example). On each trial a standard object is described by a unique novel word following one of three predicates: "*looks like a(n)____*," "*is made of____*," or "*has a(n)____*." Children must generalize each word to a comparison object, based on an inference about the word's referential meaning. Five- and 6-year-olds consistently generalize words predicated by "looks like a..." to same-shape objects, words predicated by "is made of..." to same-material objects, and words predicated by "has a..." to same-part objects.

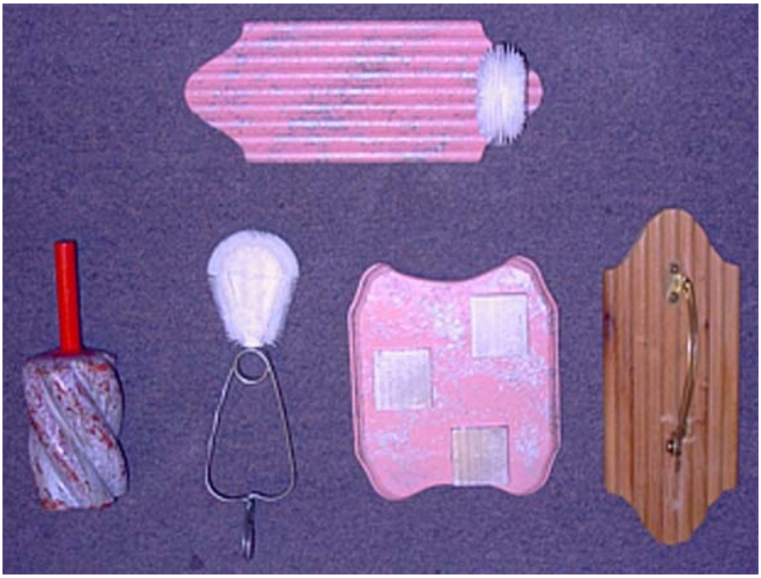

Fig. 1. Sample object set from the FIM: (top) Standard; (bottom, left to right) foil, same-part, same-material, and same-shape objects.

Three-year-olds also seem, in simplified tasks, to understand these predicates (Deák, 2003). Optimally, then, preschool children should be able to use predicate cues to correctly infer the meanings of words for properties of the objects. The FIM imposes a demand for cognitive flexibility, however, because children hear several words for each set, and so in later trials they must ignore the objects to which they generalized previous words.

Flexibility in the FIM varies over age groups and individual children. Predicate-appropriate inferences in the original sample (Deák, 2000a) increased from a mean of 7.9 (out of 18) in 3-year-olds, to 11.5 in 4-year-olds and 15.0 in 5- and 6-year-olds. A critical finding is that 4–6-year-olds made predicate-appropriate inferences about the first *and* later words for a set, whereas 3-year-olds made correct inferences only about the first word for a set. Each child could be classified as *flexible*, *perseverative*, or *indiscriminate*, based on total predicate-appropriate responses and on number of switches between inferences about successive words. Three-year-olds divided almost evenly into the three response types, whereas two-thirds of 4-year-olds and almost all 5- and 6-year-olds were flexible the remaining minority were perseverative. In all age groups, the dominant alternative to flexible inferences about word meanings was perseveration on one inferred meaning.

Perseveration is often attributed to a failure to inhibit a dominant or practiced response (e.g., Bjorkland & Harnishfeger, 1990; Dempster, 1992; Houdé, 2000). *Cognitive inhibition* is a vague construct that must be better specified to test its relation to children's language processing flexibility and perseveration. Inhibitory failure might explain children's perseveration over inferences about word meanings, but other cognitive factors, like children's *construal* of changing task cues (Deák & Bauer, 1995), might also contribute. To explore whether developmental changes in cognitive inhibition can adequately explain young children's inflexible inferences about successive words, we conducted four experiments. We tested factors (e.g., delay between problems) that should, according to a current model of inhibitory development, affect how hard it is for children to inhibit previous inferences in the FIM. If the incidence of perseveration changes with these factors, it will indicate that developing inhibitory processes, as currently understood, can explain young children's perseverative errors. If, instead, perseveration is unaffected, it will fail to confirm the presumed causal relation between perseveration and cognitive inhibition, at least with respect to flexible inferences about meaning. To contextualize the factors manipulated here, we first describe a model of cognitive inhibition and its development, then explain the utility of the FIM for evaluating this model.

What is cognitive inhibition?

Cognitive inhibition is the suppression of competing or primed responses, foci of attention, or associations. It is believed to play a significant role in visual search (Neill, Valdes, & Terry, 1995), memory (Peterson & Peterson, 1959), word retrieval (Gernsbacher & Faust, 1995), and problem solving (Duncker, 1945). Yet the foregoing definition leaves unanswered several crucial questions. For example, is cognitive inhibition one or several capacities? Is it centralized, or specialized for different response and processing systems (e.g., Olson, 1989)? Is it the same as resistance to interference? Is it deliberate and conscious, passive and unconscious, or both? *What* is inhibited in flexible performance—actions, associations, attention, representations, or all of these? Such questions underscore the vague relation between cognitive inhibition and flexible cognitive and language. Under some operationalizations of inhibition, virtually *any* error can be attributed to inhibitory failure, it is particularly tempting to attribute perseverative errors to inhibitory failure. However, perseveration can stem from non-inhibitory processes (Deák, 2003). Therefore, to sharpen the conceptual relation of cognitive inhibition to children's errors, we outline a *modal model* of inhibitory development. This model suggests factors that should moderate children's perseveration, and therefore can be tested.

Inhibitory development: A modal model

Many researchers construe inhibition as a resource- or capacity-demanding process (Dempster, 1992; Guttentag & Ornstein, 1990; Harnishfeger, 1995; Logan & Cowan, 1984; May, Kane, & Hasher, 1995; Miller, Seier, Probert, & Aloise, 1991; Neumann & Deschepper, 1992; Ridderinkhof, van der Molen, Band, & Bashore, 1997). As cognitive load or interference increases, inhibition tends to fail. Inhibition acts on activated, *interfering* responses, cues, or representations. More highly active interfering responses (or cues or representations) demand more cognitive inhibitory resources. With increasing age children gain resources, presumably as frontal cortical regions and pathways mature (Diamond & Goldman-Rakic, 1989). Perseveration thereby declines.

The modal model predicts that certain factors should moderate the difficulty of inhibition (i.e., demands for inhibitory resources) and thereby the probability of perseverative errors. For example, interference should increase with the frequency and recency of prior responses, so perseveration should increase as the interval between successive competing inferences becomes shorter, or as the number of repetitions of an earlier response or task cue increases.

Another premise of the modal model is that the inhibitory capacity is a general resource which can be construed as a stable trait that varies across age groups and individual children (Carlson & Moses, 2001; Olson, 1989; Reed, Pien, & Rothbart, 1984). By this premise, poor performance on a test of inhibition should predict perseveration in a test of flexibility. Given these premises, if perseveration does not vary with frequency and recency of prior responses, and if individual children's performance on tests of inhibition does not predict perseveration, we would have reason to question whether the modal model of inhibitory development can adequately explain children's perseverative errors.

Perseveration and induction of word meanings

The *Flexible Induction of Meaning* (FIM) test (Deák, 2000a) requires preschool children to infer meanings for several words about a stimulus array, based on changing *predicate cues* that implies a specific stimulus properties. There is evidence that even 2-year-old children can use predicate cues to infer a word meaning (Brown, 1957; Goodman, McDonough, & Brown, 1998), but this does not mean that children can use unpredictably changing contextual cues to infer meanings of several unfamiliar words heard within an episode or event. Yet, this ability is potentially important, for example, a visiting relative, new preschool setting, or new storybook might expose children to several novel words in a relatively brief time, and each word only can be understood with respect to its specific context of meaning. This compounds the problem of word learning and illustrates a more general problem: inferences must be adapted to local, changeable (linguistic and non-linguistic) contextual cues. This is the problem of flexible cognition: how can humans adapt to changing situations or problems by selecting and interpreting relevant contextual cues?

Because perseveration in the FIM varies between age groups and individual children, this task can be used to test factors predicted by the modal model to mediate inhibition. In brief, perseveration on an inferred meaning should increase with repetition of an initial predicate cue and response, and decrease with a delay and set-breaking activity between successive words for a given set. In the original FIM procedure children heard one word per set, all after the same predicate, within a block of trials (Deák, 2000a). For example, some children made inferences about six words predicated by "is made of," then six different words (for the same sets) following "has a," and finally six words following "looks like a." The modal model would predict, for example, that perseveration of an initial "made of" response set (e.g., persistent same-material choices) is less likely if a delay is imposed between "made of" and "has a" trial blocks—especially if the delay is filled with activity designed to break set by getting children to attend to the next relevant attribute (e.g., parts).

Study 1 examined whether an added between-task delay, with intervening activities to break the prior response set (i.e., habit), reduces perseveration. Study 2 tested whether blocking problems to increase the number of predicate (i.e., task) switches, or decrease the delay between responses to a given set, decreases or increases perseveration. Study 3 tested whether the number of successive problems of one type in the first (i.e., pre-switch) block influences perseveration in later blocks.

Children in Studies 3 and 4 also completed a Stroop-like test of verbal inhibition (Gerstadt et al., 1994), to assess the relation between susceptibility to verbal interference and perseveration in the FIM. Children in Study 3 completed a receptive vocabulary test, to examine the relation between flexible word learning and individual differences in lexical growth, and children in Study 4 also completed a test of their ability to judge whether a situation is determinate or indeterminate. The latter was intended to explore the alternative hypothesis that flexibility in the FIM depends on children's awareness that each novel word constitutes a new indeterminacy, rather than a repetition of an already-solved problem (Deák, 2003).

Study 1

Preschoolers might be most likely to perseverate on recent responses. The modal model predicts that interference from active responses should decay with time and with set-breaking activity. For example, adults have been found to perseverate less in interpreting ambiguous figures following a delay (Epstein & Rock, 1960), and preschoolers benefit from delays between dimension shift problems (Cameron, 1984). Set-breaking cognitive activity also influences adults' flexibility. In task-switching studies, interference from prior responses lasts for only 1–2 trials before the old response is inhibited (Allport, Styles, & Hsieh, 1994; Gilbert & Shallice, 2000). By the same token, interference from earlier inferences in the FIM might decay over time and/or over intervening trials. Notably, many children who perseverate in the FIM respond correctly to the first word given for an object set, then continue to make the same inference (i.e., choose the same attribute) in later blocks (Deák, 2000a). A between-block delay filled by activity that primes attention to the upcoming predicate, and the attribute it implies, might reduce 3-year-olds' perseveration.

To assess this possibility, in the FIM test, one group of children was trained in between blocks on the upcoming predicate. This imposed a delay of about 3 min (minimum = 90 s). Maljkovic and Nakayama (2000) reported that in adults, priming of visual features decays within 90 s, suggesting that children's perceptual attention to specific physical features should decay over a delay of this length. However, because there are few experimental data on the duration of semantic priming in young children, the adequacy of this delay is a matter of speculation. Consequently, this study only begins to establish a chronology of semantic priming in preschool children.

The addition of training during the delay yields a strong test of one prediction of the modal model of interference. During training the impending predicate was used to describe examples of the relevant attribute. For example, if the next predicate was "is made of...," pairs of same-material training objects were presented during the delay. These were described as, for instance, "made of porcelain," and relevant features were pointed out (e.g., "See, they're both hard and smooth and shiny!"). This was the pretest training procedure from Deák (2000a, Experiment 1), but distributed between blocks rather than massed before the test blocks. If perseveration stems from failure to inhibit prior responses or prior attention to the attribute implied by the first predicate, then delay and set-breaking activity should reduce perseveration.

Method

Participants

Fifty-four 3-year-olds (29 girls; mean age = 3.6, range 3.3–3.9) were recruited from a database at the University of Minnesota (Twin Cities). Most children were Caucasian and middle class. Eighteen 3-year-olds, recruited along with the 36 described in Deák (2000a, Experiment 1), were randomly assigned to a delay group.

Materials

Six sets of novel objects were used, plus six novel object pairs for the set-breaking activity (see Deák, 2000a; for detailed description). Each set included a standard, three comparison objects matching in body shape, material, or affixed part, and a dissimilar foil. Matching properties were unfamiliar. Set-breaking objects included two same-shape, two same-material, and two same-part pairs, with properties different from the test objects. Eighteen unfamiliar words were randomly assigned to test trials; six additional English words (e.g., *porcelain*) were used to label training pairs' common properties.

Procedure

Children generalized 18 novel words (one per trial) from a standard to a comparison objects. Each block of six trials included one unique novel word per set. Each block featured a different predicate: "looks like a," "is made of," or "has a." Predicate order was counterbalanced across children. Sets were presented in a random order that was repeated in each block, for a constant interval between successive trials with a given set. Words were randomized over trials. In each trial the child examined all objects in the set, and the experimenter then named property of the standard, saying either "This *looks like a(n)* [NOVEL WORD]," or "This *is made of* [WORD]," or "This *has a(n)* [WORD]." The experimenter then pointed to the comparison objects and asked "Which one of these also [*predicate*] [WORD]?" After the child responded, she or he was told, "Good job. Thank you."

Before each block children in the *delay* group completed two trials with two object pairs each matched on the attribute implied by the upcoming predicate. The experimenter named the attribute associated with the relevant predicate (e.g., "This one *is made of* porcelain...this one is made of porcelain, too"), and pointed out the perceivable properties of the attribute (e.g., "see, they're hard... and shiny"). The experimenters were to administer the two training trials in no less than 90 s, but the actual between-block interval was partly child-determined (e.g., by duration of object examination), so the mean delay varied across children. The delay in fact added a mean interval of almost 3 min between successive questions, compared to the control group. Children in the latter *no-delay* group (described in Deák, 2000a, Experiment 1) also received training trials, but completed all six before the first test block began.

Results

Children's inferences about novel word meanings were considered appropriate if they matched conventional inferences by older children and adults about the properties most strongly implied by the predicates. Predicate-appropriate responses are (1) same-shape objects for "looks like a(n)..." words, (2) same-material objects for "is made of..." words, and (3) same-part objects for "has a(n)..." words.[1] Flexibility is measured by the number of correct switches between successive inferences about words for a set, where the later response is appropriate for the new predicate.

Preliminary tests revealed that no gender difference approached significance; all further tests therefore combine girls' and boys' data. Mean number of correct switches was 3.8 ($SD = 2.7$) out of 12 in the delay group, versus 3.0 (2.9) in the no-delay group. Scores were entered into an analysis of variance, with group (delay vs. no delay) between subjects, and age in months and number of appropriate first block responses (a baseline index of predicate sensitivity) as covariates. The group effect was not significant, $F(1, 50) = 1.1$, nor was either covariate, age: $F(1, 50) = 1.3$; first responses: $F(1, 50) < 1$. Thus, younger 3-year-olds did not perseverate more than older 3-year-olds, and children who made more predicate-appropriate responses in the first block perseverated no more than children who made fewer. The latter finding is notable because the delay group made more appropriate first block responses (mean $= 4.6$ out of 6, $SD = 1.2$) than the non-delay group (3.2 $SD = 1.7$), $t(52) = 2.9$, $p < .01$. This makes the null group effect easier to interpret: there was sufficient power to detect a group difference, and sampling error, if any, would actually have favored the delay group. Because delay varied within each group (due to child-controlled trials), it was treated as continuous and entered in a partial correlation with number of appropriate response switches, and appropriate responses in the last two blocks (another index of flexibility). Age (months) and number correct in the first block were partialled out. The partial correlation between two FIM performance measures was $r = .92$, indicating high internal coherence. However, the measures' correlations with delay length were $r = .09$ and $r = .08$, *ns*, respectively. Finally, no group differences were found in any alternate dependent measures (e.g., total correct in later blocks).

[1] Although same-part responses would be syntactically "legal" inferences about words following "looks like a...," children do not choose this response (Deák, 2000a). Also, similar results were obtained with "is a" instead of "looks like a."

Table 1
Study 1: Summary of 3-year-old delay and control (Deák, 2000a, Experiment 1) groups

	Timing condition	
	Control	Delay
Delay (mean seconds)	411	589
Dependent variable (mean and SD)		
Correct first inferences	3.2 (1.7)	4.6 (1.2)
Correct switches	3.0 (2.9)	3.8 (2.7)
Response pattern (proportion)		
Flexible	.31	.39
Perseverative	.36	.39
Indiscriminate	.33	.22

Note. Delay group: $n = 18$; no-delay group: $n = 36$.

A final prediction of the inhibitory model is that children should perseverate less after a longer delay and intervening activity. Because inflexible 3-year-olds can be perseverative or indiscriminate, perhaps relatively more inflexible children in the delay group were perseverative. This would not necessarily show up as a group difference in number of correct switches. To test this, numbers of *perseverative*, *indiscriminate*, and *flexible* children were compared between groups. Numbers are shown in Table 1.[2] The distribution does not differ between groups, χ^2 $(df = 2, N = 54) = 0.8$. In sum, a 3-min delay with set-breaking activity did not reduce perseveration across inferences about novel words.

Discussion

The results do not suggest that perseveration in the FIM is reduced by a delay between successive responses, even when activity to prime the next predicate is imposed during the delay. Predicate-appropriate responses in the first block were higher among children in the delay group, perhaps due to the recency of same-predicate training trials. The null effect of a delay-plus-set-breaking is underscored by post hoc comparison of the delay group to 4-year-olds reported in Deák (2000a, Experiment 1). The latter performed significantly better not only than the no-delay 3-year-olds, but also better than the delay group, though *not* in the first block, $t(52) = 1.3$, *ns*. That is, 4-year-olds with no delay or set-breaking activity started out indistinguishable from the 3-year-old delay group, but thereafter made significantly more correct switches (means of 6.1 vs. 3.8), $F(1, 51) = 7.5$, $p < .01$. Thus, a delay-plus-set-breaking did not reduce a previously reported age difference in flexibility.

[2] A *flexible* pattern was defined as 10 or more (out of 18) predicate-appropriate responses (binomial $p < .04$ at $P = .33$). A *perseverative* pattern was defined as fewer than 10 appropriate responses and fewer than 4 total response switches (range = 0–12). An *indiscriminate* pattern was defined as 10 or fewer appropriate responses and at least four response switches.

Perhaps an added delay of about 3 min was insufficient to release interference from the prior set. Hypothetically, a very long delay (e.g., a year) would reduce perseveration, but this is trivial because flexibility implies selection from *active* responses. That is, prior responses must be recent enough to remain active as memory traces that might influence future responses. To confirm that such interference occurs in the FIM, a new group of 3- and 4-year-olds ($N = 12$) completed two blocks of the FIM test, but in the second block they were simply asked to recall their first responses. They correctly recalled a mean of 5.7 out of 6 responses (1.5 would be expected by chance). Thus, in the FIM test, children's first responses remain primed during later trials.

Because the incidence of perseveration in the FIM varies across children, the modal model would predict that factors that influence priming, like delay and intervening set-breaking activity, should affect perseveration. This prediction is disconfirmed by the current data. However, it is possible that failure to inhibit prior responses contributes to 3-year-olds' errors in the FIM, but this failure is moderated by temporal or sequential factors other than delay. Inhibitory processes might be more sensitive to repetition priming, operationalized either as the number of successive words with the same predicate (hypothetically related to response set activation strength), or number of trials between successive inferences about a set. These possibilities were tested in Study 2.

Study 2

FIM trials blocks in Deák (2000a) and Study 1 were blocked into same-predicate trials: children heard one word per set, all with the same predicate, then heard six new words (one per set) with another predicate, and so on. In this procedure some inflexible children are *attribute perseverative*: they focus on one attribute, usually the one implied in the first block (e.g., if "is made of" is first, they continue choosing same-material objects in later blocks). The modal inhibitory model could explain this as a learned response set for an abstract dimension that interferes with responses implied by later predicates. Other inflexible children are *item-perseverative*: they persistently choose the same object from a set, but across sets these objects do not match the standard objects on the same dimension. The modal model could explain this as idiosyncratic preferences for specific objects producing proactive interference in later trials.

These explanations are untested because the FIM has only been administered with predicate-blocked trials, thus confounding two factors that might affect children's inhibition of prior responses: (1) number of trials since the previous inference about a word for a given set; and (2) frequency of predicate changes, that is, the mean run length of same-predicate trials. If some children cannot inhibit a primed attribute type (e.g., material), and others cannot inhibit the priming of a specific object, then manipulating the aforementioned factors should affect the incidence and type (i.e., attribute-based or item-based) of perseveration.

We tested whether FIM perseveration is sensitive to changes in trial spacing and predicate change frequency by manipulating the trial blocking procedure. Two new

Table 2
Study 2: Summary of three different blocking conditions [predicate blocked (Deák, 2000a, Experiment 1), mixed predicate blocked, set blocked]

	Condition		
	Predicate blocked	Mixed predicate blocked	Set blocked
Delay and switches (mean)			
Between-trial delay	400.1	367.8	64.1
Number of switches	2.0	14.3	15.4
Dependent variable (mean and SD)			
Correct first inferences	3.6 (1.9)	3.3 (1.5)	3.3 (1.2)
Correct switches	4.6 (3.8)	4.2 (3.2)	5.0 (2.9)
Response pattern (proportion)			
Flexible	.46	.53	.53
Perseverative	.32	.25	.14
Indiscriminate	.22	.22	.33

Note. Mixed predicate and set blocked groups each $n = 36$; predicate blocked group $n = 72$.

groups were tested and compared to the 3- and 4-year-olds in Deák (2000a, Experiment 1). In a *Mixed predicate blocked* group, each block included one problem about each of six sets, but predicates changed randomly within blocks, so each block included two words with each of the three predicates ("looks like a," "is made of," and "has a"). Thus, number of trials between successive trials of a given set remained the same as in the *predicate blocked* groups (Deák, 2000a), but the number of predicate switches dramatically increased. In the *Set blocked* group, all three problems about a set were given in succession, thus minimizing the delay and intervening trials between problems about with the set, but maintaining the high rate of predicate switches. The differences between conditions are summarized in Table 2 (top panel). The modal model would predict more flexibility and less attribute perseveration in the Mixed predicate group, because there are fewer trials over which to build a highly activated response set to an attribute. By contrast, no change in item perseveration would be predicted because the delay between questions about a set is unchanged. The modal model would also predict more item perseveration in the Set blocked group, because the last response to a given set is quite recent and there is no intervening activity to reduce proactive interference. For the same reason, overall perseveration might increase in this group.

Method

Participants

Thirty-six 3-year-olds (18 girls, mean age = 3.6, range 3.1–3.11), and 36 4-year-olds (23 girls, mean age = 4.6, range 4.1–4.11) were recruited from a database at the University of Minnesota, and preschools in Nashville, TN. Most were Caucasian and middle class.

Materials

The test stimuli from Deák (2000a) were used.

Procedure

The procedure was modified from Deák (2000a, Experiment 1), where one predicate was used per block, for a total of two predicate switches. The *mixed predicate* group completed three blocks (six trials each), with two trials per block containing each predicate. Predicates were ordered quasi-randomly, with each given once per set, and no predicate given more than twice in a row. Set order was repeated in each block to match the delay between presentations of each set (see Table 2 for summary). The *set blocked* group completed all three trials with a given set in succession. To control test length, sets were removed and replaced between trials. This procedure, though awkward, did not seem to confuse children. Predicate and word order were randomized.

Results

Data from the mixed predicate and set blocked groups were compared to the original predicate blocks group (Deák, 2000a, Experiment 1).[3] Preliminary analyses showed no gender difference, so girls' and boys' data were combined. Number of predicate-appropriate inferences about the first word for a set, and number of correct switches, were analyzed in separate ANOVAs with group (predicate, mixed, or set) and age (3 or 4 years) between subjects (see Fig. 2 for group means and *SE*s). First inferences differed between age, $F(1, 138) = 9.8$, $p < .005$, but not blocking condition, $F(2, 138) < 1$. The interaction was not significant, $F(2, 138) < 1$. Correct switches, with first correct responses covaried, increased significantly with age, $F(1, 137) = 6.8$, $p = .01$. However, there was no group effect, $F(2, 137) < 1$, or group-by-age interaction, $F(2, 137) = 1.7$, $p > .17$.[4]

Numbers of flexible, perseverative, and indiscriminate responders (see Table 2) did not differ between groups; χ^2 $(df = 4,\ N = 144) = 4.7$, $p > .25$. There was a non-significant trend towards more indiscriminate responders in the set-blocked group. Also, attribute perseveration seemed to depend on blocking trials by predicate: 48% of perseverative children in the predicate-blocked group were attribute perseverative, versus only 14% in the mixed- and set-blocked groups. This trend, though not significant, χ^2 $(df = 2,\ n = 37) = 4.4$, $p = .11$, suggests that predicate repetition might weakly prime children's attention to an implied property dimension.

[3] Although between-experiment comparisons risk sampling error and unintended confounds, the FIM results reported by Deák (2000a) have been replicated, showing very similar results, across different training procedures, experimenters and sites. Although this does not completely address the problem of between-experiment sampling error, that problem is largely obviated by the results.

[4] An alternate index of flexibility, number of post-switch (i.e., first and second block) predicate-appropriate inferences, also showed a significant age effect but no blocking effect or blocking-by-age interaction.

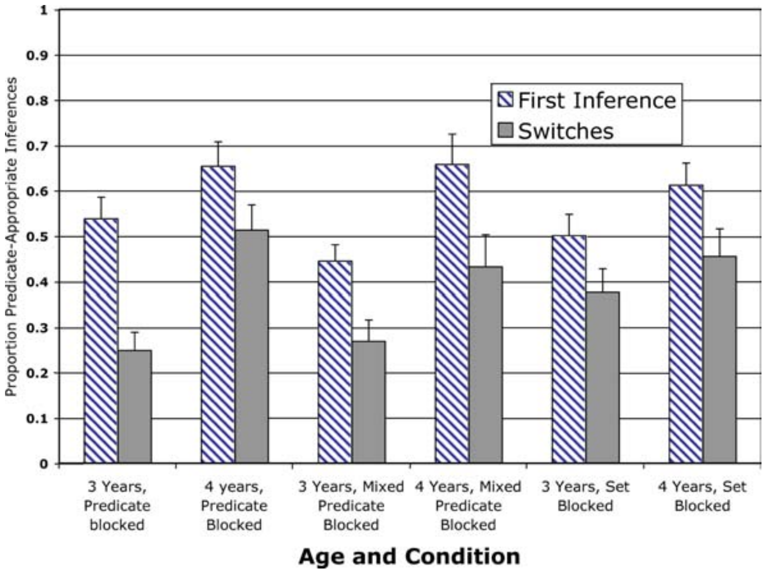

Fig. 2. Mean proportion of predicate-appropriate first responses to each set, and proportion of correct switches (error bars = *SE*), by age and condition in Study 2. The predicate blocked groups are 3- and 4-year-olds from Deák (2000a, Experiment 1).

Results from Studies 1–2

The manipulations of delay and trial order in Studies 1 and 2 allow a more powerful analysis of the effects of inter-trial delay on children's FIM test flexibility. For all 3- and 4-year-olds in Deák (2000a, Experiment 1) with useable videotape, and from Studies 1 and 2 ($N = 129$), mean delay between successive same-set trials was coded. Number of predicate-appropriate switches was submitted to a stepwise regression, with gender, age (in months), mean delay between successive questions about the same set, and number of appropriate first inferences (out of 6) entered as factors. The best model, with adjusted $R^2 = .18$, $F(2, 126) = 15.0$, $p < .001$, included age, $\beta = .35$, $t = 4.3$, $p < .001$, and number of appropriate first responses, $\beta = .21$, $t = 2.6$, $p = .01$. The two other factors—gender and mean delay—together accounted for about 1% of unique variance.

Discussion

The set-blocked group heard three successive facts with different words and predicates for a given set. If cognitive resources are needed to inhibit interfering responses, and interference accrues with response (or cue) repetition, a high frequency of cue changes should limit the activity strength of a response set. To test this prediction, a mixed predicate group was included also with a high rate of cue changes, but with fairly long delays between successive questions about a set. This is relevant because the modal model also predicts that interference decays over time and/or intervening activity. Thus, long delays should reduce perseveration. Contrary to these predictions, neither manipulation reduced perseveration nor increased predicate-appropriate response switches.

A limitation of the data is that more extreme changes in order and set length might influence children's flexible inductions of word meanings. Though we cannot impose larger manipulations of the relevant factors without adding trials, in Study 3 same- or mixed-predicate problems were added before the first block, to manipulate set length. Nonetheless, the findings are informative because they challenge an intuitive characterization of cognitive inhibition and perseveration. In some tests, blocking problems that require the same cognitive strategy causes perseveration in subsequent problem solving (Luchins, 1942). Also, a few findings suggest that problem solving by infants or preschool children is affected by the number, order, and timing of successive problems (Cameron, 1979; House, 1973; Marcovitch & Zelazo, 1999; Reutener & Fang, 1985; Toppino, Lee, Johnson, & Shishko, 1979). Finally, another possibility was that perseverative responses in the FIM were an artifact of blocking trails by predicates. In answer to these questions, there was no evidence that set length or cue switching rate affected perseveration in the FIM. In fact, few if any prior findings specifically show effects of timing or cue switch rate on flexible problem solving in children or adults. It is of course possible that a sufficiently repetitive task could "lull" adults and children into ignoring an unexpected cue change, thereby eliciting perseveration. Yet this would best be described not as inhibitory failure (i.e., inability to suppress the practiced response), but as an effect of inattention, inadequate activation, or distraction.

A clear implication is that the interval (from 1 to 8 min) between questions about an array does not influence children's flexibility in inferring different meanings. This has implications for children's linguistic skills, because inferences about word meanings *should* ignore a wide range of variables, including other recent inferences. Young children would misconstrue many utterances if they took into account contextual information like other recent novel words. Perseveration in the FIM signals a failure, in a challenging task, to decontextualize inferences about novel words (Olson, 1977). Apparently, however, even 3- and 4-year-olds are not influenced by potentially interfering contextual factors like the frequency or recency of a change in linguistic cues (e.g., predicates). An intriguing question for future research is whether children who perseverate do not know that word meanings are independent of such contextual variables. Another is whether flexible inferences about meaning depend on the same cognitive processes, and contextual factors, as flexible problem solving in other, non-lexical or non-verbal task domains.

The data also address the utility of the FIM test for studying children's word learning in rich contexts where they might hear many novel words. These findings (and Deák, 2000a) indicate age differences in flexible language processing that are not sensitive to training trials, interval between novel words, or the order in which various words are presented. The FIM test is therefore fairly robust.

In sum, preschool children's tendency to perseverate on an inferred meaning for several words for an array does not fit a model in which interference builds over repetitions of a cue or type of inference, or in which interference degrades with time or intervening activity. An alternative account (see General discussion) consistent with these findings is that children's understanding of a task, and the logical implications of between-problem differences, governs successive inferences or changes in their problem solving strategy. In Study 4 we tested this idea. First, though, we examine another hypothesis of the modal model: individual differences in FIM performance are related to a specific cognitive skill: cognitive inhibition. This is a focus of Studies 3 and 4.

Study 3

A few prior findings suggest that adding trials to a task increases the difficulty of shifting to a new response or strategy (Anderson, Choi, & Lorch, 1987). Marcovitch and Zelazo (1999), for example, found that infants' perseverative searching increases with the number of successive trials in which an object was found at the first location. Cameron (1979) found that older children establish a stronger learning set when given more trials. Toppino et al. (1979) found that multiple preliminary trials could compel children to respond to a non-preferred stimulus dimension. These findings suggest that difficulty of inhibiting a prior cue or response increases with cue or response repetition.

We tested this by having children complete the FIM in one of two randomly assigned groups: *High-Interference* or *Control*. Both completed six preliminary problems using object sets with familiar properties (shape, material, and parts); for instance, a square made of fur with a button attached (see Deák, 2000a, Experiment 3). The high-interference group's preliminary trials all had the same predicate as the first FIM test block. For example, if the first predicate was "is made of," children completed six problems of choosing comparison objects, for instance, "made of fur," "made of wood." They then completed the "is made of" test block, with novel materials and words, before switching to the next predicate. This yielded 12 consecutive same-predicate trials, of which the first six were unambiguous (because the property name was familiar; e.g., "wood") before a task switch. The control group completed six preliminary trials, including two with each predicate in the same order as the test blocks. These children therefore had three cue changes before the first block of test trials. The modal model would predict a stronger response set in the high-interference group, and consequently more perseveration after the first test block.

Individual differences in flexibility and perseveration also were investigated. The modal model casts inhibition as a centralized capacity that limits performance in a wide range of tasks (Logan & Cowan, 1984), including situations that require self-control, problem solving, and social cognition (Olson, 1989; Welch, Pennington, & Groisser, 1991). A related premise is that inhibitory capacity is a stable inter-individual trait. This is based on positive correlations between tasks that seem to require inhibition (Carlson & Moses, 2001; Reed et al., 1984). An unexamined question is whether purported tests of inhibition can predict children's flexibility in a complex inductive task like the FIM. To address this, we gave children a Stroop-like test of verbal response inhibition (Gerstadt et al., 1994), as well as the FIM. Because both tests involve lexical/semantic processes, children also completed the Peabody Picture Vocabulary Test (PPVT-R, Dunn & Dunn, 1981), which can be construed as an assessment of a child's history of successful lexical inference, and a fair index of verbal IQ. We expect a correlation between receptive vocabulary and the ability to make semantic inferences as required by the FIM, because the latter presumably contributes to the former.

Method

Participants

Twenty-four 3-year-olds (14 girls, mean age = 3.7 years; range 3.1–3.11) and 24 4-year-olds (12 girls, mean age = 4.6 years; range 4.1–4.11) were recruited from preschools in Nashville, TN. Most were Caucasian and middle-class. Five more children were replaced because they did not complete the tasks, and one because his PPVT-R score was more than 2 *SD* below the mean.

Materials

FIM. The test sets from Studies 1 and 2 were used. Six more object sets (Deák, 2000a, Experiment 3) were used in preliminary trials. Each, like the test sets, had a standard, same-shape, same-material, and same-part comparison objects, and a distractor. Critical properties were familiar: standard shapes were square, circle, triangle, egg, star, and heart; materials were glass, paper, wood, Play-Doh, sponge, and fur; and parts were bells, small teddy-bears, butterflies, dinosaurs, shells, and "googly eyes."

Stroop day/night. Two 13.5 × 10 cm pictures, mounted on cardboard, showed a crescent moon and stars on a black background ("day" card), and a yellow sun in a blue sky ("night" card).

PPVT-R. Each of a series of plates shows four line drawings.

Procedure

FIM interference test. Each age group was randomly assigned to a *high interference* or *control* group. Both completed six preliminary trials and 18 test trials. The high interference group did six easy (preliminary) and six hard (test) trials, all with a single predicate, before switching to another predicate in the second test block. The control group did two preliminary trials with each of the three predicates, then switched back to the first predicate for the first test block, then switched as usual for the second test block. Both groups switch again after the second test blocks.

Stroop-like day/night task. As in Gerstadt et al. (1994), children were trained to say "day" when they saw the moon picture and "night" when they saw the sun picture.

They were prompted if necessary, and praised for correct naming. Children then completed up to four practice trials, with feedback and reminders as needed, until at least two cards were correctly named. Children then completed 16 test trials (8 per card), in random order, with no feedback. Children were prompted, if necessary, to maintain attention. Number of rule-correct responses were recorded, and naming latency was later coded from videotape.

PPVT-R. Children were given words of increasing difficulty and asked to choose one of four referent pictures from a corresponding plate. The words become progressively less common. Mean age-normed receptive vocabulary is 100 ($SD = 15$). Tests were given (along with fourth, reported elsewhere) over two sessions on different days. Test order was counterbalanced.

Results

Preliminary analyses showed no sex differences, so all analyses combined girls and boys.

FIM

Mean predicate-appropriate responses (and *SD*s) in the preliminary trials and in the first test block, and correct switches in the last two test blocks, are shown in Table 3. Children were close to ceiling on preliminary trials (i.e., all group means $\geqslant 93\%$ correct), so the first predicate or property dimension should be strongly primed in the high interference group.

Predicate-appropriate responses in the first test block were compared in a 2 (age) $\times$ 2 (group) ANOVA. The age effect was not significant, $F(1,44) = 2.4$, $p < .13$ (with preliminary trial accuracy partialled out; however, age and first block accuracy were marginally correlated; $r = .27$, $p = .06$). The high interference group showed a trend towards more correct first block responses, $F(1,44) = 2.9$, $p = .10$. Age and condition did not interact, $F(1,45) < 1$. Number of correct switches in later blocks were also tested in a 2×2 ANOVA. The predicted age effect was obtained, $F(1,44) = 5.0$, $p < .04$, but the group factor was non-significant, $F(1,44) < 1$, as was its interaction with age, $F(1,44) < 1$.

Table 3
Study 3: Mean number (and SDs) of predicate-appropriate responses in the preliminary trials and the first test block of the FIM (maximum = 6), and the mean number of predicate-appropriate response switches after the first block (maximum = 12), by age and group

	Preliminary block	First test block	Correct switches
High interference			
3-year-olds	5.7 (0.6)	4.1 (2.3)	4.6 (3.8)
4-year-olds	5.9 (0.3)	5.3 (1.3)	5.6 (3.7)
Control			
3-year-olds	5.6 (0.8)	3.3 (1.7)	3.9 (2.6)
4-year-olds	5.9 (0.3)	4.4 (1.9)	5.9 (4.2)

Non-parametric analyses also showed null group differences. A Kolmogorov–Smirnov test showed no significant separation in the between-group distributions of appropriate switches ($Z = .3$, *ns*). Also, the distribution of flexible, perseverative, and indiscriminate response patterns did not differ between conditions, χ^2 $(df = 2,\ N = 48) < 1$. Thus, there is no hint of more perseveration in the high-interference group.

Stroop

Performance was evaluated in terms of total correct (i.e., rule-based) responses (out of 16), difference between accuracy in the first and last four trials (younger children's accuracy tends to decline), and response latency. Total correct averaged 9.1 ($SD = 6.3$) and 10.9 (5.2) in 3- and 4-year-olds, respectively, $t(46) = 1.1$, *ns*. Difference between first and last trials averaged 0.8 in both 3- and 4-year-olds (SDs = 1.1 and 1.5), $t(46) < 1$, *ns*. Overall latency averaged 2.7 s ($SD = 1.4$) and 2.3 s (0.9) in 3- and 4-year-olds; $t(38) = 1.1$, *ns*. These data are similar to those reported by Gerstadt et al. (1994). Stroop performance did not differ between the high-interference and control groups: mean correct, $t(46) < 1$; first-last difference, $t(38) < 1$; latency, $t(38) < 1$, all *ns*.

PPVT-R

Mean standardized PPVT scores were 103.7 ($SD = 10.3$) for 3-year-olds, and 103.5 (14.0) for 4-year-olds, suggesting that our cohorts were comparable and typical in terms of verbal intelligence. There was no vocabulary difference between the high-interference and control groups (means = 105.7 and 101.5, respectively), $F(1, 44) = 1.4$, *ns*.

Between-task relations

Partial correlations were computed among five FIM measures (total predicate-appropriate responses, number in the first block, number in last two blocks, difference between first and last blocks, total appropriate response switches), and eight Stroop measures (total correct, first four trials, last four trials, difference from first to last four trials, mean response time in all trials, mean RT in first four trials, mean RT in last four trials, difference between the log of mean response times in first and last four trials). Three factors were partialled out: child's age, receptive vocabulary (PPVT-R raw score), and number of correct preliminary FIM trials. Because the high-interference and control groups performed similarly in FIM and Stroop tasks, their data were combined to increase statistical power. Although many significant positive correlations were found among within-task measures, ranging as high as $r = .89$, there was *no* significant association between any of the 40 pairs of FIM and Stroop measures. The highest inter-task partial correlation was $r = .15$, *ns* (number of appropriate first-block FIM responses and mean Stroop RT). The mean correlation among all 40 inter-task pairs was $r = -.01$ ($SD = .14$). It is notable that by chance we would expect to find two significant and two marginal correlations out of 40.

For a finer-grained analysis, we compared Stroop performance of children who perseverated on the FIM ($n = 7$) with those who responded indiscriminately ($n = 8$).

These groups had similar error rates but different numbers of perseverative errors. If perseveration stems from inhibitory failure, we would expect worse Stroop performance in the former group. The difference in Stroop errors was non-significant, however, $t(13) = 0.7$, and in the opposite direction than predicted (means of 3.7 and 5.1). The groups also did not differ in proportion of reversal errors (means of 0.7 in both groups). Thus, FIM response pattern is unrelated to Stoop error rate and type.

The predicted relation between vocabulary and FIM performance was confirmed: the partial correlation between standardized PPVT scores and appropriate FIM responses was $r = .28$, $p < .06$, with age and correct FIM preliminary responses controlled. Also, children who responded more quickly on the Stroop had larger receptive vocabularies, $r = .36$, $p < .03$.

Discussion

Adding preliminary trials to extend the initial run of same-predicate problems did not increase perseveration after a predicate switch. Group differences were trivial and not in the direction predicted by the modal model. Most children responded correctly in preliminary trials, ostensibly building a response set, but this did not influence later inferences. Perhaps perseveration does not stem from an interfering response set that is proportional to the number of prior responses, or at least not in verbal inference tasks. Notably, the predicted age difference in flexibility was obtained, suggesting that statistical power was adequate to detect group differences.

Stroop test performance did not predict perseveration in inferences about word meanings. This is notable because both tasks elicit perseveration in preschoolers, and both require selection of word-referent relations. If some inhibitory resource contributed to individual and age differences in performance, we would expect the tests to be correlated. However, the 40 between-task correlations ranged narrowly around zero. Yet children's performance in the tasks was not random: there were many significant positive correlations among within-task measures. It seems the FIM and Stroop tests do not tap into a common trait that varies across children, such as a general inhibitory capacity.

This conclusion contradicts reports of moderate within-child stability across purported tests of response inhibition (Carlson & Moses, 2001; Reed et al., 1984). We therefore re-examined those reports. Both showed low intra-individual stability between tasks, and the highest inter-task correlations were mostly between tasks with similar incidental features. Reed et al.'s (1984) highest correlations (partial $r \approx 0.50$) were between, for example, a task that required delaying an action until a verbal signal was given (Pinball), and a simplified "Simon Says" task that also requires waiting for a specific verbal signal. Also, the authors did not control for verbal memory, verbal knowledge, speed of processing, or other factors, so the nature of the correlation is ambiguous. Carlson and Moses (2001) tested preschoolers on 10 purported tests of cognitive inhibition. The mean partial correlation, when age, gender, and verbal IQ were controlled, was only $r = .16$ ($SD = 0.12$). Thus, if there is a central inhibitory capacity, it accounts for less than 3% of unique shared variance among tests of inhibition. The authors did identify two clusters of tasks with different inhibitory demands: those tasks that impose conflict between competing responses, and those tasks that require delaying a desired action. The first cluster, which includes Stroop tasks, has a mean partial inter-task correlation of only $r = .22$ ($r^2 < .05$), so even these more narrowly defined tasks show little coherence. Of course, the FIM differs considerably from any of the tasks given by Carlson and Moses. For example, the Stroop requires remembering a simple, repetitive rule, whereas the FIM requires the use of probabilistic verbal cues (i.e., predicates) to make semantic inferences. Yet both would seem to require suppression of familiar or prior lexical associations. Also, both tasks are positively related to receptive vocabulary. Thus, variability in children's FIM flexibility is related to other abilities (e.g., word learning) but is unrelated to the capacity to inhibit primed lexical associations.

Study 4

The results of Study 3 suggested several further goals and questions. One concern is whether the independence of Stroop and FIM performance is replicable. Study 4 included both tasks, without manipulating any between-subjects factors. Study 3 also left open the question of what cognitive skills contribute to children's word learning flexibility. One possibility is that children who perseverate do not notice that each trial (or word) in the FIM constitutes a new, or independent, problem. There is evidence that young children misconstrue vague or difficult questions as easy or determinate (Cosgrove & Patterson, 1977; Revelle, Wellman, & Karabenick, 1985; Speer, 1984). In particular, children as old as 5 years fail to notice when there are several possible solutions to a problem or question (Fabricius, Sophian, & Wellman, 1987). It is as if they do not readily perceive the logical indeterminacy of an inductive problem. This might contribute to perseveration, because flexibility requires construing *each* inductive problem in a series of related problems—as a new indeterminacy. The solution to one problem does not "close bidding" on all related problems. If children expect successive problems to be independent (i.e., possibly have different answers), they should analyze whatever cues pertain to each problem, and adapt their responses accordingly. In the FIM, such children should respond flexibly to changing predicate cues, and infer a unique meaning for the word associated with each cue. By contrast, children who mistakenly construe inductive problems as determinate should tend to perseverate on their first solution to a series of problems about a stimulus array.

To test this possibility we administered a *Detection of Indeterminacy* test (modified from Deák, Ray, & Brenneman, 2003). Children saw several paired versions of scenarios that foreshadowed an outcome or answer. One version of each scenario was *determinate*: an unambiguous outcome was made obvious by information in the stimulus array. The other version was *indeterminate*: the result could not be inferred from available information. Children were not asked to infer the result; rather, they judged whether it had to be guessed (i.e., indeterminate) or could be known "for sure" (i.e., determinate). At issue is whether children who can judge a situation as determinate or indeterminate are more flexible across successive inferences about words for a stimulus array.

Method

Participants

Twenty-four 3-year-olds (13 girls, mean age = 3.8 years; range 3.1–3.11) and 24 4-year-olds (12 girls, mean age = 4.8 years; range 4.1–5.0) were recruited from preschools in San Diego, CA. Most were Caucasian and middle-class.

Materials

FIM stimuli and Stroop test cards were used. The Indeterminacy Detection test used two boxes of poker chips and two board game spinners with stickers of cartoon characters around the edges.

Procedure

FIM. The test was given as in Deák (2000a) but with no preliminary trials.

Stroop day/night. The test was administered as in Study 3.

Detection of indeterminacy. Children were shown three scenario pairs. One version of each implied a deterministic outcome; the other implied an indeterminate outcome. Children were asked, of each scenario version, whether the outcome was known with certainty (i.e., "Do you know for sure...") or indeterminate (i.e., "Do you have to guess?"). A correct response was to say that the indeterminate version outcome must be guessed, and the determinate version outcome is known.

In the *finger task* the experimenter held up several fingers. Children were asked whether they "know for sure" or "[have to] guess" how many fingers are up. In the *determinate version* the hand was held out on the table; in the *indeterminate version* it was held under the table, out of sight.

In the *poker chip* task, the experimenter showed children boxes filled with small plastic discs. She held each box so they could not see inside, and stirred up the chips while saying that she would draw a chip from the box. Children were asked whether they "know for sure" or "have to guess" what color would be picked from the box. In the *determinate version* all chips were the same color. In the *indeterminate version* the chips were divided between three different colors.

In the *spinner* task children saw two spinners. The experimenter spun each one and, while it was spinning, asked children if they "have to guess" or "know for sure" who the spinner would point to when it stops. The *determinate* spinner had stickers of Winnie the Pooh all around the edge; the *indeterminate* one had pictures of six different Winnie the Pooh characters.

Results

Preliminary tests of sex differences showed that girls made significantly more correct Stroop responses (mean = 14.0, $SD = 2.4$) than boys (10.5, $SD = 6.1$), $t(45) = 2.6, p < .02$. This fits with findings that sex differences in verbal ability, when obtained, favor girls. There were, however, no sex differences in correct FIM responses, $t(46) = 1.2$, *ns*, or indeterminacy detection, $t(46) < 1$, so data from boys and girls were combined.

FIM

Three-year-olds made fewer predicate-appropriate responses than 4-year-olds in the first FIM block (means = 3.5 and 4.5, *SD*s = 1.7 and 1.6); $t(46) = 2.0, p = .05$. To test age differences in flexibility, correct switches after the first block were entered in an ANOVA, with correct first-block responses covaried. Correct switches increased with age, $F(1, 45) = 7.8, p < .01$: means for 3- and 4-year-olds were 3.7 and 6.1 (*SD*s = 3.2 and 4.1). The distribution of response patterns also differed with age, χ^2 $(df = 1,\ N = 48) = 6.6, p < .04$: 8 3-year-olds and 15 4-year-olds were flexible; 9 3-year-olds and 2 4-year-olds perseverated; 7 of each were indiscriminate. This replicates previous results, though with a higher proportion of indiscriminate 4-year-olds.

Stroop day/night

Three- and 4-year-olds made means of 11.7 ($SD = 5.9$) and 12.9 (3.8) correct responses, a non-significant difference, $t(45) < 1$. Number of responses declined slightly from the first four to the last four trials, but this trend did not differ by age.

Detection of indeterminacy

Three-year-olds correctly responded to a mean of 0.6 ($SD = 0.7$) out of 3 situation pairs; 4-year-olds correctly responded to 1.1 (1.2). The age difference approached significance, $t(46) = 2.0$, $p < .06$.

Between-task relations

The partial correlation (controlling for age) was calculated among three FIM outcomes (total predicate-appropriate; first block appropriate; number of correct switches), and three Stroop outcomes (total correct, correct in first four trials, correct in last four trials). Though 5 of the 6 within-task measure pairs were significantly and positively related (mean $r = .66$; range $r = .24$ to .87), none of the between-task pairs approached significance (mean $r = -.01$; range $r = -.20$ to .22). (Note, however, the Stroop total correct distribution is skewed, so correlation coefficients should be interpreted cautiously.) There was no Stroop accuracy difference between flexible, indiscriminate, and perseverative FIM groups, $F(2, 46) = 1.3$, *ns*.

There is evidence, albeit equivocal, that flexible induction of word meanings is associated with ability to judge the determinacy of a situation's outcomes. Correct FIM responses in the last two blocks (i.e., after predicate switches) were correlated with indeterminacy detection scores at $r = .26$, with age and correct first block FIM responses partialled out. The relation is marginal ($p = .09$), but indeterminacy detection scores were not normally distributed, so correlations are hard to interpret. However, children who correctly judged the determinacy of both versions of 2 or 3 scenarios made more appropriate FIM switches (mean = 6.1, $SD = 4.1$) than children who correctly judged only 0 or 1 scenario pair (3.7, $SD = 3.2$), $t(46) = 2.1$, $p < .05$. Similarly, flexible FIM responders correctly evaluated the determinacy of more scenario pairs (1.2, $SD = 1.1$) than perseverative or indiscriminate responders (0.6 and 0.4; SDs $= 0.8$ and 0.7), $F(2, 47) = 3.4$, $p < .05$. The modal model would suggest the hypothesis that perseverative children are poorest at detecting indeterminacy, because to do so requires inhibiting imagined or inferred outcomes. However, indiscriminate responders did no better than preservative ones.

Finally, a linear regression was conducted on number of appropriate post-switch FIM responses (i.e., last two blocks), with age, predicate order, number of correctly judged scenario pairs (Indeterminacy Detection), and Stroop correct responses entered as factors. In the most parsimonious model, age ($\beta = .32$), and indeterminacy detection ($\beta = .26$), together yielded $R^2 = .23$, $F(2, 42) = 6.4$, $p < .005$. Other variables did not predict significant additional variance (complete model: $R^2 = .25$).

Discussion

The results confirm the findings of Study 3 and offer new evidence of a possible connection between logical reasoning and flexible problem solving. First, the absence of a relation between FIM and Stroop task performance was confirmed, though many significant within-task correlations were found. Gerstadt et al. (1994) claim that the Stroop task requires efficient inhibition of a prepotent verbal association, plus working memory of a naming rule. Both the Stroop test and the FIM elicit perseverative errors from some 3- to 5-year-old children. If this were due to a deficit in some inhibitory ability, as the modal model of inhibition predicts, a relation should have been found between the two tasks. This finding, together with our reanalysis of data from tests of cross-task stability of cognitive inhibition (Carlson & Moses, 2001), fails to support the premise of a stable "cognitive inhibitory capacity" that varies between preschool children. Such a trait might exist, but additional specification is required. For example, it might be volatile—highly sensitive to task context—or have a narrow range of variability that is not measurable by tests like the FIM or Stroop. It is possible that a minority of children have stable deficits of cognitive inhibition (Olson, 1989), but even this is ambiguous because other factors (e.g., verbal intelligence, processing speed, verbal comprehension, social knowledge, motivation, emotional factors) tend to be conflated in clinical subgroups of young children defined as impulsive (i.e., having inhibitory problems).

Though there is no compelling evidence of a central, stable inhibitory trait, there is some evidence that sensitivity to whether problems or situations are indeterminate or determinate is related to cognitive flexibility. Children who could judge the indeterminacy of ambiguous outcomes were better able to adapt inferences about word meaning to changing predicate cues.[5] A caveat is that when age was entered first in a step-wise

[5] One reviewer inquired whether indeterminacy detection could be re-described more simply as awareness that problems are independent. Deák (2000a, Experiment 3) found that in a non-inductive test with several independent questions about the same stimulus array (like the FIM), virtually all 3- and 4-year-olds responded flexibly. Moreover, many studies have shown that preschool children can select non-obvious information to solve inductive problems (e.g., Deák & Bauer, 1996). Thus, it is a sequence of logically independent problems, each of which requires a selective inference (e.g., a non-trivial forced-choice question), that seems to elicit preschoolers' perseveration.

regression, indeterminacy detection predicted only marginally significant unique variance in flexibility. Thus, the relation requires confirmation. Notably, Deák et al. (2003) found that preschoolers' ability to judge indeterminacy, plus receptive vocabulary, predicted performance in the Appearance-Reality test, which requires flexible selection of object labels. However, several possible interpretations of the relation between flexible cognition and indeterminacy detection are possible. For example, a general ability to interpret questions might explain the relation in question. Because vocabulary and indeterminacy detection contributed separate variance to appearance-reality performance in Deák et al. (2003), and because in the current study first-block FIM accuracy was partialled out of the marginal correlation between flexibility and ability to detect indeterminacy, there is some indication that general verbal comprehension cannot fully explain the relation. It seems that both linguistic and logical skills contribute to children's ability to flexibly use and make inferences about words and meanings. The exact relation, however, remains a question for future studies.

General discussion

The *modal model of inhibitory development* is the common but often implicit view that inhibition of prepotent responses—a common demand of complex tasks—demands an unspecified cognitive resource. This resource is limited in young children, who therefore make impulsive responses and perseverative errors when inhibition is required. The degree of interference posed by prior responses, or "response set," should vary with factors such as similarity of previous and current response context, frequency of the prepotent response, and time (and unrelated activity) since the prior response. Manipulating these factors should therefore influence children's tendency to perseverate.

The findings described here do not support the idea of a general inhibitory capacity that is time and repetition dependent, and that causes age and individual differences in perseveration. Flexibility in the FIM was unaffected by interval between successive problems about a stimulus (up to 9 min), number of successive same-cue problems (up to 12) before a task shift, and number of intervening trials (1 or 5) between successive questions about a set. Even 3 min of preparatory training between blocks for the upcoming predicate cue did not facilitate switching, suggesting that "release from interference" is not a necessary and sufficient condition of flexible post-switch responses. Nor did six "easy" problems before the first test block, all with the same predicate cue, cause subsequent perseveration, suggesting that cue repetition is a trivial factor.

The evidence also does not suggest a stable inhibitory capacity trait that varies between children (Carlson & Moses, 2001; Nigg, 2000; Olson, 1989; Reed et al., 1984). There was no hint of a relation between flexibility in children's word meaning inferences, and their ability to switch word-meaning associations (i.e., Stroop day/night), though both tasks should require semantic or lexical inhibition. It is not that the FIM is unrelated to other measures: it a marginally significantly related to vocabulary, with age and other factors controlled, and to ability to detect indeterminacy.

In sum, our findings replicate previously reported age and individual differences in 3- to 4-year-olds' FIM performance, demonstrating the test's robustness in the face of procedural and sampling variation. The data also begins to establish the external validity of the FIM with respect to other cognitive and linguistic measures. They fail, though, to support any prediction of the modal model of inhibitory development. Several concerns, however, might qualify this interpretation. Perhaps, for instance, we lacked statistical power to detect temporal and sequential effects (i.e., Type II error). Though we cannot conclusively refute this, two counterarguments seem relevant. First, the numerous significant and marginal effects found in Studies 1-4 indicate adequate power: besides age differences in every study, in Study 1 the delay group outperformed the control group in the first block; in Study 3 there were marginal or significant relations of vocabulary to FIM flexibility and Stroop accuracy; in Study 4 there was a significant difference in FIM flexibility between good and poor detectors of indeterminacy. Second, effects predicted by the modal model were never even marginally significant: all showed small differences, some in the opposite direction than predicted.

Another critique could be that we did not adequately manipulate the factors that influence inhibition (e.g., delay). This is difficult to refute, because factors like between-block delay certainly are amenable to extreme manipulation. However, the FIM elicits quite variable responses, including perseverative ones, from 3- and 4-year-olds, so even modest changes in *important* parameters should have some effect. If flexibility is affected only by extreme manipulations of delay or number of trials, it will demand a more specific model that yields more specific predictions than any current model of the development of inhibitory resources. In addition, it is worth noting that other studies have reported similar negative findings. For instance, Zelazo, Frye, and Rapus (1996) also found no effect of the number of pre-switch sorting trials using one rule, on subsequent switching to another rule. Also, in the FIM, preschoolers make no more perseverative errors on the first trials of a new block than on the last trials (Deák, 2000a), contrary to the idea that interference from a prepotent response "set" decays over time or intervening trials. Finally, Jacques, Zelazo, Kirkham, and Semecesen (1999) found that children who perseverated in a card-sorting task also judged a perseverating puppet to be correct, but judged a correctly response-shifting puppet to be incorrect. Thus, perseverative children are not simply unable to inhibit their own prior response; rather, it is their incorrect *construal* of the task that governs their responses.

Such results narrow the scope of plausible models of the relation between cognitive flexibility and inhibition, and underscore a central question: if perseveration is due not to inhibitory failure, then to what? Perhaps perseveration reflects an inability to adapt to cue changes when a task has overly complex cue/response contingencies (Zelazo & Frye, 1996). Or perhaps perseveration reflects a difficulty in reflecting on alternate perspectives (Perner, Stummer, Sprung, & Doherty, 2002). Neither of these explains why some children perseverate and others are indiscriminate. Perner et al.'s account might be adequate if recognizing indeterminacy requires explicitly realizing the uncertainty of one's perspective on some scenarios, and the FIM requires shifting perspectives between properties of a set of objects. Another possibility, which does not require an abstract construct like complexity or perspective, is that perseveration comes from not knowing what cues, or changes between questions, signal a new problem (i.e., a new indeterminacy). In the FIM, many variables change between trials, so children who do not notice predicate switches are likely to repeat familiar or confident responses. Because most children can understand the predicate cues, some respond correctly to the first predicate and then maintain attention to the relevant property, producing an attribute-perseverative pattern. Others focus on the repeated presentation of specific items, producing an item-perseverative pattern. Most generally, children who perseverate in the FIM blend responses across problems, rather than construing each question as logically independent from previous, superficially similar questions. If this is an apt description, awareness of the indeterminacy of successive problems should increase flexibility, and there is some evidence that this is so (Study 4).

Finally, the data have implications for children's language abilities. The FIM test focuses on a basic problem of comprehension: messages received from speakers have unpredictable meanings, implications and intentions. Fluent listeners must rapidly form and revise representations of meaning during discourse. This poses a special problem for preschool children because they hear many unfamiliar words. They must infer word meanings based on whatever linguistic, paralinguistic, and non-linguistic cues are associated with each word. One problem is that these cues change unpredictably. Another is that different cues have different scope: for instance, a marker like stress might influence a single utterance, whereas an inflection for gender can influence the form and interpretation of several subsequent utterances (e.g., by substituting gendered pronouns for a previously elaborate noun phrase). Typically, the referential meaning of a major sentence constituent (e.g., noun phrase) is most constrained by its predicate context. Herein lies a problem: how do children know the scope of a predicate cue? How do they know that meaningful cues from one sentence have limited relevance to the meanings of different unfamiliar words in later sentences—even if the same topic is re-established, for instance by pointing to the physical referent of the earlier sentence? This is essentially what occurs in the FIM: children must limit the scope of semantic predicate cues as the experimenter repeatedly says things about the same objects. Some children misjudge the influence of previous predicate cues or their own responses on later questions. The question is how various perceptual and cognitive processes contribute to such errors of semantic (or syntactic) inference (Deák, 2000b). Perhaps inhibitory difficulty contributes, at least in some circumstances. The current data show, though, that we lack an adequate description of inhibitory processes' contribution to developing ability to decontextualize semantic inferences. The data do suggest, in contrast, that some higher-order cognitive skills, like detecting indeterminacy, might contribute to the ability.

Acknowledgments

Thanks to all of the children who participated, their parents, and the preschool staff who facilitated recruitment and testing. Thanks to Laura Kilcoyne, Jennifer Wagner, and Loulee Yen for assistance with data collection. Thanks to Monica

Sweet and three anonymous reviewers for helpful suggestions and comments. This research was supported by a Vanderbilt University Research Council grant and National Science Foundation award BCS-0092027 to the first author. Some data were reported at the Psychonomics Society meeting, Dallas, TX, 1998.

References

Allport, A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umilta, & M. Moscovitch (Eds.), *Attention and performance XV* (pp. 421–452). Cambridge, MA: MIT Press.

Anderson, D. R., Choi, H. P., & Lorch, E. P. (1987). Attentional inertia reduces distractability during young children's TV viewing. *Child Development, 58*, 798–806.

Bjorkland, D. F., & Harnishfeger, K. K. (1990). The resources construct in cognitive development: Diverse sources of evidence and a theory of inefficient inhibition. *Developmental Review, 10*, 48–71.

Brown, R. (1957). Linguistic determinism and the parts of speech. *Journal of Abnormal and Social Psychology, 55*, 1–5.

Cameron, C. A. (1979). Trials per problem and age as factors in learning set formation of children. *Journal of Experimental Child Psychology, 27*, 401–422.

Cameron, C. A. (1984). Interference effects on preschool children's learning sets. *Journal of Experimental Child Psychology, 37*, 251–261.

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development, 72*, 1032–1053.

Cosgrove, J. M., & Patterson, C. J. (1977). Plans and the development of listener skills. *Developmental Psychology, 13*, 557–564.

Deák, G. O., & Bauer, P. J. (1995). The effects of task comprehension of preschoolers' and adults' categorization choices. *Journal of Experimental Child Psychology, 60*, 393–427.

Deák, G. O., & Bauer, P. J. (1996). The dynamics of preschoolers' categorization choices. *Child development, 67*, 740–767.

Deák, G. O. (2000a). The growth of flexible problem solving: Preschool children use changing verbal cues to infer multiple word meanings. *Journal of Cognition and Development, 1*, 157–192.

Deák, G. O. (2000b). Hunting the fox of word meaning: Why 'constraints' fail to capture it? *Developmental Review, 20*, 29–80.

Deák, G. O. (2003). The development of cognitive flexibility and language abilities. In R. Kail (Ed.), *Advances in Child Development and Behavior* (Vol. 31, pp. 271–327). San Diego, CA: Academic Press.

Deák, G. O., Ray, S. D., & Brenneman, K. (2003). Children's perseverative appearance-reality errors are related to emerging language skills. *Child Development, 74*, 944–964.

Dempster, F. N. (1992). The rise and fall of the inhibitory mechanism: Toward a unified theory of cognitive development and aging. *Developmental Review, 12*, 45–75.

Diamond, A., & Goldman-Rakic, P. S. (1989). Comparison of human infants and Rhesus monkeys on Piaget's AB task: Evidence for dependence on dorsolateral prefrontal cortex. *Experimental Brain Research, 74*, 24–40.

Duncker, K. (1945). On problem solving. *Psychological Monographs, 58* (5, Whole No. 270).

Dunn, L. M., & Dunn, L. M. (1981). *Peabody picture vocabulary test—Revised*. Circle Pines, MN: American Guidance Service.

Epstein, W., & Rock, I. (1960). Perceptual set as an artifact of recency. *American Journal of Psychology, 72*, 214–228.

Fabricius, W. V., Sophian, C., & Wellman, H. M. (1987). Young children's sensitivity to logical necessity in their inferential search behavior. *Child Development, 58*, 409–423.

Fauconnier, G. (1997). *Mappings in thought and langauge*. Cambridge: Cambridge University Press.

Gernsbacher, M. A., & Faust, M. (1995). Skilled suppression. In F. Dempster & C. Brainerd (Eds.), *Interference and inhibition in cognition* (pp. 295–327). San Diego, CA: Academic Press.

Gerstadt, C. L., Hong, Y., & Diamond, A. (1994). The relationship between cognition and action: Performance of children 3 1/2-7 years old on a Stroop-like day-night task. *Cognition, 53*, 129–153.

Gilbert, S. J., & Shallice, T. (2000). Task switching: A PDP model. *Cognition, 44*, 297–337.

Goodman, J. C., McDonough, L., & Brown, N. B. (1998). The role of semantic context and memory in the acquisition of novel nouns. *Child Development, 69*, 1330–1344.

Guttentag, R., & Ornstein, P. (1990). Attentional capacity and children's memory strategy use. In J. T. Enns (Ed.), *The development of attention: Research and theory* (pp. 210–305). Amsterdam: Elsevier.

Harnishfeger, K. K. (1995). The development of cognitive inhibition: Theories, definitions, and research evidence. In F. N. Dempster & C. J. Brainerd (Eds.), *Interference and inhibition in cognition* (pp. 175–204). San Diego, CA: Academic Press.

Houdé, O. (2000). Inhibition and cognitive development: Object, number, categorization, and reasoning. *Cognitive Development, 15*, 63–73.

House, B. J. (1973). Problem length and multiple discrimination learning in retarded children. *American Journal of Mental Deficiency, 78*, 255–261.

Jacques, S., Zelazo, P. D., Kirkham, N., & Semecesen, T. (1999). Rule selection versus rule execution in preschoolers: An error-detection approach. *Developmental Psychology, 35*, 770–780.

Johnson, C. J. (1994). Inhibitory mechanisms in selection among multiple object names. *Cognitive Development, 9*, 293–309.

Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review, 91*, 295–327.

Luchins, A. S. (1942). Mechanization in problem solving: The effect of Einstellung. *Psychological Monographs*, 54 (6, Whole No. 248).

Luria, A. R. (1959). The directive function of speech in development and dissolution. *Word.*

Maljkovic, V., & Nakayama, K. (2000). Priming of popout: III. A short-term implicit memory system beneficial for rapid target selection. *Visual Cognition, 7*, 571–595.

Marcovitch, S., & Zelazo, P. D. (1999). The A-not-B error: Results from a logistic meta-analysis. *Child Development, 70*, 1297–1313.

May, C. P., Kane, M. J., & Hasher, L. (1995). Determinants of negative priming. *Psychological Bulletin, 118*, 35–54.

Miller, P. H., Seier, W. L., Probert, J. S., & Aloise, P. A. (1991). Age difference in the capacity demands of a strategy among spontaneously strategic children. *Journal of Experimental Child Psychology, 52*, 149–165.

Neill, W. T., Valdes, L. A., & Terry, K. M. (1995). Selective attention and the inhibitory control of cognition. In F. N. Dempster & C. J. Brainerd (Eds.), *Interference and inhibition in cognition* (pp. 207–261). San Diego, CA: Academic Press.

Neumann, E., & Deschepper, B. G. (1992). An inhibition-based fan effect: Evidence for an active suppression mechanism in selective attention. *Canadian Journal of Psychology, 46*, 1–40.

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology, and a working inhibition taxonomy. *Psychological Bulletin, 126*, 220–246.

Olson, D. R. (1977). From utterance to text: The bias of language in speech and writing. *Harvard Educational Review, 47*, 257–281.

Olson, S. (1989). Assessment of impulsivity in preschoolers: Cross-measure convergence, longitudinal stability, and relevance to social competence. *Journal of Child Clinical Psychology, 18*, 176–183.

Perner, J., Stummer, S., Sprung, M., & Doherty, M. (2002). Theory of mind finds its Piagetian perspective: Why alternative naming comes with understanding belief? *Cognitive Development, 17*, 1451–1472.

Peterson, L. R., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology, 58*, 193–198.

Piaget, J. (1954). *The construction of reality in the child.* New York: Ballantine.

Reed, M. A., Pien, D. L., & Rothbart, M. K. (1984). Inhibitory self-control in preschool children. *Merrill-Palmer Quarterly, 30*, 131–147.

Reutener, D., & Fang, J. (1985). Encoding processes and release from proactive interference in short-term memory of preschool children. *Journal of Genetic Psychology, 112*, 343–348.

Revelle, G. L., Wellman, H. M., & Karabenick, J. D. (1985). Comprehension monitoring in preschool children. *Child Development, 56*, 654–663.

Ridderinkhof, K. R., van der Molen, M. W., Band, G., & Bashore, T. R. (1997). Sources of interference from irrelevant information: A developmental study. *Journal of Experimental Child Psychology, 65*, 315–341.

Sophian, C., & Wellman, H. (1983). Selective information use and perseveration in the search behavior of infants and young children. *Journal of Experimental Child Psychology, 35*, 369–390.

Speer, J. R. (1984). Two practical strategies young children use to interpret vague instructions. *Child Development, 55*, 1811–1819.

Toppino, T. C., Lee, N. D., Johnson, P. J., & Shishko, S. A. (1979). Effect of perceptual pretraining on children's concept performance with nonpreferred relevant dimensions: Evidence for the role of attentional strategies. *Developmental Psychology, 15*, 190–196.

Welch, M., Pennington, B., & Groisser, D. (1991). A normative-developmental study of executive function: A window on prefrontal function in children. *Developmental Neuropsychology, 7*, 131–149.

Wertlieb, D., & Rose, D. (1979). Maturation of maze behavior in preschool children. *Developmental Psychology, 15*, 478–479.

Zelazo, P. D., & Frye, D. (1996). Cognitive complexity and control: A theory of the development of deliberate reasoning and intentional control. In M. Stamenov (Ed.), *Language structure, discourse, and the access to consciousness*. Amsterdam: John Benjamins.

Zelazo, P. D., & Reznick, S. J. (1991). Age-related asynchrony of knowledge and action. *Child Development, 62*, 719–735.

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