Journal of Experimental Psychology: Human Perception and Performance
(in press)

Symbolic Distance Between Numerosity and Identity Modulates Stroop-Like Interference

Antonella Pavese
University of Oregon

Carlo Umiltà
University of Padua

Thirty undergraduates participated in an experiment investigating the effect of the arithmetic difference between stimulus identity and stimulus numerosity in a numerical version of the Stroop task. It was found that digits symbolically close to the enumeration response reliably produced larger interference than digits that were farther from the enumeration response. This semantic distance effect (SDE) was found with different numerosities (1÷9) and different enumeration processes (counting and subitizing), and it increased as a function of numerosity in the subitizing range. These findings suggest that digit identity autonomously activates a magnitude representation organized as a compressed number line.

Interference paradigms have been widely used to investigate which representations are automatically activated when nonrelevant or interfering visual objects are displayed. Among the most important interference paradigms is the Stroop task (Stroop, 1935; see MacLeod, 1991, for a comprehensive review). In Stroop's original work, participants named the ink color of incongruous color-word stimuli. Relative to control stimuli consisting of color patches, slower response times (RTs) were reported when the word (the nonrelevant dimension) was the name of a different color than that of the ink (the relevant dimension). These results showed that, when participants were required to name the color, the nonrelevant word was processed and interfered with the production of the correct response. Evidence that the semantic representation of irrelevant words is activated in the Stroop task comes from experiments that have shown a "semantic gradient" of interference. These studies showed that interference in the color-naming task increases as a function of the semantic association between the word and the concept of color (e.g., Klein, 1964).

Comparison judgment is another task in which the degree of symbolic similarity between stimuli affects performance. The symbolic distance effect (SDE; Moyer & Bayer, 1976) is a general phenomenon that has been found when individuals compare numbers or other symbols that can be associated with continuous object dimensions. Typically, the time required to compare two symbols on a given dimension varies inversely with the distance between their referents. When numbers are used, the latency of the comparative judgment is an inverse function of the arithmetic difference between the two numbers (Moyer & Landauer, 1967; Parkman, 1971). Although the SDE was originally found in comparison tasks, several studies have shown that this effect is a more general property of number representation and is also found in tasks in which no comparison is required (Brysbaert, 1995; Dehaene & Akhavein, 1995). It has been suggested that the SDE is the result of the activation of a magnitude representation associated with numerals and

We thank Steven Keele and Michael Posner for the use of the laboratories and for their helpful comments. We are indebted to Amy Hayes and Scott Doran for editing suggestions. We would also like to thank Stan Dehaene, Lana Trick, Raymond Klein, and an anonymous reviewer for their helpful comments and suggestions on earlier drafts of this article This research was supported by grants from the Italian Consiglio Nazionale delle Ricerche.

Correspondence concerning this article should be addressed to Antonella Pavese, Department of Psychology, College of Arts and Sciences, 1227 University of Oregon, Eugene, Oregon 97403-1227. Electronic mail may be sent to pavese@darkwing.uoregon.edu.

analogous to a compressed number line (Moyer & Landauer, 1967; Restle, 1970).

In the present study, we wanted to replicate the effect of symbolic distance between stimulus dimensions on Stroop interference. Previous studies (Hock & Petraseck, 1973; Pavese & Umiltà, 1997) have found that, when participants are asked to enumerate inconsistent digits in a display, the amount of Stroop-like interference depends on the arithmetic difference between the identity of the irrelevant digits and the number of items in the display (e.g., the enumeration response). Our main purpose was to investigate whether the modulation of the interference effect would be consistent with the activation of a magnitude representation organized as a compressed number line.

Stroop Interference and Semantic Association

Klein (1964, Experiment 1) measured Stroop interference in six conditions that differed in the semantic association between nonrelevant words and the relevant color dimension: (a) color words included in the response set (the standard incongruent condition); (b) color words not included in the response set; (c) common English words semantically associated with the colors; (d) common English words not associated with the colors; (e) rare English words; (f) nonsense syllables; and (g) groups of asterisks. The results revealed a continuum of interference from nonsense syllables, which produced a small but reliable delay in the color-naming latency, to the standard incongruent condition, which showed the strongest interference effect. Klein concluded that the amount of interference was influenced by the frequency of the word as well as its relatedness to the concept of color.

Several studies have confirmed this semantic gradient effect and have shown that interference increases as a function of the strength of the association between irrelevant words and the concept of color (Fox, Shor, & Steinman, 1971; Proctor, 1978; Redding & Gerjets, 1977; Scheibe, Shaver, & Carrier, 1967). Some studies have extended this result to other domains, such as numerosity and spatial direction (e.g., Fox et al., 1971).

In his review on the Stroop effect, MacLeod (1991) concluded that: "Compared with naming the ink color alone, irrelevant verbal stimuli unrelated to the concept of color interfere only minimally with color naming. However, as the word's semantic association to the concept of color increases, so does its potential to interfere." (p. 173). According to this view, interference depends on the strength of the association between the category of the nonrelevant dimension (colors) and the category of the response (color names). However, it is also possible that the degree of association between the particular values of relevant and nonrelevant dimensions in a certain trial might determine the amount of Stroop interference. This hypothesis is supported by recent evidence showing that perceptual similarity between the color of the ink and the color designated by the word affects the amount of Stroop interference. Klopfer (1996) has shown that words that denote a color that is highly similar to the color-naming response (e.g., the word GREEN in blue) yield more interference than words that denote a color that is dissimilar from the color-naming response (e.g., the word ORANGE in blue).

Numerical Variations of the Stroop Task

Windes (1968) first reported that performance in an enumeration task was slower when the stimuli being counted were incompatible Arabic numerals, and several studies have replicated these results (Shor, 1971; Flowers, Warner, and Polansky, 1979). Using a card-sorting task in which participants were required to order a group of cards according to the number of symbols printed on them, Morton (1969) found that number words and digits caused interference and noted that this interference effect was larger when the nonrelevant digits belonged to the response set.

Fox et al. (1971) found that an interference gradient similar to the semantic gradient found by Klein (1964) in the domain of color also occurred in the domain of numerosity. They reported increasing interference as the semantic association between the concept of number and the symbols to be counted increased. These symbols included circles, abstract shapes, letters, common words, Arabic numerals not included in the response set, and, within the response set, incongruent Roman numerals, Arabic numerals and number names.

Another numerical variation of the Stroop paradigm was developed by Francolini and Egeth (1980, Experiments 2 and 3), who instructed participants to enumerate red items in a circular display consisting of red and black items. Compared with a neutral condition in which letters had to be enumerated, Stroop-like interference was found when the red items were digits that were inconsistent with the enumeration response and a facilitation effect was found when the red items were compatible digits.

The SDE in Number Comparison

Moyer and Landauer (1967) first reported that the RT to judge which of two digits was the larger was an approximately inverse linear function of the numerical difference between the two stimulus digits[1] and named this effect the SDE. The SDE also occurs when participants are required to compare objects in memory on a certain dimension. For instance, Moyer (1973) found that in judging which of two animal names represented the larger animal, RT varied as an inverse linear function of the logarithm of the estimated difference in animal size.

Several researchers have investigated the SDE for numerical comparisons (Aiken & Williams, 1968; Banks, Fujii, & Kayra-Stuart, 1976; Dehaene, Dupoux, & Mehler, 1990; Duncan & McFarland, 1980; Folz, Poltrock, & Potts, 1984; Henik & Tzelgov, 1982; Hinrichs, Yurko, & Hu, 1981; Parkman, 1971; Sekuler & Mierkiewicz, 1977; Sekuler, Rubin, & Armstrong, 1971) and found a similar relationship between RT and numbers to be compared. The SDE appears to be continuous in two-digit numbers, with a significant influence of the units and with little or no discontinuity at decade boundaries (Dehaene et al., 1990; Hinrichs et al., 1981). The SDE does not disappear with extensive practice (Poltrock, 1989), is observed early in childhood (Duncan & McFarland, 1980; Sekuler & Mierkiewicz, 1977), and can be found in different linguistic communities (Dehaene et al., 1990).

The SDE has also been found with paradigms that used numbers but did not require any comparison, such as naming tasks. Marcel and Forrin (1974, Experiment 4) presented digits between 2 and 9 and asked participants to name them. They found a priming effect that varied as a function of the distance between the target digit and the prime digit. In a similar experiment, den Heyer and Briand (1986) asked participants to name single letters, asterisks, and digits. Naming a number was facilitated if the previous stimulus was another number rather than a letter or an asterisk. Furthermore, the amount of priming was larger for close digits and decreased with the distance between prime and target digits. More recently, Brysbaert (1995) found that when participants read a sequence of numbers, reading was facilitated if the previous number had a close value. Using a different experimental paradigm, in which participants had to judge whether a probe digit was included in a previously presented set of target digits, Morin, Derosa, and Stultz (1967) found that latencies for the "no" responses varied as a function of the distance between the probe digit and the target set: Close probes were rejected slower than far probes. Duncan and McFarland (1980) found that same-different judgments were also affected by the numerical distance between two numbers. The wide generality of this phenomenon led Dehaene (1992) to contend that the SDE is a universal characteristic of human numerical cognition.

The SDE and Stroop Tasks

In numerical variations of the Stroop task, participants respond to the number of items in the display and ignore their identity. In the semantic gradient version of this paradigm, two variables are usually manipulated: (a) the association between printed symbols and the concept of number and (b) the congruence between the enumera-

1. In particular, the equation proposed by Welford (1960) best describes the experimental results of Moyer and Landauer (1976) and Parkman (1971). Reaction times were defined as RT = $\underline{a} + \underline{b} \log [(\text{minimum}/\text{distance}) + 1]$, where distance is the difference in stimulus values and minimum is the smaller of the two stimulus values.

tion response and numerical symbols. The literature on the SDE suggests that, in an enumeration task, the strength of the association between a relevant dimension (i.e., the number of items in the display) and a nonrelevant dimension (i.e., the identity of the items) can be manipulated by varying the arithmetic distance between the correct response and the magnitude represented by the digits. The semantic gradient effect suggests that this manipulation should result in a change in the interference effect. Displays in which the items to be counted represent a quantity that is symbolically close to the enumeration response (e.g., a display with four 5s) should be enumerated slower than displays in which the items represent a quantity that is symbolically far from the enumeration response (e.g., a display with four 7s)

The presence of the SDE in a numerical Stroop task would demonstrate that the amount of interference is related not only to the strength of the association between the nonrelevant dimension and the relevant domain (e.g., between the nonrelevant word and the concept of color in a color-naming task) but also between the particular value of the nonrelevant and relevant dimensions in that trial (e.g., the color denoted by the word and the color of the ink; Klopfer, 1996).

Hock and Petrasek (1973, Experiment 3) first reported that the arithmetic distance between item identity and numerosity can affect enumeration latencies. In their experiment, participants were presented with lists of digit strings that had to be enumerated, ignoring their identity. When digit identity was close to the enumeration response (e.g., 33), response latencies were longer than when digit identity was far from the enumeration response (e.g., 55). Pavese and Umiltà (1997) used Francolini and Egeth's (1980) paradigm to verify the effect of symbolic distance on a numerical version of the Stroop task. Circular arrays of green and red items were presented for 200 ms and then masked. The task was to enumerate the red items and ignore the green ones. Red items could be letters (the neutral condition), digits consistent with the enumeration response, or digits inconsistent with the enumeration response. Inconsistent digits could be symbolically close (e.g., a display with four 3s) or symbolically far (e.g., a display with four 1s) from the enumeration response. The results showed that inconsistent close digits always yielded greater interference than inconsistent far digits regardless of the number of items to be counted (four or five). Furthermore, digit identities larger and smaller than the enumeration responses yielded a similar amount of interference.

In the current study, this effect was further explored to verify the hypothesis that the SDE is related to the activation of the magnitude representation of irrelevant Arabic numerals.

This experiment was designed to investigate two important characteristics of the effect of symbolic distance on Stroop interference in enumeration tasks. First, we wanted to verify whether the effect of symbolic distance on interference was affected by the type of enumeration process (subitizing or counting) that participants used. Subitizing is the effortless, confident, fast, and accurate enumeration process for a small number of items (Kaufman, Lord, Reese, & Volkmann, 1949; Mandler & Shebo, 1982). The subitizing range is widely defined as 1÷4, although the literature reports different estimates, and relevant individual differences have been found (Atkinson, Campbell, & Francis, 1976; Mandler & Shebo, 1982; Trick & Pylyshyn, 1993; 1994). Counting is a process that can handle a larger number of items but is slow, effortful, and error prone. Both processes can be defined as "enumeration" (Trick & Pylyshyn, 1994). Subitizing and counting are characterized by typical patterns of latencies and error rate. In the subitizing range RT increases reliably but slowly (slope = 40-100 ms) as a function of numerosity, whereas in the counting range RT increases faster (slope = 250-350 ms; Trick & Pylyshyn, 1994). Errors are typically low for numerosities of 1-3 and increase for larger numerosities (Mandler & Shebo, 1982). In this experiment, we used the first nine digits and several combinations of enumeration responses and nonrelevant digits. The effect of symbolic distance on interference was tested for numerosities that belong to the subitizing range (1-5) and to the counting range (5-9).

Another purpose of this study was to investigate the characteristics of the representation underlying the SDE, measuring interference variations as a function of numerosity and item identity. It has been proposed that the magnitude of numbers is represented as a compressed number line, in which the symbolic distance between one number and the next decreases as a function of numerosity (Dehaene, 1992; Restle, 1970). Stroop interference is known to increase as a function of the degree of association between relevant and nonrelevant dimensions (Klein, 1964; MacLeod, 1991). If one assumes that symbolic distance is a measure of the strength of the association between two number representations, one should expect that the amount of interference caused by nonrelevant digits would not only be a function of the arithmetic distance between the nonrelevant digit identity and the enumeration response but also of the absolute value of the enumeration response (e.g., its position on the number line). For instance, according to the compressed number line hypothesis, because the distance between 1 and 2 is larger than the distance between 8 and 9, the digit 1 should produce less interference on the enumeration of a two-item display than the digit 8 on the enumeration of a nine-item display.

Method

Participants

Thirty undergraduate students at the University of Oregon participated in the experiment. All had normal or corrected-to-normal visual acuity. They were tested individually in two sessions of approximately 50 min each. Participants were randomly assigned to one of two experimental groups: subitizing and counting.

Design

Participants in the subitizing group were presented with numerosities between 1 and 5, whereas participants in the counting group were presented with numerosities between 5 and 9. Therefore, the subitizing group responded to numerosities that were included in the subitizing range, whereas the counting group responded to numerosities that were included in the counting range. Letters were used as items to be counted for the neutral condition. Within each experimental group all the possible combinations of number of items and identities were used, for a total of 30 cells (5 display numerosities $\times$ 6 nonrelevant item identities) in each group.

Apparatus and Materials

The experiment was carried out on a Macintosh IIci. Stimulus displays were generated and controlled by the software Psyscope (Cohen, MacWhinney, Flatt & Provost, 1993), and presented on an Apple color monitor. The display was a standard phosphorous display with a graphic resolution of 640×480. The computer recorded vocal enumeration RTs using a microphone connected to the computer through a response box. The accuracy of the recorded latency was $\pm$ 1 ms. The identity of the participant's vocal response was manually entered by the experimenter at the end of each trial. The screen intensity was adjusted to an easy reading level and was maintained at that level throughout the experiment. Stimuli appeared in red against a black background. Each element was located at one of 18 equally spaced locations on the circumference of an imaginary circle. At the viewing distance of 65 cm, the center-to-center distance between the two diametrically opposed stimulus elements subtended a visual angle of approximately 4.1°. The mean visual angle between the edges of two adjacent positions was approximately 0.5°. Each item subtended a visual angle of approximately 0.2° in height and 0.4° in width. The symbols were displayed using the Macintosh system font Times (type size = 14 points).

The items were randomly distributed on the circumference. In the subitizing condition, two items were never presented in adjacent positions. Stimuli were either randomly selected uppercase letters (A, C, G, H, K, L, M, P, R, U, V, Y, and Z)[2] or the digits 1, 2, 3, 4 and 5 for the subitizing group and 5, 6, 7, 8, and 9 for the counting group. A white fixation cross was presented in the center of the imaginary circle for the duration of the trial.

Procedure

The experiment took place in a sound-attenuated, dimly lit room. Participants viewed the stimuli binocularly at a distance of about 65 cm from the display. They were tested in two experimental sessions. In each session, participants performed seven blocks of 60 trials each. At the end of each block, visual feedback was presented that informed the participant of his or her average response latency and percentage of correct trials.

The procedure for each trial was the following: (a) A fixation cross was presented for 800 ms, (b) the stimulus display was presented until the vocal response was recorded, (c) using the computer keyboard, the experimenter entered the identity of the response, and (d) 2.5 s elapsed from the end of a trial to the onset of the next trial.

In each session, participants began with a practice block of 30 trials followed by the seven experimental blocks. Participants were allowed to rest as long as desired between trial blocks. They responded to 420 trials in each session, for a total of 840 trials. Therefore, each participant contributed with 28 trials to each of the 30 experimental conditions.

Results

Median RTs and error percentages were calculated for each condition for each participant after removing trials containing incorrect responses (see Table 1). Both RTs and error percentages were entered into analyses of variance (ANOVAs) carried out separately on each experimental group to assess the effect of numerosity on RT and accuracy. Additional ANOVAs were performed to investigate the effects of symbolic distance on Stroop interference.

Numerosity Analyses

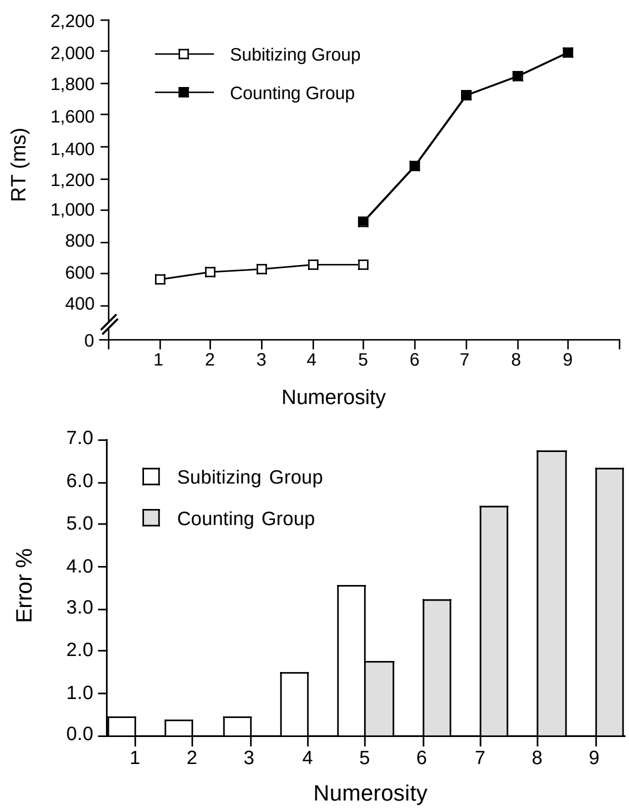
Separate analyses were carried out on median RTs and error rates for each experimental group (subitizing and counting). The average RTs and percentages of errors are plotted in Figure 1.

RT data. In the subitizing group, the main effect of numerosity was significant, $F(4, 56) = 33.23$, $MSe = 4{,}666$, $p < .0001$. Planned comparisons showed that all the differences were significant ($ps < .05$), with the exception of four- and five-item displays, which did not significantly differ from each other. Linear, $F(1, 56) = 118.8$, $p < .0001$, and quadratic, $F(1, 56) = 11.6$, $p < .005$, contrasts were significant. The increment of RT as a function of numerosity averaged 25 ms.

In the counting group, numerosity was also significant, $F(4, 56) = 132.3$, $MSe = 121{,}648$, $p < .0001$. Planned comparisons showed that RTs for each numerosity significantly differed from all the other numerosities ($ps < .005$). Linear, $F(1, 56) = 495.0$, $p < .0001$, and quadratic, $F(1, 56) = 28.6$, $p < .001$, contrasts were significant. The average slope of RT as a function of numerosity was 256 ms.

Error data. The same analyses were carried out on error percentages. In the subitizing group, the main effect of numerosity was significant, $F(4, 56) = 23.63$, $MSe = 7.085$, $p < .0001$. Planned

Figure 1. Mean reaction times (RT; in milliseconds) and error percentage as a function of numerosity.



2. We excluded the letters that are visually similar to numbers (e.g., I or O) and the letters that are initials of the number names included in the response set (e.g., O, T, F, S, E, and N).

Table 1
Mean latencies (in ms) and error rate (in percentage) as a function of Numerosity and Digit Identity.

Subitizing Group

	Digit Identity					
Numerosity	1	2	3	4	5	Letter
One	537 (0.0)	568 (1.3)	560 (0.7)	563 (0.2)	567 (0.5)	566 (0.0)
Two	615 (0.2)	588 (0.7)	620 (0.0)	616 (0.7)	611 (0.2)	606 (0.2)
Three	624 (0.2)	630 (0.5)	607 (0.0)	645 (1.2)	627 (0.2)	619 (0.5)
Four	658 (1.2)	656 (1.2)	667 (1.7)	624 (0.7)	680 (2.7)	651 (1.4)
Five	656 (5.5)	653 (4.6)	658 (2.4)	674 (5.0)	631 (1.7)	653 (2.1)

Counting Group

	Digit Identity					
Numerosity	5	6	7	8	9	Letter
Five	900 (0.5)	959 (2.7)	928 (1.8)	933 (2.2)	918 (2.0)	918 (1.1)
Six	1274 (3.9)	1272 (3.9)	1299 (2.9)	1246 (3.7)	1287 (4.7)	1289 (3.1)
Seven	1708 (7.8)	1676 (5.0)	1684 (6.1)	1737 (5.5)	1687 (5.7)	1710 (7.2)
Eight	1814 (6.5)	1798 (8.4)	1864 (7.3)	1745 (6.9)	1799 (9.5)	1857 (8.3)
Nine	1981 (4.7)	1930 (7.7)	1937 (9.7)	1934 (6.4)	1920 (5.1)	2011 (6.0)

comparisons showed that error rates for displays with one, two, and three items did not differ but that they were significantly lower than error rates for displays with four and five items ($p < .05$). Five-item displays were responded to less accurately than four-item displays ($p < .001$).

In the counting group, numerosity was also significant, $F(4, 56) = 10.434$, $MSe = 52.324$, $p < .0001$. Planned comparisons showed that displays with five and six items had a significantly lower error rate than displays with seven, eight, and nine items.

Symbolic Distance and Interference Analyses

To investigate the effect of symbolic distance and enumeration process on Stroop interference, we conducted two-way mixed ANOVAs on median RTs and error percentages. The between-subjects factor was group (subitizing vs. counting) and the within-subjects factor was condition (consistent, neutral, inconsistent close, and inconsistent far). For each participant, each level of condition was obtained by collapsing results from five different numerosities. The type of trials averaged for each condition is reported in Table 2. For the numerosities 3 and 7, for which inconsistent digits could either be greater or smaller than the enumeration response, the average of the two inconsistent close (±1) and inconsistent far conditions (±2) were averaged together[3].

RT data. The main effect of group was significant, $F(1, 28) = 280.2$, $MSe = 89719$, $p < .0001$. The subitizing group was faster than the counting group (618 and 1,534 ms, respectively). The main effect of condition was also significant, $F(3, 84) = 12.20$, $MSe = 891.2$, $p < .0001$. The mean RTs were 1,051, 1,088, 1,093, and 1,072 for the consistent, neutral, inconsistent close, and inconsistent far conditions, respectively. Planned comparisons indicated that consistent trials were faster than the other trials, $t(29) = -4.83$, $p < .0001$. The inconsistent close condition was significantly slower than the inconsistent far condition, $t(29) =$

3. For example, the inconsistent close condition for the subitizing group would be the average of displays with one 2, two 3s, three 4s (identity smaller than numerosity), three 2s, four 3s, and five 4s (identity larger than numerosity), that is, the average of 568, 620, 645, 630, 667, and 674 ms, respectively.

Table 2
Trials collapsed in the conditions Inconsistent Close and Inconsistent Far, for the Subitizing and Counting Groups.

	Conditions			
Numerosity	Close (+1)	Far (+2)	Close (-1)	Far (-2)
One	2	3	-	-
Two	3	4	-	-
Three	4	5	2	1
Four	-	-	3	2
Five	-	-	4	3
Five	6	7	-	-
Six	7	8	-	-
Seven	8	9	6	5
Eight	-	-	7	6
Nine	-	-	8	7

2.47, $p < .02$, but did not differ from the neutral condition $t(29) = -0.53$, $p > .5$.

The condition by group interaction approached significance, $F(3, 84) = 2.49$, $MSe = 891.2$, $p = .065$. This interaction was caused by the particular behavior of the neutral condition, which was faster than the inconsistent close condition in the subitizing group and slower than the inconsistent close condition in the counting group. A similar ANOVA, carried out by excluding the neutral condition, showed no significant interaction between condition and group ($F < 1$).

To further investigate the effect of symbolic distance as a function of group, we conducted two separate one-way repeated measures ANOVAs for the subitizing and for the counting group. In the subitizing group, the effect of condition was significant, $F(3, 42) = 40.8$, $MSe = 83.9$, $p < .0001$. The mean RTs were 597, 619, 633, and 623 for the consistent, neutral, inconsistent close, and inconsistent far conditions, respectively. Planned comparisons showed that, except for the difference between the neutral and inconsistent far trials ($p > .25$), all other differences were significant ($ps < .01$). In particular, inconsistent close trials were significantly slower than inconsistent far trials, $t(14) = 3.27$, $p < .006$.

In the counting group, the effect of condition was also significant, $F(3, 42) = 5.693$, $MSe = 1{,}698$, $p < .005$. The mean RTs were 1,504, 1,557, 1,553, and 1,521 for the consistent, neutral, inconsistent close, and inconsistent far conditions, respectively. Planned comparisons revealed that the consistent condition did not differ from the inconsistent far condition and that the neutral condition did not differ from the inconsistent close condition ($p > .25$). Inconsistent close trials were significantly slower than inconsistent far trials, $t(14) = 1.915$, $p = .038$, one-tailed .

Error data. A condition by group mixed ANOVA was carried out on error percentages. The only significant effect was the main effect of group, $F(1, 28) = 21.25$, $MSe = 23.46$, $p < .0001$. The subitizing group produced fewer errors than the counting group (1.1% and 5.2%, respectively). Error percentages for condition were 2.6, 3.0, 3.3, and 3.6 for the consistent, neutral, inconsistent close, and inconsistent far conditions, respectively. However, neither the main effect of condition, $F(3, 84) = 1.79$, $MSe = 3.12$, $p > .15$, nor the condition by group interaction, $F(3, 84) = 1.46$, $MSe = 3.12$, $p > .23$, reached significance.

Separate one-way repeated measures ANOVAs were carried out for the two groups. In the subitizing group, the main effect of condition was significant, $F(3, 42) = 3.847$, $MSe = 0.936$, $p < .05$. The percentages of errors were 0.6, 0.9, 1.8, and 0.9 for the consistent, neutral, inconsistent close, and inconsistent far conditions, respectively. Planned comparisons revealed that the inconsistent close condition was less accurate than the other conditions ($p < .05$). In the counting group, the effect of condition was not significant, $F(3, 42) = 1.242$, $MSe = 5.436$, $p > .3$. The percentages of errors were 4.5, 5.1, 4.9, and 6.1 for the consistent, neutral, inconsistent close, and inconsistent far conditions, respectively.

Discussion

The results confirm the presence of the SDE on Stroop interference, as has been found in previous experiments (Hock & Petrasek, 1973; Pavese & Umiltà, 1997). Analyses carried out separately by group indicated that in both the subitizing and the counting groups, digits close to the enumeration response (±1) produced more interference than digits that were far from the enumeration response (±2).

In the subitizing group, neutral trials were slower than consistent trials and faster than inconsistent close trials, but inconsistent far trials did not differ from neutral trials. More important, the inconsistent close condition was significantly slower than the inconsistent far condition, and error rates were higher in the inconsistent close condition than in the other conditions. Also in the counting group, the inconsistent close condition was slower than the inconsistent far condition. However, in this group the neutral condition was particularly slow and did not differ from the inconsistent close condition.

There can be little doubt that interference was greater for the inconsistent close condition than for the inconsistent far condition. The results of the subitizing group, in which the inconsistent far condition was as fast as the neutral condition, suggest that interference was limited to the inconsistent close condition. Unfortunately, a similar comparison for the counting group was not informative because of problems with the neutral condition (discussed later).

Enumeration processes

The results nicely replicate the typical findings on counting and subitizing. In the subitizing group, there was a small but reliable difference in RTs when the numerosity of the display was increased by one item (average increment = 25 ms). The significant linear contrast demonstrated that RTs increased as a function of numerosity, although the significant quadratic trend indicated that the slope was not constant. This might have been the result of the limited number of items in the response (and stimulus) set. Because only five possible responses were possible for this group, each intermediate value (2, 3, and 4) had to be discriminated from two other possible adjacent responses. The extreme high and low values had to be discriminated only from a single similar value. Therefore, it was relatively easier to respond to displays with one or five items than to displays with two, three, or four items (the end effect; Folk, Egeth, & Kwak, 1988; Mandler & Shebo, 1982). The error data showed another typical result: Accuracy was almost perfect for small numerosities (1-3) but decreased for displays of four and five items (Mandler & Shebo, 1982).

In the counting group, RT increment as a function of numerosity was much larger (average increment = 256 ms), and the error rates increased overall. The pattern of errors was different from that found in the subitizing group. There was a significant increase in error rates between 5 and 7, but the error rate remained stable between 7 and 9. Also in this group, the linear and quadratic contrasts on latencies were significant, suggesting a combination of a linear increment attributable to numerosity and an effect of the serial position in the response set. Extreme values (5 and 9) were relatively easier to respond to than intermediate values.

It is interesting to examine the results of the five-item displays because they were presented to both experimental groups. As shown in Figure 1, RTs to five-item displays were much faster (657 vs. 926 ms) but also less accurate (3.6% vs. 1.7%) in the subitizing group than in the counting group. This pattern suggests that participants in the two groups enumerated the five-item displays in different ways. Five is often mentioned as being the boundary between the subitizing and the counting range. Participants likely used a “subitizing” strategy or a “counting” strategy with five-item displays depending on the set of possible numerosities that was used in the task (Mandler & Shebo, 1982). In this experiment, the subitizing group may have been induced to use a strategy of “direct apprehension” (Gelman & Gallistel, 1978), which is faster and highly efficient with smaller numerosities, but yielded increasing error rates as the numerosity increased, starting from four-item displays. In contrast, the counting group might have been induced to use a more conservative strategy, which consisted of also counting the fove-item displays, resulting in a slower but more accurate performance.

In summary, these results show that the two groups differed in the type of enumeration process used. The subitizing group used a faster process, accurate for small numerosity but with error rates that increased as a function of numerosity for four- and five-item displays. The counting group used a slower process that was more accurate for smaller displays and had a higher cost in latencies for each additional unit. In the counting group, error rates increased between 5 and 7, but was almost constant for numerosities between 7 and 9. The slope of the two curves is consistent with other results in the enumeration literature (Jensen, Reese, & Reese, 1950; Mandler & Shebo, 1982; Trick & Pylyshyn, 1994).

Despite these clear differences, the effect of symbolic distance on interference was similar in the two groups. The difference between the close and far inconsistent conditions was somewhat larger in the counting group than in the subitizing group (32 and 10 ms, respectively), but the condition by group interaction was not significant when the neutral condition was excluded. The only reliable difference between the two groups was indeed in the behavior of the neutral condition, which is discussed in the next section.

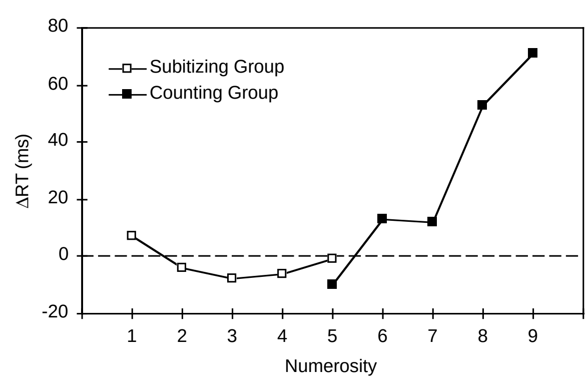
Neutral Condition and Numerosity

As mentioned earlier, the neutral condition could not be considered a reliable baseline in the counting group. In the subitizing group, the average RTs of neutral trials for each numerosity were close to the average RTs for all other conditions in which the items to be counted were digits. In the counting group, however, the neutral condition tended to become increasingly slower than the average RT of the other conditions for the same numerosity (see Figure 2). An analysis of the difference between the neutral condition and the average of the other conditions revealed a significant linear trend across numerosities in the counting group, $\underline{F}(1, 56) = 4.478$, $\underline{p} < .05$.

It is noteworthy that in the counting group the high degree of interference in the neutral condition was found only in the first session. An ANOVA that examined the effect of condition and session revealed that the RT was 198 ms faster in the second session than in the first session, $\underline{F}(1, 14) = 65.066$, $\underline{MSe} = 18{,}198$, $\underline{p} < .0001$, and that session interacted with condition, $\underline{F}(3, 42) = 2.903$, $\underline{MSe} = 2{,}086$, $\underline{p} < .05$. In the first session, the average RTs were 1,614, 1,683, 1,676, and 1,660 ms for the consistent, neutral, inconsistent close and inconsistent far conditions, respectively, whereas in the second session the average RTs for the same conditions were 1,404, 1,450, 1,498, and 1,487 ms, respectively. These results suggest that (a) counting performance remarkably improves with practice and (b) the high level of interference from letters occurs only at the beginning of the practice with the task. These changes in performance may reflect a change in the counting process from an algorithmic computation to a memory retrieval process, as proposed by Lassaline and Logan (1993).

The increasing interference effect from letter displays as a function of numerosity is a new and unexpected finding. However, to our knowledge, none of the researchers investigating numerical Stroop paradigms used numerosities larger than 6. Some researchers have investigated numerosities in the 0 or 1-3 range (Flowers et al., 1979; Francolini & Egeth, 1980; Windes, 1968) and others have used numerosities between one and six (Fox et al., 1971; Morton, 1969; Shor, 1971). Furthermore, only some studies used letter as a control condition (Fox et al., 1971; Francolini & Egeth, 1980; Morton, 1969), whereas others used symbols such as circles, asterisks, or compatible digits as control against which to measure interference. In our experiment, in which numerosities between 1 and 5 were tested, a "traditional" congruency effect was found in the subitizing group. An analysis comparing the average of all the inconsistent conditions with

Figure 2.
Mean reaction times for the Neutral condition minus the average of reaction times for the other three conditions (Consistent, Inconsistent Close and Inconsistent Far) as a function of Numerosity and Group.



consistent and neutral trials in the subitizing group showed that letters were counted slower than consistent digits, $t(14) = -8.742$, $p < .0001$, and faster than inconsistent digits, $t(14) = 4.857$, $p < .005$.

There are two possible expalanations of the longer latency of the neutral condition in the first session of the counting group: (a) The neutral condition produced more interference for larger displays or (b) the inconsistent conditions became less interfering for larger displays. Remember that the stimulus array was displayed until participants responded and it was not masked and that larger displays were enumerated more slowly than smaller displays. Therefore, larger numerosities differed from smaller numerosities in both enumeration latency and exposure time.

A tentative explanation of the increasing interference for neutral trials may assume that, because the task was enumeration, the number domain was more activated, or primed, than the letter domain (MacLeod, 1991). The higher default activation of digits compared with the default activation of letters may explain why digits caused more interference than letters, at least for small numerosities. However, the representations of nonrelevant letter identities might have reached a higher level of activation with longer enumeration latencies, yielding an increase in their interference effect. Previous studies have reliably shown that letters produce a significant interference effect in enumeration tasks when compared with abstract symbols or circles (Fox et al., 1971; Morton, 1969).

An alternative possibility is that for larger numerosities, inconsistent digits produced less interference, yielding a relative decrease in their latencies and a relative increase in neutral trial latencies. It has been suggested that nonrelevant dimensions are inhibited during attentive selection (Tipper, 1985), and that inhibition needs time to develop (Neill & Westberry, 1987). It is possible that, for longer latencies, inhibition of a nonrelevant digit identity yielded a selective reduction of interference. The prediction that the relative delay of the neutral condition over the inconsistent conditions is caused by inhibition requires the additional assumption that inhibition is more strongly associated with digits than with letters (for a similar suggestion, see Neill, Valdes, & Terry, 1995; Tipper, Weaver, & Houghton, 1994). Additional investigations are required to confirm these results and to lend support to one of these alternative hypotheses.

The SDE, Stroop interference, and the Compressed Number Line

It has been proposed that the representation of magnitude associated with numbers can be thought of as a compressed number line (Dehaene, 1992; Restle, 1970). According to this model, the distance between one quantity and the next on the number line decreases as a function of numerosity. For example, the distance between 2 and 3 would be larger than the distance between 3 and 4, which in turn would be larger than the distance between 4 and 5, and so on. If interference is a function of the symbolic distance between digit identity and enumeration response, one should also expect that the difference in interference between inconsistent close and inconsistent far digits would not be constant. Rather, it should increase as a function of numerosity, because the symbolic distance between adjacent digits decreases as the absolute value increases.

Two predictions of the compressed number line hypothesis are as follows: (a) For a given arithmetic difference between digit identity and enumeration response, interference should increase with numerosity and (b) interference should be greater when the digits to be counted are larger than the enumeration response than when they are smaller than the enumeration response (e.g., the symbolic distance between 3 and 4 is larger than the symbolic distance between 4 and 5). These predictions were tested on the RTs of the subitizing group, because in the counting group the unreliability of the neutral condition as a baseline made it difficult to assess the amount of interference.

<u>Effect of numerosity</u>. Interference effects (differences between RTs in the neutral condition and RTs in the inconsistent close condition) were computed for each numerosity and for each participant in the subitizing group. Mean interference effects were 1, 12, 18, 22, and 22 ms for

numerosities one, two, three, four, and five, respectively. The linear contrast was significant, $F(1, 56) = 9.452$, $p < .005$, indicating that interference linearly increased as a function of numerosity. Neither the quadratic, $F(1, 56) = 1.632$, $p > .20$, nor the cubic contrasts ($F < 1$) were significant.

Effect of larger and smaller digit identity. In the subitizing group, three numerosities—2, 3 and 4—were displayed using both digits that were greater than the enumeration response (e.g., three 4s) and digits that were smaller than the enumeration response (e.g., three 2s). A two-tailed t test was carried out on the interference effect for larger and smaller digit identities across these three numerosities. Larger digits produced significantly more interference (M = 23 ms, SD = 18) than did smaller digits (M = 12 ms, SD = 11), $t(14) = 3.045$, $p < .01$.

Both predictions of the compressed number line hypothesis were supported in the subitizing group: (a) Interference from symbolically close digits linearly increased as a function of numerosity and (b) interference was greater when the digits were larger than the enumeration response than when they were smaller than the enumeration response. These findings support the hypothesis that the SDE found in the subitizing group is similar to that found in comparison tasks and in priming tasks. In particular, these three phenomena are likely to be associated with the activation of the same magnitude representation, organized as a compressed number line.

General Discussion

In this study, we found that Stroop-like interference in an enumeration task was affected by the arithmetic difference between the relevant dimension (numerosity) and the nonrelevant dimension (digit identity). The effect was present with different numerosities and different enumeration processes. Furthermore, the results of the subitizing group suggest that this effect is based on the same type of magnitude representation—a compressed number line—proposed to explain the SDE in comparison tasks: (a) For a given numerosity, interference was greater when the digits to be counted were larger than the enumeration response than when the digits were smaller than the enumeration response and (b) interference from symbolically close digits increased as a function of numerosity. As found in the color-word Stroop task by Klein (1964), and more recently by Klopfer (1996), the results of our study confirm that interference reflects the degree of similarity between the representations of relevant and nonrelevant information.

Another interesting aspect of our results is that they also confirm that the SDE is not limited only to comparison judgment tasks. Although discrete models of the SDE (Banks, 1977) have emphasized the importance of the comparison process, our findings suggest that the SDE is a more basic effect related to the characteristic of number representation rather than to the specifics of the task. Not only is the SDE found in tasks that do not require comparison, as suggested by data from priming experiments (Brysbaert, 1995; den Heyer & Briand, 1986; Marcel & Forrin, 1974), but it is also found in selective attention tasks in which one of the numerical dimensions is nonrelevant to the task and should be ignored.

These results will be discussed with reference to the processing and representation of numerical information.

Autonomous Processing of Numerical Information

Zbrodoff and Logan (1986) referred to involuntary or unintentional processing as “autonomous”. The processing of a stimulus dimension is autonomous whenever such a dimension, although irrelevant to the task, affects performance. Several studies have shown that numerical magnitude information is autonomously activated even when it is irrelevant to the task. For instance, Duncan and McFarland (1980, Experiment 2) and Henik and Tzelgov (1982) found that the numerical difference between digits affects RTs even when the task can be performed on the basis of perceptual properties. Sudevan and Taylor (1987) showed that the larger-smaller status of the target digit interfered with the odd-even classification, suggesting a mandatory activation of numerical comparison. Dehaene, Bossini, and Giraux (1993, Experiment 1) also reported a compatibility effect in a parity judgment task. When large numbers were presented latencies were shorter with right-hand key responses, whereas with small numbers shorter latencies were found with left-hand key responses (the Spatial-Numerical Association of Response Codes [SNARC] effect).

However, some authors have questioned the specificity of the numerical information that is autonomously activated. Tzelgov, Meyer, and Henik (1992) asked participants to evaluate either the physical or the numerical size of digits by comparing them with a standard presented at the beginning of the block. They found that when numerical size was irrelevant only a crude, dichotomous representation of numerical size was encoded. Digits 1-4 were classified as "small" and digits 6-9 were classified as "large", with a neutral midpoint around 5. Tzelgov et al. suggested that the processing of nonrelevant information varies according to the operations that must be performed on the relevant information.

In our study, significant differences in interference were found with nonrelevant digit identities included in the ranges 1-5 and 5-9. This result is not consistent with a strong version of Tzelgov et al. (1992) hypothesis, stating that distractor numbers smaller than 5 are categorized as small and numbers larger than 5 are categorized as large. A weaker version of this theory, proposing that the small-large categorization is relative to the stimulus set used in the experiment, it was supported either. For example, in the subitizing group, three-item displays were responded to 18 ms faster, $t[14] = 2.23$, $p < .05$, when the items to be counted were 5s than when they were 4s, which both should be categorized as large in this group. Our results show that more specific magnitude information is autonomously activated. Note, however, that in our experiments accurate magnitude information might have been relevant because the task required participants to enumerate the items in the display.

It would seem likely that under our experimental conditions, information that was necessary to perform the task (i.e., magnitude information) was encoded for both relevant and nonrelevant dimensions of the stimuli. Activation of magnitude information associated with the relevant dimension produced the correct enumeration response, whereas the activation of magnitude information associated with the irrelevant dimension was responsible for the interference effect.

Number representation

Campbell and Clark (1988) excluded the existence of a central abstract representation and suggested that visuo-spatial, verbal, and other modality-specific number codes are associatively connected and activate each other during retrieval and calculation. McCloskey, Caramazza, and Basili (1985), in contrast, proposed a model in which an amodal, abstract representation of numbers constitutes the entry to calculation routines and to stored number knowledge. This model, is different from Campbell and Clark's model in that it postulates a mandatory access to the magnitude code before any further number processing.

Dehaene (1992) proposed a triple-code model of number processing. This model assumes three main mental representations of number: (a) an auditory verbal word form in which numbers are represented in verbal notation, (b) a visual Arabic number form in which numbers are represented in Arabic notation, and (c) an analog magnitude representation. Following the original work of Moyer and Landauer (1967) and Restle (1970), this magnitude representation is thought to take the form of an analog number line oriented from left to right. Its quasi-spatial characteristics can account for the number and space interaction in the SNARC effect and similar results (Dehaene et al., 1993)

The results of our experiment support McCloskey e al.'s (1985) and Dehaene's (1992) views that a magnitude representation of numbers exists and can be autonomously activated. It is, however, possible that this representation is activated (for both relevant and nonrelevant numerical information) only when the task to be performed requires the manipulation of magnitude information related to numbers (Tzelgov et al., 1992).

Conclusions

The results of this study show that (a) Stroop-like interference is affected by the arithmetic distance between the enumeration response and item identity and (b) for numerosities 1-5, interference seems to reflect the activation of a magnitude representation of digit identity that is organized as a compressed number line. These results suggest that interference effects found in Stroop-like enumeration tasks depend on the trial-by-trial activation of magnitude representations of both relevant and nonrelevant dimensions. Therefore, they support models of numerical processing that assume that nonrelevant digit identity autonomously activates its associated magnitude representations.

Our results also support the existence of an analog magnitude representation, such as Restle's (1970) number line, at least for small numerosities, and confirm that the SDE found in comparison judgment tasks reflects a general property of number representation (Dehaene, 1992).

REFERENCES

Aiken, L. R., & Williams, F. N. (1968). Three variables related to reaction time to compare single-digit numbers. Perceptual & Motor Skills, 27, 197-206.

Atkinson, J., Campbell, F. W., & Francis, M. R. (1976). The magic number 4 +/- 0: A new look at visual numerosity judgements. Perception, 5, 327-334.

Banks, W. P. (1977). Encoding and processing of symbolic information in comparative judgments. In W. Bower (Ed.), The psychology of learning and motivation, (Vol. XI, pp. 101-159). San Diego, CA: Academic Press.

Banks, W. P., Fujii, M., & Kayra-Stuart, F. (1976). Semantic congruity effects in comparative judgments of magnitudes of digits. Journal of Experimental Psychology: Human Perception and Performance, 2, 435-447.

Brysbaert, M. (1995). Arabic number reading: On the nature of the numerical scale and the origin of phonological recoding. Journal of Experimental Psychology: General, 124, 434-452.

Campbell, J. I. D., & Clark, J. M. (1988). En encoding-complex view of cognitive number processing: Comment on McCloskey, Sokol and Goodman (1986). Journal of Experimental Psychology: General, 117, 204-214.

Cohen, J., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphical system for designing and controlling experiments in the psychological laboratory using Macintosh computers. Behavioral Research Methods, Instrumentation, and Computation, 25, 251-271.

Dehaene, S. (1992). Varieties of numerical abilities. Cognition, 44, 1-42.

Dehaene, S., & Akhavein, R. (1995). Attention, automaticity, and levels of representation in number processing. Journal of Experimental Psychology: Learning, Memory, and Cognition, 21, 314-326.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. Journal of Experimental Psychology: General, 122, 371-396.

Dehaene, S., & Changeux, J.-P. (1993). Development of elementary numerical abilities: A neuronal model. Journal of Cognitive Neuroscience, 5(4), 390-407.

Dehaene, S., Dupoux, E., & Mehler, J. (1990). Is numerical comparison digital? Analogical and symbolic effects in two-digit number comparison. Journal of Experimental Psychology: Human Perception and Performance, 16, 626-641.

den Heyer, K., & Briand, K. (1986). Priming single digit numbers: Automatic spreading activation dissipates as a function of semantic distance. American Journal of Psychology, 99(3), 315-340.

Duncan, E. M., & McFarland, C. E. J. (1980). Isolating the effects of symbolic distance and semantic congruity in comparative judgments: An additive-factors analysis. Memory and Cognition, 8, 612-622.

Flowers, J. H., Warner, J. L., & Polansky, M. L. (1979). Response and encoding factors in "ignoring" irrelevant information. Memory and Cognition, 7, 86-94.

Folk, C. L., Egeth, H., & Kwak, H. (1988). Subitizing: Direct apprehension or serial processing? Perception & Psychophysics, 44, 313-320.

Folz, G. S., Poltrock, S. E., & Potts, G. R. (1984). Mental comparison of size and magnitude: Size congruity effects. Journal of Experimental Psychology: Learning, Memory, and Cognition, 10, 442-453.

Fox, L. A., Shor, R. E., & Steinman, R. J. (1971). Semantic gradients and interference in naming color, spatial direction and numerosity. Journal of Experimental Psychology, 91, 59-65.

Francolini, C. M., & Egeth, H. E. (1980). On the nonautomaticity of “automatic” activation: Evidence of selective seeing. Perception & Psychophysics, 27, 331-342.

Gelman, R., & Gallistel, C. R. (1978). The child’s understanding of number. Cambridge: Harvard University Press.

Henik, A., & Tzelgov, J. (1982). Is three greater than five: The relation between physical and semantic size in comparison tasks. Memory and Cognition, 10, 389-395.

Hinrichs, J. V., Yurko, D. S., & Hu, J. (1981). Two-digit number comparison: Use of place information. Journal of Experimental Psychology: Human Perception and Performance, 7, 519-584.

Hock, H. S., & Petrasek, J. (1973). Verbal interference with perceptual classification: The effect of semantic structure. Perception & Psychophysics, 13, 116-120.

Jensen, E. M., Reese, E. P., & Reese, T. W. (1950). The subitizing and counting of visually presented fields of dots. Journal of Psychology, 30, 363-392.

Kaufman, E., Lord, M., Reese, T., & Volkmann, J. (1949). The discrimination of visual number. American Journal of Psychology, 62, 498-525.

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

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

Lassaline, M. E., & Logan, G. D. (1993). Memory-based automaticity in the discrimination of visual numerosity. Journal of Experimental Psychology: Learning, Memory, and Cognition, 19, 561-581.

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

Mandler, G., & Shebo, B. J. (1982). Subitizing: An analysis of its component processes. Journal of Experimental Psychology: General, 111, 1-22.

Marcel, T., & Forrin, B. (1974). Naming latency and the repetition of stimulus categories. Journal of Experimental Psychology, 103, 450-460.

McCloskey, M., Caramazza, A., & Basili, A. (1985). Cognitive mechanisms in number processing and calculation: Evidence from dyscalculia. Brain and Cognition, 4, 171-196.

Morin, R. E., DeRosa, D. V., & Stultz, V. (1967). Recognition memory and reaction time. Acta Psychologica, 27, 298-305.

Morton, J. (1969). Categories of interference: verbal mediation and conflict in card sorting. British Journal of Psychology, 60, 329-346.

Moyer, R. S. (1973). Comparing objects in memory: Evidence suggesting an internal psychophysics. Perception and Psychophysics, 13(2), 180-184.

Moyer, R. S., & Bayer, R. H. (1976). Mental comparison and the Symbolic Distance Effect. Cognitive Psychology, 8, 228-246.

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgements of numerical inequality. Nature, 215, 1519-1520.

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.), New perspectives on interference and inhibition in cognition, (pp. 207-261). New York: Academic Press.

Neill, W. T., & Westberry, R. L. (1987). Selective attention and the suppression of cognitive noise. Journal of Experimental Psychology: Learning, Memory, and Cognition, 13, 327-334.

Parkman, J. M. (1971). Temporal aspects of digit and letter inequality judgments. Journal of Experimental Psychology, 91, 191-205.

Pavese, A., & Umiltà, C. (1997). Symbolic Distance Effect in an enumeration task. Manuscript submitted for publication.

Poltrock, S. E. (1989). A random walk model of digit comparison. Journal of Mathematical Psychology, 33, 131-162.

Proctor, R. W. (1978). Sources of color-word interference in the Stroop color-naming task. Perception & Psychophysics, 23, 413-419.

Redding, G. M., & Gerjets, D. A. (1977). Stroop effect: Interference and facilitation with verbal and manual responses. Perceptual & Motor Skills, 45, 11-17.

Restle, F. (1970). Speed of adding and comparing numbers. Journal of Experimental Psychology, 83, 274-278.

Scheibe, K., Shaver, P. R., & Carrier, S. C. (1967). Color association values and response interference on variants of the Stroop test. Acta Psychologica, 26, 286-295.

Sekuler, R., & Mierkiewicz, D. (1977). Children's judgments of numerical inequality. Child Development, 48, 630-633.

Sekuler, R., Rubin, E., & Armstrong, R. (1971). Processing numerical information: A choice time analysis. Journal of Experimental Psychology, 90, 75-80.

Shor, R. E. (1971). Symbol processing speed differences and symbol interference effects in a variety of concept domains. Journal of General Psychology, 85, 187-205.

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

Sudevan, P., & Taylor, D. A. (1987). The cueing and priming of cognitive operations. Journal of Experimental Psychology: Human Perception and Performance, 13, 89-103.

Tipper, S. P. (1985). The negative priming effect: Inhibitory priming by ignored objects. The Quarterly Journal of Experimental Psychology, 37A, 571-590.

Tipper, S. P., Weaver, B., & Houghton, G. (1994). Behavioural goals determine inhibitory mechanisms of selective attention. The Quarterly Journal of Experimental Psychology, 47A, 809-840.

Trick, L. M., & Pylyshyn, Z. W. (1993). What enumeration studies can show about spatial attention: Evidence for limited capacity preattentive processing. Journal of Experimental Psychology: Human Perception and Performance, 19, 331-351.

Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. Psychological Review, 101, 80-102.

Tzelgov, F., Meyer, J., & Henik, A. (1992). Automatic and intentional processing of numerical information. Journal of Experimental Psychology: Learning, Memory, and Cognition, 18, 166-179.

Welford, A. T. (1960). The measurement of sensory-motor performance: Survey and reappraisal of twelve years' progress. Ergonomics, 3, 189-230.

Windes, J. D. (1968). Reaction time for numerical coding and naming numerals. Journal of Experimental Psychology, 78, 318-322.

Zbrodoff, N. J., & Logan, G. D. (1986). On the autonomy of mental processes: A case study of arithmetic. Journal of Experimental Psychology: Learning, Memory, and Cognition, 16, 83-97.