Contributions of response set and semantic relatedness to cross-modal Stroop-like picture–word interference in children and adults

Julie B. Hanauer a,*, Patricia J. Brooks b,*

a Department of Social Sciences, Suffolk Community College, Selden, NY 11784, USA
b Department of Psychology, College of Staten Island and Graduate Center of the City University of New York, Staten Island, NY 10314, USA

Received 27 December 2003; revised 2 August 2004
Available online 1 October 2004

Abstract

Resistance to interference from irrelevant auditory stimuli undergoes development throughout childhood. To test whether semantic processes account for age-related changes in a Stroop-like picture–word interference effect, children (3- to 12-year-olds) and adults named pictures while listening to words varying in terms of semantic relatedness to the pictures and response set membership. In Experiments 1 and 2, with animal and clothing pictures, the interference effect observed in children, but not in adults, depended on the distractors’ status as members of the response set. In Experiment 3, with unrelated pictures, adults, but not children, showed greater interference for trials with distractors in the response set. These results indicate developmental changes in picture–word interference involving the establishment of a response set in working memory.

© 2004 Elsevier Inc. All rights reserved.

Keywords: Picture naming; Selective attention; Suppression; Resistance to auditory interference; Stroop effect

* Corresponding authors.
E-mail addresses: hanauerj@sunysuffolk.edu (J.B. Hanauer), brooks_pj@yahoo.com (P.J. Brooks).

0022-0965/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
Introduction

The child’s environment is filled with information that he or she must learn to process selectively. In a noisy world, the concepts of selective attention, resistance to interference, inhibition, and working memory have important implications for task performance, allowing one to respond judiciously to relevant information while ignoring irrelevant sources of information. The Stroop task (Stroop, 1935) is the most widely used measure of selective attention and inhibitory control (MacLeod, 1991). In this task, participants are instructed to identify the print color of a color word (e.g., say “green” when shown the word blue printed in green). This requires them to inhibit reading the color word, a difficult task given that reading tends to be automatized (MacLeod, 1991; MacLeod & MacDonald, 2000). Hence, there is a prepotent tendency to read the word itself rather than state its color.

A typical finding (e.g., Knopler, 1996) is that interference increases (as measured by response latencies) as the strength of the semantic relation between the to-be-ignored word and the correct response increases. This has led researchers (e.g., Luo, 1999) to argue that semantic competition, as opposed to response competition, is the main source of Stroop interference. However, in the original Stroop task, the four color words (blue, green, red, and yellow) also made up the response set of correct responses. Therefore, the color words may have interfered more with identification of print colors simply because they were primed as responses relative to other words. Klein (1964) tested this response set hypothesis, using color words within the response set (e.g., green) and color words outside of the response set (e.g., orange), and found much greater interference when the written word was a member of the response set than when it was not. This early study demonstrated that response set interference is an important component of the Stroop effect.

To study the development of selective attention and inhibitory control, the Stroop task has been adapted for use with children in various ways. This work has generated debate as to whether the interference effect shows developmental continuity versus change. Comalli, Wapner, and Werner (1962) were the first to use the Stroop task with literate school-aged children and found significant developmental differences in the magnitude of the interference effect, with the largest effect occurring with 7-year-olds and decreasing throughout childhood until 17–19 years of age. In contrast, in related picture–word interference studies where participants were instructed to name pictures on which printed words were superimposed, Rosinski, Golinkoff, and Kukish (1975) and Rosinski (1977) found equivalent semantic interference effects in elementary school children and adults, suggesting continuity in semantic processing. In contrast, Ehri (1976) and Schiller (1966) found that children who were not yet proficient readers failed to show a semantic interference effect at all.

For the purpose of exploring earlier developments in selective attention, a number of researchers created versions of the Stroop task that do not require reading and, therefore, can be administered to very young children. For example, Jerger, Martin, and Pirozzolo (1988) adapted procedures of Green and Barber (1981, 1983) for a study of 3- to 6-year-olds, who were given the incongruous task of pressing a “Mom-
my” button whenever a male voice was heard and pressing a “Daddy” button whenever a female voice was heard. As compared with both the congruent trials (e.g., hearing Mommy spoken in a female voice) and neutral trials (e.g., hearing ice cream spoken in either a female or male voice), Jerger et al. (1988) found slower response latencies for incongruent trials (e.g., hearing Daddy spoken in a female voice). In addition, the magnitude of the interference effect (i.e., the reaction time [RT] difference between conflict and neutral conditions) decreased significantly between 3 and 4 years of age.

The debate concerning the relative contributions of semantic relatedness and response set membership to Stroop interference resurfaced in the developmental literature surrounding the “day–night” Stroop task. In the original study, Gerstadt, Hong, and Diamond (1994) instructed 3½- to 7-year-olds to say “day” to a black card with a white moon and to say “night” to a white card with a yellow sun. In a control condition, children were told to say “day” and “night” to abstract geometric designs. Gerstadt and colleagues found a significant difference between the experimental and control conditions, with children’s ability to produce the correct nonmatching response increasing with age. Taking Gerstadt and colleagues’ study a step further, Diamond, Kirkham, and Amso (2002) used several variations of the day–night Stroop task with 4-year-olds. Instead of saying “day” to the black–moon card and “night” to the white–sun card, children were told to say “pig” to the moon and “dog” to the sun. Although this task still required children to maintain two rules in working memory, the pictures and responses were not semantically related; therefore, Diamond and colleagues proposed that the inhibitory demands of the task would be reduced. Children’s better performance in the dog–pig condition in comparison with the day–night condition led to the conclusion that young children have difficulty in inhibiting a word that is semantically related to the target word. That is, activating the word night automatically activated the semantically related word day, thereby generating conflict.

Simpson and Riggs (2003) proposed an alternate explanation for children’s differential performance of the day–night and dog–pig conditions of Diamond et al. (2002). In the day–night condition, children’s responses were in fact candidate names for the pictures of the sun and the moon, whereas in the dog–pig condition, they were not. Simpson and Riggs (2003) hypothesized that response set interference, as opposed to semantic conflict, might be the source of children’s difficulty. To test this hypothesis, they removed the confound of semantic relatedness and response set by setting up four conditions: same response set, semantically related (see red, say “blue;” see blue, say “red”); same response set, semantically unrelated (see car, say “book;” see book, say “car”); different response set, semantically related (see black, say “yellow;” see white, say “green”); and different response set, semantically unrelated (see sun, say “pig;” see moon, say “dog”). Testing 4- to 11-year-olds, Simpson and Riggs found a significant effect of response set and no effect of semantic relatedness, demonstrating that response set interference accounted for children’s poorer performance on the day–night Stroop task. They suggested that children differentially attend to the stimuli within the response set rather than to specific semantic features of the words.
This same confound between semantic relatedness and response set membership appeared in the cross-modal Stroop task of Elliot, Cowan, and Valle-Inclan (1998). In this study, adults were asked to name patches of color as quickly as possible while ignoring related or unrelated auditory distractors. The related distractors were color names, all within the response set of the visual stimuli but never congruent with the visual stimulus presented on a given trial (e.g., the word red was never paired with a red color patch). The unrelated distractors were adjectives (e.g., dry, short, best). Across trials, participants named color patches that were presented without an auditory distractor (silent condition), with distractor onset preceding the color patch by 500 ms (−500-ms condition), or with distractor onset simultaneously with the appearance of the color patch (0-ms condition). Elliott and colleagues observed greater interference with color distractors versus noncolor distractors (e.g., seeing a blue color patch and hearing yellow versus hearing single) only when the distractor occurred simultaneously with the color patch. They suggested that the lack of a Stroop-like effect with the word presented prior to the color patch was due to adults having sufficient time to suppress their representation of the distractor. Therefore, they claimed that the source of the cross-modal Stroop effect resided in inhibitory mechanisms, although they could not rule out other theoretical notions, such as concurrent semantic processing, as the source of interference.

To test the suppression versus concurrent semantic processing accounts of the interference effect, Hanauer and Brooks (2003) adapted Elliot et al. (1998) cross-modal Stroop task for 4- to 11-year-olds. Because children process information more slowly than do adults (Kail, 1991, 1992), Hanauer and Brooks (2003) adjusted the timing of the −500-ms condition so that the color patch occurred 500 ms after the offset of the auditory distractor. This was done to allow even the youngest children sufficient time to fully activate the meaning of the distractor prior to seeing the color patch. Hanauer and Brooks argued that a Stroop-like interference effect in the −500-ms condition would be incompatible with the concurrent semantic processing hypothesis but would support the suppression account due to children having less efficient suppression mechanisms (Tipper, Bourque, Anderson, & Brehaut, 1989). When the distractor occurred simultaneously with the color patch, Hanauer and Brooks (2003) found a significant Stroop-like interference effect across child and adult groups that decreased markedly in size with age. This pattern of greater interference from auditory distractors in children than in adults is highly consistent with recent work demonstrating greater salience of the auditory modality versus the visual modality in children (Sloutsky & Napolitano, 2003). Even more important, Hanauer and Brooks (2003) observed a Stroop-like interference effect in 4- to 5-year-olds with the distractor presented 500 ms in advance of the color patch, supporting Elliot et al. (1998) suppression account of this effect.

Unfortunately, because the color–word distractors named colors in the response set (i.e., the color patches), Hanauer and Brooks (2003) could not determine whether the interference effect was due to the semantic relatedness of the distractors to the target colors or to their being members of the response set. In the current experiments, we sought to resolve this issue using a Stroop-like cross-modal picture–word
interference task to systematically examine semantic versus response set components of the interference effect.

To date, there have been just a few studies using the cross-modal picture–word interference paradigm with children (e.g., Brooks & MacWhinney, 2000; Jerger, Lai, & Marchman, 2002a, 2002b; Jerger, Martin, & Damian, 2002). Typically in the picture–word interference paradigm used with adults (cf. Cutting & Ferreira, 1999; Damian & Martin, 1999; Schriefers, Meyer, & Levelt, 1990), target pictures have been selected from a number of different semantic categories, in contrast to Stroop-like tasks where all target stimuli are semantically related (e.g., colors, numbers). In addition, stimulus asynchrony is usually blocked, in contrast to cross-modal Stroop experiments (Elliot et al., 1998; Hanauer & Brooks, 2003) where timing of the distractor varies within each block of trials. It is generally assumed (e.g., Cutting & Ferreira, 1999; Damian & Martin, 1999; Jerger et al., 2002; Schriefers et al., 1990) that semantic competition between the visual stimuli and auditory stimuli at the stage of lemma\(^1\) retrieval creates interference when participants are asked to name a picture semantically related to an auditory distractor. Jerger et al. (2002) found statistically equivalent effect sizes for semantic interference in children (i.e., 5- to 7-year-olds) and adolescents (i.e., 12- to 14-year-olds), suggesting continuity in linguistic knowledge and a similar time course of semantic processing during lexical access. Both the adults in Damian and Martin (1999) and the children and adolescents in Jerger et al. (2002) were slower at picture naming when target pictures and auditory distractors shared semantic features, with the interference effect peaking when the distractor onset was slightly (i.e., 150 ms) in advance of the picture (see also Schriefers et al., 1990). This pattern of stability in the semantic interference effect contrasts with Hanauer and Brooks (2003), who observed a marked developmental progression in the magnitude and time course in the cross-modal Stroop effect using semantically related distractors in the response set.

Caramazza and Costa (2000) addressed the issue of response set membership in adults with respect to semantic interference in a purely visual picture–word interference task in which printed words are superimposed over pictures to be named. They found that the magnitude of the semantic interference effect did not vary as a function of whether or not the distractor words were members of the response set. To our knowledge, the issue of response set membership has not been addressed in experiments using the cross-modal picture–word interference paradigm, nor has it been examined from a developmental perspective.

In sum, the existing literature is divided with respect to whether children show developmental change or continuity in performing Stroop-like selective attention tasks and whether distractor difficulty reflects semantic or response set competition. If there is continuity in the psycholinguistic processes involved in picture naming from early childhood to adulthood, as suggested by Jerger et al. (2002), the striking

\(^1\) In the Levelt (1989) speech production model, the term “lemma” refers to the nonphonological part of an item’s lexical information, that is, “those aspects of a word’s stored information that are relevant for the construction of the word’s syntactic environment” (p. 6).
developmental changes observed in the cross-modal Stroop effect (Hanauer & Brooks, 2003) must stem from another source besides semantic competition. The current experiments explored this issue further.

**Experiment 1**

Experiment 1 tested for the existence of a Stroop-like cross-modal picture–word interference effect in children and adults, using a variant of the cross-modal Stroop task of Elliot et al. (1998) and Hanauer and Brooks (2003). In contrast to these earlier studies of color naming that used distractors that were adjectives (i.e., color words such as purple for the related condition and other adjectives such as fewer for the unrelated condition), our participants named pictures of objects while listening to the names of other objects over headphones. The important difference in procedure here was the use of distractors, for both the related and unrelated conditions, that named objects that could easily be rendered as pictures. In each of two tasks, nouns of one semantically defined superordinate category (animals or clothing) comprised target pictures and related distractors, and the nouns of the remaining category served as the unrelated distractors. As in Elliot et al. (1998) and Hanauer and Brooks (2003), we varied stimulus asynchrony within each block of trials. Picture-naming latencies for trials with pictures paired with related words were compared with picture-naming latencies for trials with pictures paired with unrelated words to determine the occurrence of semantic interference.

**Method**

**Participants**

A total of 19 3- to 5-year-olds (7 boys and 12 girls, mean age = 4 years 11 months, range = 3 years 2 months to 5 years 11 months), 19 6- and 7-year-olds (7 boys and 12 girls, mean age = 7 years 2 months, range = 6 years 0 months to 8 years 0 months), 20 8- to 11-year-olds (9 boys and 11 girls, mean age = 10 years 5 months, range = 8 years 9 months to 11 years 11 months), and 30 adults (11 men and 19 women, mean age = 20 years, range = 18 to 39 years) took part in the study. All participants had corrected to normal vision, had normal hearing, and were native speakers of English. The sample was ethnically diverse and predominantly of middle-class backgrounds. The children were recruited using flyers posted at the College of Staten Island of the City University of New York. They were tested in a psychology laboratory and received stickers and T-shirts for their participation. The adults were undergraduates who were recruited from introductory psychology classes and received extra credit for their efforts. We tested a greater number of participants in the adult group than in the child groups due to the small effect sizes obtained in previous studies with adults (Elliot et al., 1998; Hanauer & Brooks, 2003).
Materials and apparatus

The experiment was conducted using a Macintosh Power PC computer equipped with PsyScope experimental software (Cohen, MacWhinney, Flatt, & Provost, 1993) and SoundEdit 16. The visual stimuli were selected from the Snodgrass and Vanderwart (1980) set of standardized black and white line drawings and consisted of six animal pictures and six clothing pictures, each measuring 3 × 3 in. and surrounded by a black background. We selected stimulus pictures that are easily recognized by children (Cycowicz, Friedman, Rothstein, & Snodgrass, 1997). In selecting the animal and clothing stimuli, we tried to avoid words with identical onset consonants or consonant clusters and words with identical rhymes. We also attempted to create animal and clothing sets that were roughly comparable in terms of their acoustic–phonetic onset characteristics.

The animal stimuli were bear, chicken, dog, fish, rabbit, and snake. The clothing items were dress, hat, mitten, pants, shoe, and sock. The auditory stimuli, consisting of the words corresponding to the 12 pictures, were presented over headphones in a naturally spoken female voice. These stimulus words have an early age of acquisition (Morrison, Chappelli, & Ellis, 1997). The duration of the animal words ranged from 581 to 709 ms with a mean of 657 ms, and the duration of the clothing words ranged from 582 to 756 ms with a mean of 658 ms. The PsyScope software controlled the presentation of the spoken words and pictures.

A microphone was set up in front of participants with a voice key triggering a button box timer that recorded response latencies in milliseconds for picture naming. RTs for correct responses were measured from the onset of the picture to the triggering of the voice key by the onset of participants’ responses.

Procedure

Participants were tested individually in a quiet room. Participants were seated in front of the computer monitor with a microphone adjusted to be approximately 1 in. away from their mouths. Participants performed two tasks that followed the procedures of Hanauer and Brooks (2003). Each task consisted of five blocks of trials. In one task, participants were instructed to name animal pictures, whereas in the other task, they were instructed to name clothing pictures (with order of tasks counterbalanced across participants). Across trials, each stimulus picture was paired with each stimulus word except its own (e.g., the picture of a mitten was never paired with the word mitten). That is, in each block of trials, items were randomly selected with replacement from lists combining each picture with each clothing word and each animal word (except for the matching word).

To prepare participants for the naming task, each picture was shown to the participants as the experimenter read each picture name aloud. The participants were then asked to repeat the names of the pictures to ensure recognition. On rare occasions, if a picture was not identified properly, the name of the picture was repeated until recognition was ensured. Participants then were told that they were going to
play a computer game and were asked to wear a set of headphones. Participants were
told that the object of the game was to “say what you see.” The experimenter told
participants.

In this game, pictures are going to appear on the screen, and your job is to name the pictures
as quickly as possible. Sometimes you are going to hear words over the headphones, but try
your best to ignore the word and say only what you see. The words may confuse you some-
times because they may be different from what you see, but you should concentrate and say
what you see.

Participants were instructed to name each picture as soon as it appeared and not
to talk otherwise. After 36 trials, a smiley face appeared to indicate that the block of
trials was finished. At the end of each block, participants were told that they did a
great job, and the children received stickers for their efforts.

Each trial began with a white fixation cross that remained on the screen for
2500 ms. In the 0-ms condition, the timing of the onset of the auditory distractor
was simultaneous with the presentation of the picture and coincided with the disap-
pearance of the fixation cross. In the −500-ms condition, the timing of the onset of
the auditory distractor was simultaneous with the disappearance of the fixation
cross, and the picture appeared 500 ms after the offset of the word. In the silent con-
dition, the picture replaced the fixation cross after 2500 ms and no auditory distrac-
tor occurred. The participant’s response activated a voice key that recorded the RT
for the trial and triggered the disappearance of the picture. After 200 ms, the fixation
cross for the next trial appeared. The experimenter kept a log of the participant’s re-
sponses as a record of errors and lost trials.

The animal-naming and clothing-naming tasks each consisted of five blocks of
trials. The first block consisted of nine trials of the silent condition. This allowed
participants to practice naming the pictures as they appeared on the computer
monitor and allowed the experimenter to adjust the sensitivity of the voice key. If participants made any errors during this block (i.e., misnaming a picture), the
errors were corrected. The second block consisted of nine practice trials with audi-
tory distractors to allow participants to get used to hearing the words in the head-
phones. If participants made any errors in this block (e.g., repeating the auditory
stimulus instead of naming the picture), the errors were corrected. After the two
blocks of practice trials, three blocks of 36 test trials (108 total trials) were pre-
sented, with each block consisting of 12 silent trials, 12 trials with distractors from
the same semantic category (six at each stimulus asynchrony [SA]), and 12 trials
with distractors from the other semantic category (six at each SA), occurring in
a random order. That is, in the animal-naming task, for half of the trials with
distractors, animal pictures were paired with animal words, and for the other half,
animal pictures were paired with clothing words. Likewise, in the clothing-naming
task, clothing pictures were paired with clothing words and with animal words in
equal numbers of trials. In the test blocks, participants were not corrected for any
errors. If, at any point during the experiment, a child became distracted or started
a conversation, the experimenter paused the experiment until the child calmed
down and trials could be resumed.
Results

Reaction times

Because the distributions of RTs were positively skewed, especially for the youngest children, we used median RTs for correct trials as the dependent variable. Table 1 presents RTs for each age group as a function of distractor type and stimulus asynchrony. These summary statistics are group means of individual medians. Overall, picture naming was fastest for the −500-ms condition, intermediate for the silent condition, and slowest for the 0-ms condition. This ordering of RTs as a function of distractor condition replicates Elliot et al. (1998) and Hanauer and Brooks (2003). It is likely due to the fact that the timing of onset of the picture, relative to both the fixation point and the distractor, varied randomly across trials within each block. Thus, because it was impossible for participants to predict for any trial exactly when the picture would appear, when the distractor preceded the picture (−500-ms SA), it alerted them to prepare a response. In the following statistical analyses, we examine picture-naming latencies for trials involving auditory distractors and do not consider the silent condition further.

Median picture-naming latencies were analyzed in a mixed-design analysis of variance (ANOVA) with distractor type (same semantic category or different semantic category), stimulus asynchrony (−500 ms or 0 ms), and picture type (animal or clothing) as within-subjects factors and with task order (animal naming first/clothing naming second or clothing naming first/animal naming second) and Age (3- to 5-year-olds, 6- and 7-year-olds, 8- to 11-year-olds, or adults) as between-subjects factors. The ANOVA showed main effects of age, $F(3, 80) = 69.94, \text{MSE} = 24,685,705, p < .001$, distractor type, $F(1, 80) = 16.52, \text{MSE} = 379,936, p < .001$, and stimulus asynchrony, $F(1, 80) = 206.71, \text{MSE} = 13,064,446, p < .001$. Overall, latencies for naming the pictures decreased markedly with age. Latencies tended to be slower for trials with distractors from the same semantic category than for trials with unrelated distractors and were slower when the distractor occurred simultaneous with the

<table>
<thead>
<tr>
<th>Age</th>
<th>Silent</th>
<th>Auditory stimulus condition</th>
<th>0 ms SA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-500 ms SA</td>
<td>Not in response set/ different category</td>
<td>In response set/ same category</td>
</tr>
<tr>
<td>3–5 years</td>
<td>1394 (59)</td>
<td>1188 (59)</td>
<td>1868 (84)</td>
</tr>
<tr>
<td>6–7 years</td>
<td>1033 (48)</td>
<td>854 (31)</td>
<td>1285 (59)</td>
</tr>
<tr>
<td>8–11 years</td>
<td>809 (22)</td>
<td>742 (19)</td>
<td>910 (29)</td>
</tr>
<tr>
<td>Adult</td>
<td>646 (11)</td>
<td>640 (11)</td>
<td>662 (15)</td>
</tr>
</tbody>
</table>

Note. Standard errors are in parentheses. $n = 19$ for 3- to 5-year-olds and for 6- and 7-year-olds, $n = 20$ for 8- to 11-year-olds, and $n = 30$ for adults.
picture (0-ms SA) rather than in advance of the picture (−500-ms SA). Although the main effects of task order and picture type were not significant, there was a significant two-way interaction of task order and picture type, $F(1,80) = 4.35$, $MSE = 280,409$, $p < .05$. When the animal task preceded the clothing task, latencies were considerably faster for animal trials (899 ms) than for clothing trials (967 ms). In contrast, latencies for animal and clothing trials were nearly identical when the clothing task preceded the animal task (1010 vs. 1006 ms for animal and clothing trials, respectively), suggesting a trade-off between faster naming for the first task, irrespective of semantic category, and faster naming of animals over clothing.

The main effects of age, distractor type, and stimulus asynchrony were qualified by significant two-way interactions of age and distractor type, $F(3,80) = 6.77$, $MSE = 155,661$, $p < .001$, age and stimulus asynchrony, $F(3,80) = 38.54$, $MSE = 2,435,685$, $p < .001$, and age and task order, $F(3,80) = 4.31$, $MSE = 1,522,470$, $p < .01$, a three-way interaction of age, distractor type, and picture type, $F(3,80) = 2.75$, $MSE = 33,407$, $p < .05$, a four-way interaction of age, distractor type, picture type, and task order, $F(3,80) = 3.51$, $MSE = 42,684$, $p < .05$, and a five-way interaction of age, distractor type, picture type, task order, and stimulus asynchrony, $F(3,80) = 3.88$, $MSE = 30,991$, $p < .05$. There were no other higher order interactions.

To decompose these interactions and, more important, to test for the existence of a semantic interference effect at each age, we conducted separate ANOVAs for each age group with distractor type, stimulus asynchrony, and picture type as within-subjects factors and with task order as a between-subjects factor. For the 3- to 5-year-olds, there were significant main effects of distractor type, $F(1,17) = 5.97$, $MSE = 512,871$, $p < .05$, stimulus asynchrony, $F(1,17) = 73.90$, $MSE = 11,915,642$, $p < .01$, and task order, $F(1,17) = 5.76$, $MSE = 5,182,260$, $p < .05$, with no interactions. In support of a semantic interference effect, latencies were longer for trials with same category distractors than for trials with unrelated distractors (1590 vs. 1473 ms). Picture naming was slower when the distractor appeared simultaneously with the picture than when it appeared before the picture (1813 vs. 1250 ms), and for children who performed the clothing task before they performed the animal task (overall mean = 1707 ms), as opposed to performing the tasks in the opposite order (overall mean = 1337 ms).

For the 6- and 7-year-olds, there were significant main effects of distractor type, $F(1,17) = 14.30$, $MSE = 234,584$, $p < .01$, and stimulus asynchrony, $F(1,17) = 57.05$, $MSE = 4,770,043$, $p < .001$, as well as an interaction of distractor type and stimulus asynchrony, $F(1,17) = 4.48$, $MSE = 32,117$, $p < .05$. Like the youngest group of children, the 6- and 7-year-olds showed longer latencies for trials with same category distractors (1094 vs. 1016 ms) and for trials with simultaneous presentation of the picture and distractor (1232 vs. 878 ms). Although the effect of distractor type was stronger when the picture and distractor occurred simultaneously, the semantic interference effect was nonetheless significant at both stimulus asynchrony conditions, $F(1,17) = 30.72$, $MSE = 32,117$, $p < .001$, for the 0-ms SA condition and $F(1,17) = 6.50$, $MSE = 32,117$, $p < .05$, for the −500-ms condition. In addition to these effects involving distractor type and stimulus asynchrony, the
6- and 7-year-olds showed an unexpected interaction of picture type and stimulus asynchrony, $F(1,17) = 8.33$, $MSE = 151,767$, $p < .05$. With simultaneous presentation of the target picture and distractor, animals were named faster than clothing (1176 vs. 1287 ms), $F(1,17) = 12.08$, $MSE = 151,767$, $p < .01$. However, with the distractor occurring before the picture, latencies for animal and clothing trials were statistically equal (885 and 871 ms, respectively). Given the unexpected nature of this interaction, we hesitate to draw any conclusion from it.

The ANOVA conducted for the 8- to 11-year-olds revealed a significant main effect only of stimulus asynchrony, $F(1,18) = 27.31$, $MSE = 1,122,418$, $p < .001$, with longer latencies for trials with simultaneous presentation of the picture and distractor (914 vs. 746 ms).

For the adults, there was a main effect of stimulus asynchrony, $F(1,28) = 8.40$, $MSE = 46,176$, $p < .01$, and an interaction of distractor type and stimulus asynchrony, $F(1,28) = 5.23$, $MSE = 4,463$, $p < .05$. Counter to predictions regarding semantic interference, RTs for the 0-ms SA condition were faster for trials with same category distractors than for trials with unrelated distractors, $F(1,28) = 7.45$, $MSE = 4,463$, $p < .05$. For the −500-ms SA condition, the effect of distractor type was not significant, $F(1,28) < 1$.

Taken together, these analyses showed that the effect of distractor type varied dramatically as a function of participant age. The youngest two groups of children showed the predicted semantic interference effect, with slower picture-naming latencies for trials with same category distractors than for trials with different category distractors. In contrast, the oldest children showed no effect of distractor type, and the adults showed the opposite effect, with faster latencies for trials with same category distractors. Although the semantic interference effect was observed at both stimulus asynchrony conditions in the youngest two groups of children, the opposite semantic “priming” effect in adults was present only when the distractor and picture occurred simultaneously. The analyses further indicated that the effect of stimulus asynchrony was highly significant in each age group but varied in its magnitude with age. The effects of task order and picture type tended to vary with age in idiosyncratic ways. Importantly, task order did not interact with the effect of distractor type at any age.

**Errors**

Trials were counted as errors when a participant misnamed the picture. If a participant stuttered (e.g., by saying “ah” prior to naming the picture), failed to make any response within 6 s, did not speak loudly enough for the voice key to activate, or caused the voice key to be triggered prematurely (e.g., by heavy breathing in the direction of the microphone, by kicking or hitting the desk), the trial was considered lost. Overall, both children and adults were highly accurate in performing the task. Given 216 total test trials (i.e., 108 in each task), the 3- to 5-year-olds averaged 7.3 errors and 13.4 lost trials, the 6- and 7-year-olds averaged 2.2 errors and 9.5 lost trials, the 8- to 11-year-olds averaged 1.8 errors and 5.4 lost trials, and the adults averaged 1.8 errors and 4.1 lost trials. One-way ANOVAs indicated that
numbers of errors, $F(3,84) = 7.58$, $MSE = 19.0$, $p < .001$, and lost trials, $F(3,84) = 7.33$, $MSE = 53.4$, $p < .001$, decreased significantly with age. To further examine error rates for trials involving auditory distractors, error proportions (corrected for numbers of lost trials) were subjected to a mixed-design ANOVA with distractor type, stimulus asynchrony, and picture type as within-subjects factors and with task order and age as between-subjects factors. This ANOVA yielded no significant effects.

**Discussion**

The youngest groups of children (3- to 5-year-olds and 6- and 7-year-olds) showed a cross-modal semantic interference effect in picture naming that resembled the cross-modal Stroop effect of Hanauer and Brooks (2003). Pictures paired with distractors from the same semantic category were named more slowly than were those paired with unrelated distractors. This interference effect was present at both stimulus asynchronic conditions: with simultaneous presentation of the target picture and distractor and with the picture occurring 500 ms after the offset of the distractor. Hanauer and Brooks argued that the basis of the interference effect was inefficient suppression of the distractor lemma, providing evidence that the effect decreased in magnitude with development and was observed at a wider range of stimulus asynchronies in children than in adults. The findings of Experiment 1 support this account.

As in Hanauer and Brooks (2003), Experiment 1 confounded response set membership with semantic relatedness of targets and distractors, thereby precluding precise determination of the contribution of response set to the effect. Experiment 2 combined the animal-naming and clothing-naming tasks of Experiment 1 into a single task such that, across trials, participants randomly viewed 12 different pictures (six animal pictures and six clothing pictures) paired with semantically related and unrelated distractors that are always members of the response set. If the interference effect observed in Experiment 1 was due to response set interference rather than semantic competition, this effect should disappear in Experiment 2.

In Experiment 1, the oldest children (8- to 11-year-olds) and adults failed to show a semantic interference effect, and the adults showed the opposite effect, with faster latencies for related trials. The priming effect in adults was observed only with simultaneous presentation of the picture and distractor. Delaying the picture by 500 ms provided sufficient time for adults, but not young children, to fully inhibit the distractor. This lack of a semantic interference effect in the oldest children and adults was surprising. We suspect that several factors may have contributed to this null effect. One such factor is the small set size used in the current experiment. La Heij and van den Hof (1995) showed, in the purely visual picture interference task, that the magnitude of the semantic interference effect in adults was greatly reduced when the response set size was decreased (e.g., from 16 to 4 pictures) and the stimulus pictures were repeated often. Previous cross-modal picture word interference experiments with adults (Cutting &
Ferreira, 1999; Damian & Martin, 1999; Schriefers et al., 1990) have used far more stimuli, with lists consisting of items selected from many different superordinate categories. A second factor may be the use of a categorized stimulus list (e.g., all pictures selected from the one or two superordinate categories). Kroll and Stewart (1994) observed that picture-naming latencies were longer for semantically categorized lists than for randomized lists of items. They suggested that categorized lists create intertrial interference at the conceptual level, leading to slower lemma selection of the target picture name. In the context of the current experiment, the use of categorized lists may have increased the amount of time required to access the names of the target pictures. If the time course of lexical access for the target picture name was delayed in our task, the timing of the distractor word relative to the picture might not have been ideal for generating an interference effect.

The fact that we observed a priming effect in the adults was even more puzzling. The literature on semantic relatedness effects does, however, cite examples of both semantic interference and priming in picture naming (Glaser, 1992). In the picture–word translation task, a “semantic relatedness paradox” was reported by Bloem and La Heij (2003), who observed that the occurrence of semantic interference versus facilitation in word translation depended on whether the distractor was a picture or a word. Adults were instructed to translate English words (e.g., spoon into Dutch lepel) while accompanied by a semantically related context picture (a fork) or printed word (a vork, which is the Dutch for fork). Word translation latencies were facilitated by a picture of a fork but were inhibited by the word vork. Bloem and La Heij explained this pattern through a lemma competition model in which activated concepts do not automatically launch their corresponding lemmas. They argued that a distractor picture creates facilitation because it generates activation only at the conceptual level and is never put into words. A visually presented distractor word, in contrast, creates interference because it continually accesses its corresponding lemma, thereby competing with the target lemma (e.g., lepel) during lexical access.

This experiment provided us with important insights regarding our results. The use of categorized lists would be expected to increase activation at the conceptual level. Related distractor words could enhance this activation in the same manner as the related picture primes in the Bloem and La Heij (2003) experiments. The use of auditory rather than visual distractors would lead to relatively short-term activation of the lemma corresponding to the distractor (especially if adults are quick to suppress an irrelevant auditory stimulus). For the auditory distractor to produce interference, activation of its corresponding lemma would have to be coincidental with access to the lemma of the target picture name. In cases where the timing of the distractor did not coincide with access to the lemma for the picture name, a related distractor could still add activation at the conceptual level that would lead to conceptual priming of the picture name.

If this post hoc explanation for the priming effect in adults is correct, we would expect this effect to be replicated in Experiment 2, which holds constant the stimulus materials and timing of the auditory distractors.
Experiment 2

To examine how the establishment of a response set affects the cross-modal picture-word interference effect, in Experiment 2, trials with animal pictures and those with clothing pictures were no longer segregated into separate tasks but rather were combined into a single task such that, across trials, the pictures and distractors varied randomly between animal and clothing categories, with equal numbers of trials from each category. As in Experiment 1, RTs for trials with related distractors (e.g., an animal picture paired with an animal distractor) were compared with RTs for trials with unrelated distractors (e.g., an animal picture paired with a clothing distractor) to determine the occurrence of interference and/or priming due to semantic relatedness.

Method

Participants

A total of 19 3- to 5-year-olds (7 boys and 12 girls, mean age = 5 years 0 months, range = 3 years 10 months to 5 years 11 months), 20 6- and 7-year-olds (8 boys and 12 girls, mean age = 7 years 2 months, range = 6 years 5 months to 7 years 11 months), 20 8- to 11-year-olds (10 boys and 10 girls, mean age = 10 years 2 months, range = 8 years 4 months to 11 years 11 months), and 30 adults (9 men and 21 women, mean age = 25 years, range = 18 to 47 years) took part in the study. The participants were recruited from the same sources as in Experiment 1. All participants had corrected to normal vision, had normal hearing, and were native speakers of English. None had participated in the previous experiment.

Materials and apparatus

The materials and apparatus were exactly the same as in Experiment 1.

Procedure

The task was essentially the same as in Experiment 1 in that participants were instructed to name pictures while listening to distractor words over headphones. However, in each block of trials, half of the pictures were clothing and half were animals (in contrast to Experiment 1, where participants named only clothing pictures in one task and named only animal pictures in the other task). As in Experiment 1, for animal-naming trials with auditory distractors, the distractor was an animal word for half of the trials and was a clothing word for the other half of the trials. Likewise, for clothing-naming trials, distractors were animal and clothing words equally often.
Results

Reaction times

Median RTs for correct trials were used as the dependent variable due to the positive skew of the RT distributions. Table 2 presents RTs for each age group as a function of distractor type and stimulus asynchrony. These summary statistics are group means of individual medians. Overall, picture naming was again fastest for the −500-ms condition, intermediate for the silent condition, and slowest for the 0-ms condition, replicating the results of Experiment 1.

Median picture-naming latencies for trials with distractors were analyzed in a mixed-design ANOVA with distractor type (same semantic category or different semantic category), stimulus asynchrony (−500 or 0 ms), and picture type (animal or clothing) as within-subjects factors and with age (3- to 5-year-olds, 6- and 7-year-olds, 8- to 11-year-olds, or adults) as a between-subjects factor. The ANOVA showed main effects of age, $F(3,85) = 36.58$, $MSE = 797,851$, $p < .001$, and stimulus asynchrony, $F(1,85) = 109.03$, $MSE = 190,006$, $p < .001$, as well as a two-way interaction of age and stimulus asynchrony, $F(3,85) = 17.94$, $MSE = 190,006$, $p < .001$. No other main effects or interactions were significant.

To decompose the two-way interaction of age and stimulus asynchrony, and to confirm the absence of a semantic interference effect at any age, we conducted separate ANOVAs for each age group with distractor type, stimulus asynchrony, and picture type as within-subjects factors. For each child group, only the main effect of stimulus asynchrony was significant, $F(1,18) = 119.71$, $MSE = 156,319$, $p < .001$, for the 3- to 5-year-olds, $F(1,19) = 10.32$, $MSE = 572,715$, $p < .01$, for the 6- and 7-year-olds, and $F(1,19) = 20.34$, $MSE = 123,844$, $p < .001$, for the 8- to 11-year-olds. In contrast, for the adults, the ANOVA revealed significant main effects of stimulus asynchrony, $F(1,29) = 43.95$, $MSE = 3,522$, $p < .001$, and picture type, $F(1,29) = 14.13$, $MSE = 1,671$, $p < .001$, with faster RTs for animal pictures than for clothing pictures (669 vs. 689 ms). In addition to these main effects, there

<table>
<thead>
<tr>
<th>Age</th>
<th>Silent</th>
<th>Auditory stimulus condition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>−500 ms SA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>In response set/</td>
</tr>
<tr>
<td></td>
<td></td>
<td>same category</td>
</tr>
<tr>
<td></td>
<td></td>
<td>In response set/</td>
</tr>
<tr>
<td></td>
<td></td>
<td>same category</td>
</tr>
<tr>
<td>3–5 years</td>
<td>1510 (72)</td>
<td>1163 (50)</td>
</tr>
<tr>
<td>6–7 years</td>
<td>1228 (65)</td>
<td>1042 (84)</td>
</tr>
<tr>
<td>8–11 years</td>
<td>834 (40)</td>
<td>700 (17)</td>
</tr>
<tr>
<td>Adult</td>
<td>675 (11)</td>
<td>659 (10)</td>
</tr>
</tbody>
</table>

Note. Standard errors are in parentheses. $n = 19$ for 3- to 5-year-olds, $n = 20$ for 6- and 7-year-olds and 8- to 11-year-olds, and $n = 30$ for adults.
was a significant two-way interaction of distractor type and stimulus asynchrony, $F(1,29) = 10.38$, $MSE = 1.256$, $p < .01$, and a three-way interaction of distractor type, picture type, and stimulus asynchrony, $F(1,29) = 9.40$, $MSE = 1.206$, $p < .01$. Replicating Experiment 1 and counter to predictions regarding a semantic interference effect, for trials in the 0-ms condition, RTs were faster for trials with same category distractors than for trials with unrelated distractors, $F(1,29) = 9.05$, $MSE = 1.256$, $p < .01$. Further analysis, however, revealed that this semantic priming effect was significant only for clothing naming, $F(1,29) = 25.00$, $MSE = 1.206$, $p < .001$ (with means of 697 and 742 ms, respectively, for clothing trials with related category distractors and unrelated distractors), and was not significant for animal naming, $F(1,29) < 1$. As in Experiment 1, there was no significant effect of distractor type in the $-500$-ms condition.

Given the lack of a semantic priming effect for animal-naming trials in Experiment 2, we reexamined the adult data from Experiment 1 to determine whether the same pattern obtained. For clothing-naming trials, with simultaneous presentation of the picture and distractor, the semantic priming effect was indeed significant, $F(1,28) = 12.71$, $MSE = 605$, $p < .01$ (with means of 666 and 687 ms, respectively, for clothing trials with related category and unrelated distractors) whereas for animal-naming trials, the priming effect was not significant, $F(1,28) = 1.61$, $p > .10$. Thus, the results of the two experiments are entirely consistent with each other with respect to the semantic priming effect in adults.

**Errors**

As in Experiment 1, children and adults were highly accurate in performing the naming task. Of the 108 total test trials, the 3- to 5-year-olds averaged 1.6 errors and 8.4 lost trials, the 6- and 7-year-olds averaged 0.8 errors and 6.6 lost trials, the 8- to 11-year-olds averaged 0.8 errors and 3.6 lost trials, and the adults averaged 0.5 errors and 2.6 lost trials. One-way ANOVAs indicated that numbers of lost trials decreased significantly as a function of age, $F(3,85) = 5.49$, $MSE = 29.1$, $p < .01$, whereas numbers of errors did not, $F(3,85) = 1.87$, $p > .10$. Because of the very low numbers of errors produced, no additional analyses were conducted.

**Discussion**

With animal and clothing trials combined into a single task, no semantic interference was observed in children or adults, suggesting that the interference effect observed in Experiment 1 was produced by the combination of semantic relatedness and response set membership. For all groups of children, RTs for related trials (e.g., an animal word paired with an animal picture) and unrelated trials (e.g., a clothing word paired with an animal picture) were statistically equal. Although the change in method greatly affected the magnitude of the semantic interference effect, it had very little impact on the semantic priming effect (i.e., faster RTs for trials with semantically related distractors than for trials with unrelated distractors) that was observed in the adult sample. As in Experiment 1, the priming effect was
observed only with simultaneous presentation of the target picture and distractor word, indicating that early presentation of the distractor provided adults with sufficient time to suppress it. In Experiment 2, unlike Experiment 1, there was an overall difference in adult performance in naming animals versus clothing, with slower latencies for the latter. Furthermore, the priming effect was reliable for clothing-naming trials but not for animal-naming trials, possibly because the clothing items tended to be more closely related, both semantically and associatively, than the animals (e.g., shoe–sock vs. fish–bear).

Although combining animal and clothing pictures into a single task was sufficient to eliminate the semantic interference effect observed in Experiment 1, it has not yet been determined whether semantic relatedness among pictures in the response set is required for children to construct a response set in the first place. In a third experiment, we examined whether response set membership, in and of itself, affects the magnitude of interference from auditory distractors in the case where target pictures were not members of the same semantic category. The selection of unrelated pictures in the response set eliminated the possibility of priming at the conceptual level. Hence, adults should no longer show faster picture-naming latencies for trials with distractors in the response set.

**Experiment 3**

Experiment 3 tested for the existence of a cross-modal picture–word interference effect due solely to response set membership. Two sets of unrelated pictures were constructed, with participants naming pictures from only one set. To test for a response set interference effect, RTs for trials with distractors from the participants’ response set were compared with RTs for trials with distractors that were not in the participants’ response set.

**Method**

**Participants**

A total of 20 4- to 7-year-olds (7 boys and 13 girls, mean age = 6 years 5 months, range = 4 years 9 months to 7 years 10 months), 18 9- to 12-year-olds (12 boys and 6 girls, mean age = 10 years 7 months, range = 9 years 0 months to 12 years 5 months),\(^2\) and 30 adults (13 men and 17 women, mean age = 22 years, range = 17 to 37 years) took part in the study. The participants were recruited from the same sources as in Experiment 1. All participants had corrected to normal vision, had normal hearing, and were native speakers of English. None had participated in the previous experiments.

---

\(^2\) Because of difficulties in recruiting young participants for this study, we tested only two age groups of children (4- to 7-year-olds and 9- to 12-year-olds) in Experiment 3 rather than three age groups as in the previous experiments.
Materials and apparatus

The apparatus was the same as in the previous experiments. The stimulus items were 12 black and white line drawings selected from the Snodgrass and Vanderwart (1980) set of standardized pictures. We selected stimuli that are easily recognized by children (Cycowicz et al., 1997; Morrison et al., 1997) and members of different superordinate categories, with the further criterion that the names of the 12 pictures have distinct initial phonemes and rhymes. Pictures were arranged in two sets with six pictures in each. Each group of pictures comprised a response set. Half of the participants in each age group were shown pictures from Set A, and the other half were shown pictures from Set B. Stimulus pictures for Set A were book, chair, flower, house, pig, and window. Stimulus pictures for Set B were apple, cup, doll, moon, scissors, and train.

Spoken words corresponding to the names of the 12 pictures were presented over headphones in a naturally spoken female voice. The auditory distractors were the names of the pictures from both Sets A and B. The auditory stimulus durations ranged from 767 to 1092 ms with a mean of 952 ms for words corresponding to pictures in Set A and ranged from 790 to 1069 ms with a mean of 957 ms for words corresponding to pictures in Set B. As in Experiments 1 and 2, in each block of trials, items were randomly selected with replacement from a list combining each stimulus picture with each word except for the matching word. Thus, pictures were paired with auditory distractors that were both in and out of the participants’ response set. For example, the word doll paired with a picture of a moon would comprise an “in response set” trial because both doll and moon were members of Set B. In contrast, the word flower paired with a picture of a moon would comprise a “not in response set” trial because flower was not a member of Set B. In each block of trials, pictures were paired with distractors corresponding to items in their response set and items not in their response set equally often.

Procedure

The task was the same as in the previous experiments.

Results

Reaction times

Table 3 presents RTs as a function of distractor type and stimulus asynchrony.3 These summary statistics are group means of individual medians. As in the previous

---

3 Unfortunately, not every pairing of target picture and distracter was fully unrelated. Therefore, we eliminated trials where the target and distracter were somewhat related. It should be noted, however, that the results for statistical analyses with these items included yielded essentially identical results. Because of the association between doll and house in the compound noun dollhouse, we eliminated trials where the picture of a doll was paired with the word house and where the picture of a house was paired with the word doll. Because of a part–whole relationship, we eliminated trials where the picture of a window was paired with the word train or the word house and where the picture of a train or house was paired with the word window.
experiments, picture naming was fastest for the −500-ms condition, intermediate for the silent condition, and slowest for the 0-ms condition.

Median picture-naming latencies for trials with distractors were analyzed using a mixed-design ANOVA with stimulus asynchrony (−500 or 0 ms) and distractor type (in response set or not in response set) as within-subjects factors and with task version (Set A as target pictures or Set B as target pictures) and age (4- to 7-year-olds, 9- to 12-year-olds, or adults) as between-subjects factors. The ANOVA showed significant main effects of age, $F(2, 62) = 33.68$, $MSE = 196,671$, $p < .001$, and stim-
ulus asynchrony, $F(1, 62) = 309.39$, $MSE = 17,279$, $p < .001$, qualified by a two-way interaction of age and stimulus asynchrony, $F(2, 62) = 55.37$, $MSE = 17,279$, $p < .001$. The only other effect to approach significance was the interaction of dis-
ttractor type and age, $F(2, 62) = 2.77$, $MSE = 6,723$, $p = .07$.

As in Experiment 2, we conducted separate ANOVAs for each age group to exam-
ine the effect of distractor type at each age and to break down the interaction of age and stimulus asynchrony. In these ANOVAs, stimulus asynchrony and distractor type were within-subjects factors and task version was a between-subjects factor. For both groups of children, only the main effect of stimulus asynchrony was signif-
ant, $F(1, 18) = 120.83$, $MSE = 43,565$, $p < .001$, for the 4- to 7-year-olds and $F(1, 16) = 77.74$, $MSE = 12,621$, $p < .001$, for the 9- to 12-year-olds. In contrast, for the adults, both the main effects of stimulus asynchrony, $F(1, 28) = 132.42$, $MSE = 3,042$, $p < .001$, and distractor type, $F(1, 28) = 8.46$, $MSE = 763$, $p < .01$, were significant and were qualified by a significant two-way interaction of stimulus asynchrony and distractor type, $F(1, 28) = 16.18$, $MSE = 609$, $p < .001$. Adults named pictures paired with distractors in their response set more slowly than they did pictures paired with distractors not in their response set. This response set inter-
ference effect was observed when the picture and distractor occurred at the same time, $F(1, 28) = 26.48$, $MSE = 609$, $p < .001$, but was not significant when the distrac-
tor preceded the picture by 500 ms, $F(1, 28) < 1$.

Errors

Of the 108 total test trials, the 4- to 7-year-olds averaged 1.7 errors and 4.3 lost trials, the 9- to 12-year-olds averaged 0.8 errors and 1.2 lost trials, and the adults
averaged 0.2 errors and 1.2 lost trials. One-way ANOVAs indicated that both errors, $F(2,65) = 8.92$, $MSE = 1.4$, $p < .001$, and lost trials, $F(2,65) = 5.39$, $MSE = 12.5$, $p < .01$, decreased significantly with age. Because of the small numbers of errors produced, no additional analyses were conducted.

**Discussion**

To examine whether response set membership by itself would affect the magnitude of the cross-modal picture–word interference effect, in Experiment 3 we constructed sets of pictures containing items that were members of different superordinate categories. The use of lists consisting of unrelated pictures eliminated the conceptual basis for the priming effect observed in adults in the previous experiments. In Experiment 3, adults, but not children, named pictures more slowly when paired with unrelated distractors from their response set, in contrast to unrelated distractors that were not in their response set. This suggests that adults, but not children, spontaneously constructed an ad hoc response set in working memory even when the pictures comprising the set were unrelated to each other. Activation of a response set allowed the adults to respond differentially to distractors as a function of whether they were candidate names for the pictures. The response set interference effect in adults was present only with simultaneous presentation of the picture and distractor, indicating that presenting the distractor 500 ms in advance of the picture enabled adults to fully suppress it.

**General discussion**

This study explored the relative contributions of semantic interference and response set competition to development in a Stroop-like cross-modal picture–word interference task. In Hanauer and Brooks (2003), dramatic differences were found between children and adults in their abilities to name color patches while listening to auditory distractors (color and noncolor words) presented over headphones. Hanauer and Brooks argued that the observed age-related changes in the magnitude and time course of the cross-modal Stroop effect were due to increases in the efficiency of suppression mechanisms, but they were unable to distinguish semantic and response set components of the interference effect. This was due to their use of color distractors that were semantically related to both target color patches and members of the response set, in contrast to noncolor distractors that were related to neither the target colors nor members of the response set. In the current set of experiments, by manipulating the response set membership of related and unrelated auditory distractors, we were able to delineate response set interference as a major component of the cross-modal interference effect.

In Experiment 1, we attempted to replicate the semantic interference effect in Hanauer and Brooks (2003). As in that study, response set membership and semantic relatedness of distractors were intentionally confounded. Both of the youngest groups of children (3- to 5-year-olds and 6- and 7-year-olds), but not the oldest chil-
dren (8- to 11-year-olds) and adults, showed a cross-modal semantic interference effect in picture naming in which pictures paired with distractors from the same semantic category (e.g., animal pictures paired with animal distractors) were named more slowly than were pictures paired with unrelated distractors (e.g., animal pictures paired with clothing distractors). This interference effect was observed both with simultaneous presentation of the distractor and picture and with the distractor offset occurring 500 ms in advance of the picture. The extended time course of the interference effect in the youngest children matches that of the cross-modal Stroop effect and confirms Hanauer and Brooks’s evidence of inefficiency in children’s suppression mechanisms. These findings are important in demonstrating age-related change in children’s ability to distinguish relevant and irrelevant information and in showing developmental improvement in the efficiency of response selection procedures.

In Experiment 2, the simple act of combining the pictures of animals and clothing into a single task completely eliminated the semantic interference effect. Thus, it appears that the Stroop-like effect observed in children in Experiment 1 and in Hanauer and Brooks (2003) depended on the combination of semantically related distractors that were members of the response set. It is difficult to relate this pattern of results with the findings of Caramazza and Costa (2000), who used a visual picture–word interference task and found that response set membership did not affect the magnitude of the semantic interference effect in adults. In contrast to the current study, Caramazza and Costa’s experiments tested only adults and used a large set of stimulus pictures sampled from many different superordinate categories, with few repetitions of either pictures or distractors. Further research is needed to directly examine the effect of response set membership across cross-modal and purely visual picture–word interference paradigms, holding these additional factors constant. An important consideration here is that a visually presented distractor will continually activate its corresponding lemma (i.e., with each visual fixation), whereas an auditory distractor will have a more transient effect in terms of activation of its lemma.

In both Experiments 1 and 2, and contrary to our predictions, adults not only failed to show a semantic interference effect but rather showed an opposite priming effect with faster latencies for pictures paired with related distractors. Consistent with the suppression account, the priming effect occurred only with simultaneous presentation of the picture and distractor. In addition, the priming effect was significant only for clothing-naming trials and not for animal-naming trials. Although our explanation of why semantic priming, rather than interference, occurred in adults is post hoc, we have speculated that the effect arose at the conceptual level from semantic or associative links among pictures in each category. Because of the constraints of selecting pictures that children could easily recognize, we did not attempt to equate items for their degree of association or semantic relatedness. In the mental lexicon, the animal domain, relative to the clothing domain, has many more members. For our experiments, we used animals that tended to be visually dissimilar and less closely associated with each other. We suspect that stronger conceptual links among the clothing items (hat, mitten, shoe, sock, dress, and pants), in comparison with those among the animals (dog, snake, fish, bear, rabbit, and chicken), may have differen-
tially primed latencies. This effect of category points to the need for balancing a num-
ber of semantic factors in creating stimulus materials. Further work using sets of ani-
mals that vary in terms of their degree of association, visual similarity, and
conceptual relatedness (e.g., mammals vs. animals in general) would allow us to
examine whether the priming effect is mitigated by these factors. In addition, the
use of a greater number of semantic categories and items would help in assessing
the generalizability of the current results.

To further distinguish the contributions of semantic relatedness and response set
interference in the cross-modal picture–word interference task, in Experiment 3 we
eliminated the possibility of conceptual priming by presenting participants with
unrelated pictures selected from different superordinate categories. This experiment
allowed an investigation of whether an interference effect due solely to response set
membership would occur. Interestingly, only adults showed a pure response set inter-
ference effect, naming pictures more slowly when paired with unrelated distractors
within their response set than when paired with unrelated distractors that were
not in their response set. This response set interference effect in adults occurred only
with simultaneous presentation of the picture and distractor, providing further sup-
port for the suppression account. The finding that the children’s latencies were not
affected by response set membership, in the absence of semantic coherence among
pictures in the response set, suggests that the working memory component involved
in categorizing and holding each picture in mind was overtaxed. The ability to con-
struct an ad hoc response set in working memory (Barsalou, 1983) may be a mani-
festation of the maturation of executive functioning.

Our observation of developmental change in Stroop-like cross-modal picture–
word interference contrasts greatly with the recent work of Damian and Martin
(1999) and Jerger et al. (2002) using the related cross-modal picture–word interfer-
ence paradigm. They argued that semantic conflict plays a critical role in generating
interference in the cross-modal picture–word interference task and noted that the
semantic components of picture naming are developmentally stable from early child-
hood to adulthood. Although they found significant semantic effects comparing tri-
als with related and unrelated distractors that were never members of the response
set, there may be alternative explanations of their results besides semantic competi-
tion between related words. In particular, Damian and Martin (1999) used unrelated
distractors (e.g., bulk, louse, fate, germ, rank, mist, threat, hunch) that seem to us to
be more abstract than the related distractors (e.g., cow, elbow, gong, zebra, fox, rad-
ish, whale, skunk) paired with the same target pictures. This difference in concreteness
may have contributed to faster latencies for naming pictures paired with unrelated
words as compared with related words. More generally, in Cutting and Ferreira
(1999), Damian and Martin (1999), Jerger et al. (2002), and Schriefers et al.
(1990), different sets of words were used as related and unrelated distractors, with
the two sets of distractors potentially differing in many possible ways (e.g., concrete-
ness, length, frequency, lexical ambiguity). Furthermore, all of these studies blocked
stimulus asynchrony and used larger sets of stimuli, uncategorized lists, and few rep-
etitions of either target pictures or distractors, in comparison with our Stroop-like
experiments. Blocking stimulus asynchrony creates predictability in the timing of dis-
tractor presentation and can lead to strategic effects that differ from unblocked presentation of stimulus asynchrony conditions.

In addition, cross-modal Stroop-like tasks may show larger contributions of response set membership to the interference effect due to greater repetitions of a small number of stimulus items. The day–night Stroop task (Gerstadt et al., 1994; Simpson & Riggs, 2003) is similar to our task in that a very small response set (i.e., two) is used across many repeated trials. Simpson and Riggs (2003) observed that response set was a stronger predictor of interference effect size than was semantic relatedness in children’s performance of this closely related task. As shown by La Heij and van den Hof (1995), semantic interference decreases with smaller set sizes. Repetitions of stimulus items may also produce sequential dependencies in the data, and this may have affected our results. Unfortunately, because of the random selection of trials for each participant, it is not feasible to systematically track such dependencies in the current study. In future work, we will systematically manipulate spacing of stimulus repetitions so as to explore negative priming and related sequential phenomena. Additional experimental work is also needed to clarify the contributions of distractor status, response set size, use of categorized versus uncategorized lists, and item repetition to the cross-modal semantic interference effect.

The developmental changes we observed in picture–word interference can be explained parsimoniously as reflecting maturation of the systems underlying selective attention, working memory, resistance to interference, and inhibition. Research has pointed to the role of the anterior cingulate and the dorsolateral prefrontal cortex (PFC) in these cognitive functions (MacLeod & MacDonald, 2000), but the exact nature of their roles is not well understood. The PFC is thought to be the center of goal-directed activities, working memory, inhibition, and categorization (Miller & Asaad, 2002), playing a critical role in reducing perseveration, selectively attending to stimuli, and supporting flexible behavior (Becker, Isaac, & Hynd, 1987; Miller & Cohen, 2001; O’Reilly, Braver, & Cohen, 1999; Stuss & Benson, 1984). All of these operations are implicated in the cross-modal picture–word interference effect. The PFC sends and receives projections from nearly all cortical areas within the brain, suggesting that it is the seat of integration and coordination of neural operations. The PFC is the last area of the brain to mature, as indicated by late myelination and synaptic pruning (Anderson, 1998; Bell & Fox, 1992; Chugani, Mazziotta, & Phelps, 1993).

Miller and colleagues (e.g., Freedman, Riesenhuber, Poggio, & Miller, 2001; Miller & Cohen, 2001) have further emphasized the involvement of the PFC in categorization, which they defined as the ability to “react similarly to stimuli when they are physically distinct and to react differently to stimuli that may be physically similar” (Freedman et al., 2001, p. 312). For example, one should react similarly toward an orange and a banana even though they are dissimilar in appearance, and one should act differently toward an orange and a ball even though they may be similar in appearance. Establishment of an ad hoc category of task-relevant items that are neither similar in appearance nor related by superordinate category membership underlies the response set interference effect in Experiment 3. In contrast to adults, children did not appear to spontaneously generate task-relevant categories in working memory in the absence of preexisting semantic coherence among category members.
In Experiment 1 and in Hanauer and Brooks (2003), the semantic interference effect in children was observed across a broader range of timing conditions than were either the semantic priming or response set interference effects observed in adults in Experiments 1 to 3. Our interpretation, that the time course of interference underscores age-related change in the efficiency of suppression, concurs with research using the negative priming paradigm to study the development of inhibitory mechanisms. Negative priming occurs when a participant must respond to a stimulus that was ignored on the preceding trial. Suppose, for example, that in a Stroop task, the word *red* is printed in blue on Trial 1 and the word *green* is printed in red on Trial 2. Identifying red as the print color on Trial 2 should take longer than if it had not been the ignored stimulus in the previous trial. A number of researchers (e.g., Neill, 1977; Neill & Westberry, 1987; Tipper et al., 1989) have argued that the magnitude of the negative priming effect is a direct reflection of the efficiency of inhibitory mechanisms and is negatively correlated with susceptibility to distraction. Using a Stroop-like task, Tipper et al. (1989) observed that 7- and 8-year-olds showed greater Stroop interference than did adults and that adults showed greater negative priming than did 7- and 8-year-olds.

The ability to resist interference plays a pivotal role in enabling one to scan, absorb, and attend to the environment. The Stroop-like cross-modal picture–word interference task has proven to be a powerful technique for uncovering developmental changes in this ability. It has become apparent that children and adults differ in their abilities to selectively attend to target information while resisting auditory interference (e.g., Elliott, 2002). This developmental progression may be mediated by age-related changes in the salience of the auditory modality (Sloutsky & Napolitano, 2003) or in the speed of response activation (Band, van der Molen, Overtoom, & Verbaton, 2000) or even by increased experience in tuning out background noises (e.g., television, conversion) while working. The current series of experiments is important in separating semantic and response set components of auditory interference in picture naming. The results are complemented by the recent work of Simpson and Riggs (2003), who found response set membership to overshadow semantic effects in the day–night Stroop task. Our work identifies response set membership as an important, but frequently overlooked, factor in explicating developmental change in resistance to interference from irrelevant auditory stimuli.

**Acknowledgments**

The authors thank the children who participated in the experiments and their parents. We thank Devin Bracco, Jose Colon, Maya McLawrence, Barbara Padowska, and Ariel Sionov for assistance with data collection, and we thank Kevin Sailor for comments on the manuscript. Portions of this work were presented at the second biennial meeting of the Cognitive Development Society, Virginia Beach, VA, October 2001, and at the biennial meeting of the Society for Research in Child Development, Tampa, FL, April 2003.
References


