Psychological Res (1999) 62:62-71.

ORIGINAL ARTICLE

Antonella Pavese · Carlo Umiltà

Further evidence on the effects of symbolic distance on Stroop-like interference

Received: 10 November 1997 / Accepted: 23 June 1998

Abstract Pavese and Umiltà found that, in an enumeration task, Stroop-like interference is larger when the digit identity is symbolically close to the enumeration response than when it is symbolically far. In 2 experiments testing 49 undergraduates, we further explored this phenomenon using Francolini and Egeth's paradigm. We found that symbolic distance affected interference even when the stimulus was briefly presented and masked. In Experiment 2, which tested numerosities outside the subitizing range, individuals used a different enumeration strategy but showed the same symbolic distance effect. These results support the hypothesis that Stroop interference found in enumeration tasks depends on a rapid and automatic activation of digits' magnitude representation.

Introduction

When people attend to a target stimulus dimension, manipulations of non-relevant stimulus dimensions affect response latencies and error rates (MacLeod, 1991). For example, when people are asked to name the color in which words are written, non-relevant color words interfere with the naming task (Stroop, 1935). Color words incongruent with the color naming response produce longer latencies and higher error rate than neutral stimuli (color patches, arrays of Xs or non-color words); often, a small facilitation effect is found when the color word is congruent with the naming response (MacLeod, 1991). These results suggest that non-relevant stimulus dimensions activate associated representations that interfere with the task that the individual is performing.

A. Pavese (✉)
Moss Rehabilitation Research Institute, 1200 West Tabor Road, Philadelphia, PA 19141; Ph. (215) 456-9901 ext. 9220.
E-mail: pavese@hslc.org

C. Umiltà
University of Padua, Dept. of General Psychology
Padua, Italy

Which kind of representation is activated by non-relevant stimulus dimensions? One possibility is that only phonological representations associated with the words are activated, and that this activation interferes with the production of color word responses. This hypothesis was ruled out by Klein (1964), who varied the strength of semantic association between the words and the concept of color (e.g., color words, words strongly associated with colors such as "sky", and words not associated with colors such as "house"). Klein demonstrated that interference varies on a continuum as a function of semantic association between relevant and non-relevant dimensions. The finding of a semantic gradient of interference suggests that a semantic representation of the word is activated, and that the°ree of association between this semantic representation and the *concept of color* modulates the amount of interference (MacLeod, 1991).

It is still possible that a more precise relationship between relevant and non-relevant dimensions determines the amount of Stroop interference. For example, different color words may differentially affect the naming of a particular color as a function of the similarity between the color to be named and the color represented by the word. Klopfer (1996) reports evidence in support of this hypothesis: Words representing colors perceptually similar to the color to be named (e.g., "RED" written in orange) produced a larger amount of interference than words representing colors perceptually dissimilar from the naming response (e.g., "RED" written in blue).

Several studies have found Stroop-like interference effects in the number domain (i.e., Flowers, Warner, & Polansky, 1979; Morton, 1969; Shor, 1971; Windes, 1968). Francolini and Egeth (1980) reported results similar to those found in the color naming task: Digits congruent with the enumeration response were enumerated faster than neutral letters, which in turn were enumerated more rapidly than digits incongruent with the enumeration response. Fox, Shor, and Steinman (1971) found in the number domain a semantic gradient effect similar to that reported by Klein in the color domain. Symbols more associated to the concept of number, such as Arabic or Roman numerals, produce larger interference effects than symbols not associated to the concept of color, such as letters or common words. Finally, Pavese and Umiltà (1998) found that the arithmetic difference between the digits to be counted and the enumeration response affects interference. Digits symbolically close to the enumeration response produce larger interference than digits symbolically far from the response. This result—comparable to Klopfer's (1996) effect of similarity between color word and color to be named—suggests that also in numerical variations of the Stroop task, non-relevant digit identity activates an abstract representation associated with the "meaning" of the digit, that is, its magnitude.

The present study further investigates the effect of symbolic distance between relevant and non-relevant dimensions in an enumeration task first reported by Pavese and Umiltà (1998). In particular, the present experiments aim at testing the hypothesis that the magnitude representation of digit identity is rapid and automatic and does not depend on stimulus exposure and enumeration strategy.

The Representation of Number Magnitude

Magnitude is the most salient symbolic property of numbers (Shepard, Kilpatric, & Cunnigham, 1975). Number magnitude was first studied using comparison judgement tasks, in which individuals have to decide which of two numbers displayed is the larger. The interesting result of these studies is that the arithmetic difference between two numbers affects the time required for the judgement. For example, people are faster to decide that 5 is greater than 1 than to decide that 5 is greater than 4 (Moyer & Bayer, 1976; Moyer & Landauer, 1967). This result closely mimics results found in perceptual judgement tasks, as well as in symbolic judgements in which participants compare words on the basis of perceptual characteristics (e.g., physical size) of the objects they represent. Moyer and Bayer (1976) concluded that magnitude information associated with numbers is represented as a continuous dimension and shares similar properties with the representation of physical dimensions. For example, for a given symbolic distance, comparison latency increases as a function of the absolute magnitude of the two numbers to be compared, suggesting that magnitude is represented on a logarithmic scale, as physical properties are (Welford, 1960). Dehaene, Bossini, and Giraux (1993) showed that this analogue representation has a precise orientation, so that individuals are faster to respond to small numbers with the left hand and to large numbers with the right hand (spatial-numerical association of response codes or SNARC effect). These findings strongly suggest that magnitude is represented as a compressed number line oriented from left to right.

Evidence of activation of an analogue magnitude representation is not limited to comparison judgements. Symbolic distance effects (SDEs) have been found in same-different judgements (Duncan & McFarland, 1980), in naming tasks (Brysbaert, 1995; den Heyer & Briand, 1986; Marcel & Forrin, 1974), and in recall tasks (Morin, DeRosa & Stultz, 1967). These results suggest that the SDE does not depend on the task that is performed and that magnitude information can be activated even when is not relevant to the task. Studies that investigated the SNARC effect also support this conclusion. For example, Fias et al. (1996) found left hand advantage for responses to smaller numbers and right hand advantage for responses to larger numbers using a task, phoneme monitoring, that does not require access to numerical information. This result further demonstrates that the relative position of numbers on an oriented number line is automatically computed, even when is irrelevant to the task at hand.

Symbolic Distance Effect and Stroop interference

Pavese and Umiltà (1998) showed that symbolic distance also affects Stroop interference[1]. In this study, participants were asked to enumerate items in circular displays of different numerosities (one to nine). The results showed that the enumeration task was not only influenced by the item category (digits or letters) and by response congruency, but also by the arithmetic difference between display numerosity and digit identity. Digits that were symbolically close to the enumeration response produced larger interference than symbolically far digits, which did not produce a reliable interference effect when compared to letters.

The existence of an effect of symbolic distance on Stroop interference is interesting in several respects. First, interference effects can be used to explore the underlying structure of activated representations. For example, Pavese and Umiltà (1998) tested whether the pattern of interference was consistent with previous models of magnitude representation by manipulating numerosity and using digit identity larger and smaller than the enumeration response. The results showed that (1) interference linearly increases as a function of numerosity in the range one to five and that (2) digit identity larger than the enumeration response yields larger interference effects than digit identity smaller than the enumeration response. These results support the hypothesis that magnitude representation is organized as a compressed number line, at least for small numerosities.

1 An experiment that bears some resemblance with both Pavese and Umiltà's (in press) Exp. and the present study, was published by Washburn (1994). He tested Rhesus monkeys and humans in a variation of the Stroop task in which the subjects chose the larger of two digit arrays simultaneously presented. The results showed that both interference and facilitation increased as the symbolic distance between digit identities in the two arrays increased. It must be kept in mind, however, that, contrary to the present Exp., Washburn did not measure the effect of the distance between identity and numerosity, but rather the amount of interference and facilitation on the comparison task associated with the difference between the two arrays in number of items *or* digit identities.

Another interesting aspect of this paradigm is that it makes it possible to investigate whether the enumeration strategy used to perform the task influences the SDE. Enumeration, differently from color naming, can be performed in qualitatively distinct ways depending on the number of items to be enumerated and on display duration.

When the stimulus is presented until response, enumeration latencies monotonically increase as a function of numerosity, with a slope discontinuity around four (Mandler & Shebo, 1982): for small numerosities, each additional unit yields an increment in reaction times (RTs) around 40-100 ms, whereas for larger numerosity the increment per unit is around 250-350 ms. Accuracy is almost perfect for numerosity between one and three and slowly decreases for larger numerosities (Pavese & Umiltà, 1998). When the stimulus is briefly presented (200 ms), however, latency increases as function of numerosity only up to six items; for numerosities larger than six, RTs become stable and are not influenced by numerosity (Kaufman et al., 1949; Mandler & Shebo, 1982). Also, with short stimulus exposures, accuracy steeply decreases for displays larger than three items (Mandler & Shebo, 1982).

On the basis of these qualitatively different enumeration performances, Kaufman, Lord, Reese, and Volkmann (1949) proposed the existence of three different processes: subitizing, counting, and estimation. Kaufman et al. (1949) called subitizing the confident, accurate, and rapid enumeration process that is observed with a small number of items. The subitizing range is widely defined as one to four, although the literature reports different estimates (see, for example, Mandler & Shebo, 1982), and there is evidence of remarkable individual differences (Atkinson, Campbell, & Francis, 1976; Mandler & Shebo, 1982; Trick & Pylyshyn, 1993, 1994). Counting is the effortful, slower, and error-prone process that is observed for larger numerosities when the display is presented until response. Finally, estimation is an enumeration strategy faster than counting but remarkably less accurate, which is used when stimulus configurations with a large number of items are briefly presented.

Pavese and Umiltà (1998) investigated the effect of enumeration strategy using two different groups of numerosity, one to five, roughly within the subitizing range, and five to nine, in the counting range. The results showed that the effect of symbolic distance was similar in the two groups, suggesting that the underlying representation and not the enumeration process used determined the properties of the SDE.

Overview

The purpose of this study is to replicate the effect of symbolic distance on Stroop interference found in Pavese and Umiltà (1998) using a different paradigm. In the original study, the stimulus array was displayed until response. This procedure allows individuals to process non-relevant digit identity in a later stage, after the enumeration process has been completed. In the present study, the stimulus was briefly presented (200 ms) and masked. This variation had the purpose to verify whether activation of magnitude representation of non-relevant digit identity was rapid and automatic or whether it required a longer stimulus duration to occur. Furthermore, as already pointed out, previous research has shown that short stimulus exposure influences the enumeration process. By using different sets of numerosities, it is possible to study the effect of symbolic distance on interference when individuals use different enumeration strategies. In particular, the use of a short stimulus exposure and large numerosities would allow us to study the effect of symbolic distance on Stroop interference when an estimation process is used.

A further goal of this research is to verify whether a surprising result of our previous study (Pavese and Umiltà, 1998) could be replicated when stimulus exposure is controlled. Pavese and Umiltà found that incongruent digits symbolically far from the enumeration response did not yield larger interference than letters. This finding would suggest that interference from incongruent digits found in previous studies (e.g., Francolini & Egeth, 1980) was only due to symbolically close incongruent digits. It is possible, however, that this result was a consequence of the long exposure times of the stimulus arrays, especially for larger numerosities.

Experiment 1

In the present experiments, we used a variation of Francolini and Egeth's (1980) enumeration paradigm. Participants were presented with circular arrays containing 18 items. Target items were red and the remaining items were presented in green; participants were required to vocally report the number of red items. Relevant red items could be digits, congruent or incongruent with the enumeration response. Differently from Francolini and Egeth, in the present experiments we used two incongruent conditions: (a) incongruent red items symbolically "close" to the correct response (distance $\pm$ 1); (b) incongruent red items symbolically "far" from the correct response (distance $\pm$ 3).

The enumeration process that participants used was assessed by analyzing RTs and error rate as a function of numerosity. We expected a small but reliable linear increase in the enumeration latency as a function of numerosity when participants used a subitizing process (Mandler & Shebo, 1982), whereas we did not expect a linear increase in RTs when participants used an estimation strategy (Kaufman et al., 1949). The effect of item identity was assessed by analyzing RTs and accuracy as a function of item identity. On the basis of Francolini and Egeth's (1980) results, incongruent digits were expected to produce longer RTs than either congruent or neutral digits. Finally, we expected to replicate Pavese and Umiltà's (1998) effect of symbolic distance on Stroop interference: Incongruent close digits should be enumerated more slowly than incongruent far digits.

Method

Participants. Thirty-three students at the University of Padua served as subjects. All had normal color vision and normal or corrected-to-normal visual acuity. Sixteen participants were assigned to the "1-3-4" group and 17 participants were assigned to the "2-4-5" group.

Design. The design included four conditions, depending on the identity of the items to be counted: (a) *Neutral*, in which the item to be counted were letters; (b) *Congruent*, in which the item to be counted were digits congruent with the correct enumeration response (e.g., four red "4s"[2]); (c) *Incongruent Close*, in which the item to be counted were digits incongruent with and symbolically close to the correct enumeration response (e.g., four red "3s"); (d) *Incongruent Far*, in which the item to be counted were digits incongruent with and symbolically far from the correct response (e.g., four red "1s").

The Neutral condition (red letters) provided a baseline from which to estimate the amount of interference from incongruent red digits. The crucial comparison is between Incongruent Close and Incongruent Far conditions. On the basis of previous results, we expected mean latency to Incongruent Close trials to be longer than the mean latency to Incongruent Far trials.

Two groups of numerosities were used. One group of participants responded to displays with one, two, or four targets (the "1-3-4" group) and a second group responded to displays with two, three, or five targets (the "2-4-5" group).

Although three different target numerosities were presented to each group, only the trials in which the correct response was "four", for group "1-3-4", and "five", for group "2-4-5" were analyzed for the effect of digit identity. All the other trials (those with response "one", "three" and "two", "four") were used as fillers. Therefore, the response to experimental trials were always "four" for one group and "five" for the other group, so that the variability due to phonological characteristics of vocal responses or to other sources was eliminated. The filler displays were employed for two reasons. First, in order to have a real enumeration task a choice between different responses was necessary. Second, several studies have shown that distractors drawn from the response set are more effective in producing interference than distractors from outside the response set (e.g., Flowers et al., 1979; Fox et al., 1971; La Heij, van der Heijden, & Schreuder, 1985; Morton, 1969; Proctor, 1978). This experiment was specifically designed so that all the digit identities were members of the response set. The design also provided that each numerosity appeared an equal number of times during the experimental session.

The experimental session consisted of 3 blocks of 72 trials (in total, 216 trials). Each block included 24 experimental trials and 48 filler trials. In total, each subject provided 72 experimental enumeration RTs, 18 in each of the 4 item identity conditions. There were 72 trials for each filler numerosity. These filler trials were always neutral – that is, the red items to be counted were always letters. Trial sequences were randomised across subjects and error trials were not replaced.

Apparatus and materials. The experiment was controlled by a 486 IBM-PC. The stimulus elements were generated by the PC running the Micro Experimental Laboratory software (MEL®; Schneider, 1988), and presented on a color VGA monitor (Nec Multi-Sync 3FG). The IBM VGA-Ultra package graphics mode was used. The display was a standard phosphorous display with a graphic resolution of 640 x 480. The PC also recorded vocal enumeration RTs, with an accuracy of ±1 ms, using a microphone connected to the PC through a response box. The identity of the vocal response given by the subject was entered manually by the experimenter at the end of each trial.

Stimuli were presented centrally and appeared as red (targets) and green (distractors) against a black background. The screen intensity was adjusted to an easy reading level and was maintained at that level throughout the experiment.

Each element was located at one of 18 equally spaced locations on the circumference of an imaginary circle (Francolini & Egeth, 1979, 1980). At the viewing distance of 120 cm, the center-to-center distance between the two diametrically opposed stimulus elements subtended a visual angle of approximately 3.2° (Francolini & Egeth, 1980). The mean visual angle between the edges of two adjacent elements was approximately 0.2°. Each item subtended a visual angle of approximately 0.35° in height and 0.33° in width. A MEL package graphic font was used (Romantri.FNT).

Red items were distributed randomly in the circumference. The only constraint on the composition of the stimulus array was that two red items never occupied adjacent array positions. The green items on each trial were the repetition of a randomly selected uppercase letter. The red items were either randomly selected uppercase letters or the digits 1, 2, 3, 4, or 5. Letters that are visually similar to digits (i.e., B, I, O, Q, and S) were excluded from the set of possible stimuli. Whenever more than one red item was present in an array, the same item was repeated. The same letter never occurred in an array both as a green and a red item.

Additional materials consisted of a white fixation cross in the center of the circle and a red and green pattern mask. The mask was an 18-item pattern, identical to the stimulus display, and each item consisted of a red H and a green $ overlapped.

Procedure. The experiment took place in a sound-attenuated and dimly lit room. The participants viewed the stimuli binocularly at a distance of 120 cm from the display. To ensure a fixed eye-to-screen distance, a head-and-chin rest was used. A microphone, connected to the PC, was located in front of the subject.

The procedure for each trial was the following: (a) a 400-ms warning tone (1000 Hz) was played, (b) a fixation cross was presented for 600 ms, (c) the stimulus display was presented for 200 ms (the center of the imaginary circle corresponded to the location of the fixation cross), (d) a 36-ms blank occurred (the time needed to create and present the pattern mask), (e) the pattern mask was presented until the response was emitted, and (f) 3.5 sec elapsed from the response to a trial to the onset of the next trial. Participants received visual feedback on latency and errors. Two different sounds signalled missed trials (latency longer than 2 sec) and error trials.

Participants began the session performing a simple naming task; digits from 1 to 9 were presented and the task was to read them aloud. This task allowed the experimenter to assess visual acuity and to adjust the vocalisation level. Participants then performed a practice session of 36 trials, followed by 3 experimental blocks. Participants were allowed to rest as long as desired between trial blocks. The instructions specified that the task was to verbally report the number of red items and that this number could only be 1, 3, or 4 (or 2, 4, 5). Participants were also warned that the green items and the identity of the red items were non-relevant; they were used to make the task more difficult and were to be ignored. The instructions stressed both speed and accuracy. The entire session lasted approximately 40 min.

Results

Error percentages were computed for each subject in each condition. Incorrect responses were omitted from RT analysis. Mean RTs for each subject were computed with a cut-off of two standard deviations; this method yielded the exclusion of 3.6% of the data. RTs and error percentages were analyzed in a Group by Numerosity mixed analysis of variance (ANOVA). A second mixed ANOVA, carried out on RTs and error rate to four-item display responses, assessed the

2 The items were enumerated in Italian: "uno", "tre" and "quattro", respectively, for 1, 3 and 4 red items and "due", "quattro", and "cinque", respectively, for 2, 4, and 5 red items.

effect of Group and Item Identity (Congruent, Neutral, Incongruent Close, Incongruent Far).

Effect of Numerosity

Average RTs and error rate as a function of numerosity are reported in Table 1. The RT analysis showed a significant effect of Numerosity, $F(2, 62) = 171.25$, $MSe = 686$, $p < .0001$. The linear contrast was significant, $F(1, 62) = 296.79$, $MSe = 203669$, $p < .0001$, indicating that response latency increased as a function of numerosity. Also the quadratic contrast was significant, $F(1, 62) = 45.71$, $MSe = 31367$, $p < .0001$, indicating that the increase in RT was not constant. The interaction Group by Numerosity was also significant, $F(2, 62) = 27.00$, $MSe = 686$, $p < .0001$. RT differences between numerosities were larger in the "2-4-5" group than in the "1-3-4" group. Similar results were obtained in the analysis of the error rate. The effect of Numerosity, $F(2, 62) = 31.40$, $MSe = 8.63$, $p < .0001$, and the linear contrast were significant, $F(1, 62) = 60.51$, $MSe = 522$, $p < .0001$, indicating that also error rate increased as a function of numerosity. The quadratic contrast did not reach significance in the error analysis, $F(1, 62) = 2.29$, $MSe = 20$, $p > .1$.

Effect of Item Identity

RT Analysis. Average RTs and error rate as a function of Item Identity are reported in Table 2. The main effect of Item Identity was highly significant, $F(3, 93) = 41.10$, $MS_e = 767$, $p < .0001$. Planned comparisons revealed that congruent trials were faster than neutral trials ($p < .0001$) and that the average of the two incongruent conditions was slower than the neutral condition ($p < .0005$). What is more important, mean RTs were longer in the Incongruent Close condition than in the Incongruent Far condition ($p < .005$). The difference between Incongruent Far and Neutral conditions did not reach significance ($p > .1$).

The effect of Group was significant, $F(1, 31) = 6.89$, $MSe = 24972$, $p < .05$, indicating that the "1-3-4" group was overall faster than the group "2-4-5" (491 and 564 ms, respectively).

Error Analysis. Overall, errors averaged 5.9%. The main effect of Item Identity was highly significant, $F(3, 93) = 17.41$, $MS_e = 42.55$, $p < .0001$. Planned comparisons revealed that congruent trials were more accurate than neutral trials ($p < .05$) and that the average of the two incongruent conditions was less accurate than the neutral condition ($p < .05$). What is more important, error rate was higher in the Incongruent Close condition than in the Incongruent Far condition ($p < .0001$). The difference between Incongruent Far and Neutral conditions did not reach significance ($p > .5$).

Discussion

Experiment 1 replicated the congruency effect found in Francolini and Egeth (1980). When compared to neutral letters, incongruent digits slowed the enumeration task, whereas congruent digits had a facilitatory effect. What is more important, not all incongruent digits produced the same amount of interference. Incongruent digits close to the enumeration response (arithmetic difference: -1) yielded greater interference than incongruent digits far from the enumeration response (arithmetic difference: -3). The Incongruent Close condition also showed a remarkably greater percentage of errors than the other three conditions.

These results replicate the effect of symbolic distance on Stroop interference found by Pavese and Umiltà (1998) and indicate that a long stimulus exposure is not a necessary condition to observe the effect of symbolic distance

Table 1 Mean reaction times in milliseconds and error rate in percentages as a function of Numerosity in Exps 1, 2 and 3. Standard deviations are reported in parentheses.

Numerosity	One	Two	Three	Four	Five	Seven
Experiment 1						
Group "1-3-4"						
Mean RT	424 (68.1)	—	480 (74.5)	491 (82.4)	—	—
Error Rate	0.2 (0.7)	—	2.0 (2.0)	4.3 (4.0)	—	—
Group "2-4-5"						
Mean RT	—	408 (45.2)	—	539 (67.2)	563 (75.8)	—
Error Rate	—	0.4 (0.7)	—	2.4 (2.0)	7.5 (5.8)	—
Experiment 2						
Mean RT	—	—	—	600 (82.8)	629 (101.2)	587 (115.9)
Error Rate	—	—	—	10.8 (5.9)	18.0 (11.8)	10.9 (14.2)

Table 2 Mean reaction times in milliseconds and error rate in percentages as a function of Item Identity in ExpS. 1, 2 and 3. Standard deviations are reported in parentheses.

Item Identity	Congruent	Neutral	Incongruent Close	Incongruent Far
Experiment 1				
Mean RT	487 (77.8)	527 (92.5)	561 (95.5)	538 (90.4)
Error Rate	1.0 (2.2)	5.6 (6.8)	12.4 (12.0)	4.7 (6.7)
Experiment 2				
Mean RT	558 (75.3)	583 (80.7)	627 (78.9)	604 (85.2)
Error Rate	4.9 (6.6)	7.6 (6.2)	20.1 (14.4)	9.7 (9.2)

on Stroop interference. A 200-ms stimulus exposure is sufficient to activate the magnitude representation that modulates interference effects. It is worth noting that, likely due to the short exposure and the masking, participants reported that they were not aware of the identity of the item to be enumerated.

The enumeration process that participants used in this experiment is characterized by a linear increase in RTs and error rate as a function of numerosity, consistent with a subitizing process (e.g., Mandler & Shebo, 1982). The trend analysis also revealed a significant quadratic trend in the RTs as a function of numerosity. This result is consistent with an "end effect" (Folk, Egeth & Kwak, 1988; Mandler & Shebo, 1982), often observed in enumeration tasks when the response set is limited to few values. In this experiment, it is possible that the highest and lowest values show an advantage because they had to be discriminated from a single other numerosity, whereas the central numerosities had to be discriminated from two other possible responses. However, two other factors may have contributed to the quadratic trend. First, the arithmetic distance among the three numerosities was not constant. Middle numerosities (three and four) have a distance of 2 units from the smaller numerosities (one and two) and a distance of 1 unit from the larger numerosities (four and five). Therefore, even if the increase in RTs and errors was a linear function of numerosity, a larger difference between smaller and middle numerosities than between middle and larger numerosities should be expected. Second, if we assume that numerical values are represented on a compressed number line and that this analog representation influences enumeration, we would expect to find a greater difference between smaller and middle numerosities than between middle and larger numerosities.

Experiment 2

In Exp. 1, incongruent digit identities were smaller than the relevant enumeration response. In Experiment 2 we used digit identities larger than the enumeration response. The experimental response was "four" and digit identities were 4, 5, and 7.

The aim of this experiment was to generalize the results found in Experiment 1 to larger numerosities in conditions of short stimulus exposure. Previous studies have shown that the set of numerosities used in the experiment can affect the strategy people use to enumerate stimulus arrays (Mandler & Shebo, 1982; Pavese & Umiltà, 1998). If individuals use estimation rather then subitizing in this experiment, RTs should not increase as a function of numerosity. Therefore, in this experiment we do not predict a linear increase in latency as a function of numerosity. We do expect a significant quadratic trend of numerosity, however. It is likely that, in condition of short exposure and with numerosities outside the subitizing range, the ease of discrimination of each numerosity will affect response latency and accuracy. In this experiment, the central numerosity five should be the most difficult to respond to, because it has to be discriminated from two other numerosities (four and seven), whereas the two "end" numerosities four and seven should be easier to discriminate.

Finally, we expect to find an effect of symbolic distance on Stroop interference, because this effect should depend on the activation of a magnitude representation associated with digits, which, according to our proposal, is independent of the enumeration process used to perform the task.

Method

Sixteen new students of the University of Padua served as subjects in this experiment. The method was identical to that of Experiment 1 except for few details. First, only one group of 3 numerosities was used (4, 5, and 7). Second, incongruent digits were greater than the correct response (the experimental response was "four" and the incongruent digit identities were 5 and 7[3]). Third, filler trials consisted of stimulus displays with five and seven red items.

Results

Mean RTs with a cut-off of two standard deviations (which led to the exclusion of 3.4% of the trials) and error percentage were analyzed in a one-way repeated measures ANOVA to assess the effect of Numerosity (four, five, and seven). A second one-way repeated measures ANOVA, carried out on RTs and error rate to four-item display trials, assessed the effect of Item Identity.

3 Also these items were enumerated in Italian: "quattro", "cinque", and "sette", respectively, for 4, 5, and 7 red items.

Effect of Numerosity

Average RTs and error rate as a function of numerosity are reported in Table 1. The RTs analysis showed a significant effect of Numerosity, $F(2, 30) = 6.36$, $MSe = 1391$, $p = .005$. Planned comparisons showed that five-item displays were enumerated slower than four- and seven-item displays ($p < .05$). The linear contrast was not significant, $F(1, 30) = 1.20$, MSe = 1391, $p > .25$, whereas the quadratic contrast was significant, $F(1, 30) = 11.51$, $MSe = 13325$, $p < .005$.

Similar results were obtained in the analysis of the error rate. The effect of Numerosity was significant, $F(2, 30) = 4.93$, $MSe = 54.66$, $p < .02$. Planned comparisons indicated that participants made more errors when they responded to five-item displays than in the other two numerosity conditions. The linear contrast was not significant, $F(1, 30) < 1$, $MSe = .06$, $p > .9$, whereas the quadratic contrast was significant, $F(1, 30) = 9.85$, $MSe = 539$, $p < .005$.

Effect of Item Identity

RT Analysis. Average RTs and error rate as a function of Item Identity are reported in Table 2. The main effect of Item Identity was highly significant, $F(3, 45) = 19.94$, $MS_e = 688$, $p < .0001$. Planned comparisons revealed that congruent trials were faster than neutral trials ($p < .02$) and that the average of the two incongruent conditions was slower than the neutral condition ($p < .0005$). RTs in the Incongruent Far condition were slower than RTs in the Neutral condition and faster than RTs in the Incongruent Close condition (ps < .05).

Error Analysis. Overall, errors averaged 10.8%. The main effect of Item Identity was highly significant, $F(3, 45) = 10.28$, $MS_e = 77$, $p < .0001$. Planned comparisons revealed that this source of variance derived its significance primarily from the higher error rate in the Incongruent Close condition, which was significantly different from the other three conditions ($p < .0001$).

Discussion

As predicted, Exp. 2 showed a different pattern of latency and error rate as a function of numerosity than Exp. 1. In the first experiment, we observed a linear increase of errors and RTs as a function of numerosity. In Experiment 2, the linear trend was significant neither in RTs nor in error rate. These results are consistent with the use of an estimation process, which is usually not associated with an increase of RTs as a function of numerosity (Mandler & Shebo, 1982). In Experiment 2, we replicated the sij>ificant quadratic trend of Numerosity. Both latencies and error rate indicate that responding to the central numerosity five was particularly difficult. This result suggests that discriminability among displays, and not numerosity, was the most important factor in determining task performance in this experiment. The central numerosity five was the slowest and the least accurate; faster latencies were associated with the larger numerosity (seven), which differed of two units from five, rather then with the smaller numerosity (four), which differed of one unit from the middle numerosity five.

The effect of item identity of Exp. 2, however, closely replicated the results of Experiment 1. Congruent trials were faster than neutral trials and incongruent trials were slower than neutral trials. In addition, symbolically close digits produced more interference than symbolically far digits. This finding confirms that the effect of symbolic distance on interference does not arise from a particular enumeration strategy, but rather from the activation of an abstract representation of magnitude.

An Omnibus Analysis

In order to compare the two experiments, two Experiment by Item Identity mixed ANOVAs were carried out on RTs and error percentages. A Tukey test was used to assess the significance level of the pairwise comparisons. The main effect of Experiment was significant, $F(2, 47) = 6.55$, $MS_e = 27667$, $p < .02$. Mean RTs were 528 and 593 ms, for Exps. 1 and 2, respectively.

The main effect of Item Identity was highly significant, $F(3, 141) = 51.92$, $MS_e = 748$, $p < .0001$. Mean RTs as a function of Item Identity (Congruent, Neutral, Incongruent Close, and Incongruent Far) are shown in Figure 1. The Tukey test indicated that all the differences among the four levels of Item Identity were significant ($p < .05$). The interaction Experiment by Condition was not significant, $F(6, 141) < 1$, indicating that the effect of symbolic distance did not reliably differ in the two experiments.

A similar analysis carried out on error rate also reealed a significant effect of Experiment, $F(2, 47) = 8.53$, $MS_e = 119.62$, $p < .01$. Error percentages were 5.9 and 10.8%, for Exps. 1 and 2, respectively.

The main effect of Item Identity was highly significant, $F(3, 141) = 27.18$, $MS_e = 54$, $p < .0001$. Error percentages as a function of Item Identity (Congruent, Neutral, Incongruent Close, and Incongruent Far) are shown in Figure 1. The Tukey test indicated that the Incongruent Close condition was less accurate than the Neutral and the Incongruent Far conditions ($p < .05$), and that the Congruent condition was more accurate than the Neutral and the Incongruent Far condition. The interaction Experiment by Condition did not reach significance, $F(6, 141) = 1.47$, $p > .20$.

The results of the omnibus analysis allow us to draw two interesting conclusions. First, even though there was a main effect of Experiment in both RTs and error rate, Item Identity did not interact with Experiment, supporting the hypothesis

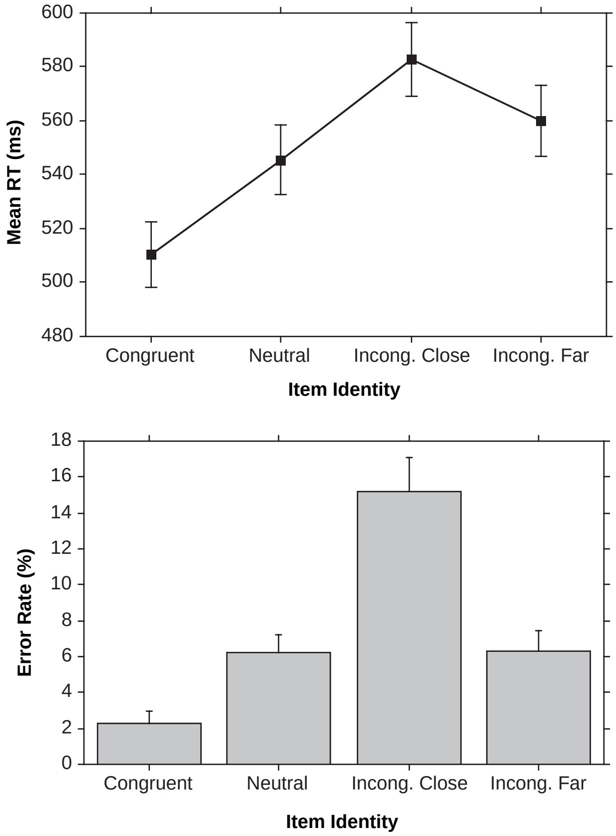

Fig. 1. Mean reaction times and error percentage collapsed across Experiments 1 and 2 as a function of item identity. Vertical bars indicate the standard error.

that the effect of identity does not depend on the set of numerosities or on the enumeration strategy. Second, this analysis revealed a significant difference between Incongruent Far and Neutral condition in RTs (but not in errors), suggesting that incongruent far digits do exert a certain level of interference on the enumeration task, at least in this paradigm in which the stimulus array was presented for only 200 ms.

These analyzes also allow us to discuss the suggestion that participants could have guessed the enumeration response of the experimental display using item identity information (Lana Trick, personal communication). Indeed, only in the experimental displays the items to be counted were sometimes digits, whereas in the two filler displays they were always letters. Thus, for example, in Exp. 2, participants could have responded "four" any time the red items were digits rather than letters, without using any enumeration process. We think that this hypothesis is not plausible, because the display was presented very briefly and masked; often participants spontaneously reported that they were not aware of the identity of the red items. However, it is important to verify whether the results of our experiments support this possibility. If participants guessed the enumeration response using digit identity information we should expect shorter RTs to experimental displays in which the items to be counted were digits rather than letters. This was not the pattern of results that we found. In Experiment 1, RTs to experimental displays were in fact longer than RTs to filler displays. In Exp. 2, the experimental four-item displays were responded to faster than the filler five-item displays, but not faster than the seven-item displays. Furthermore, in none of the experiments the neutral letter condition was slower than the incongruent digit conditions. Therefore, we can confidently conclude that, whatever process participants used to produce the correct response, they did not rely on item identity.

General Discussion

In two experiments, we replicated the effect of symbolic distance on Stroop interference in enumeration tasks found by Pavese and Umiltà (1998). Digits symbolically close to the enumeration response were enumerated significantly slower than digits symbolically far from the enumeration response. Three findings of this study are particularly relevant and will be discussed in detail. First, symbolic distance affected Stroop-like interference even when the stimulus was briefly presented and masked. Second, SDE was similar in the two experiments, even though the pattern of RTs and error rate as a function of numerosity suggests that the enumeration strategy that individuals used in Exp. 1 was different from the strategy used in Exp. 2. Third, we found a small but significant interference effect in the Incongruent Far condition in Experiment 2 and in the omnibus analysis, which is in contrast to the result reported by Pavese and Umiltà (1998) that enumerating digit symbolically far from the enumeration response did not produce more interference than enumerating letters.

The effect of stimulus presentation

An important difference between this study and Pavese and Umiltà's (1998) study was that here we used a short exposure time and a mask rather than presenting the stimulus until the response and without masking. When the stimulus is displayed until response, one cannot be certain that magnitude information is rapidly activated when a digit is presented, even when magnitude information is not relevant. In this condition, the possibility exists that magnitude information is activated only at a late stage of processing, after the enumeration task has been completed, and thus attention is free to shift to the non-relevant dimension. In contrast, when the stimulus is only briefly presented and masked, it is difficult to argue that the SDE depends on late processing of non-relevant information, rather than on a rapid and automatic activation of magnitude information from the non-relevant digit identity.

This result confirms Pavese and Umiltà's (1998) proposal that when a digit is presented, a rapid and automatic activation of magnitude representation occurs. This activation can, in turn, interfere with other magnitude-related tasks, such as enumeration. It is even possible that the rapid and automatic activation of magnitude representation can be *suppressed* with a longer stimulus presentation, as suggested by a result of Pavese and Umiltà's experiment. When participants were required to enumerate displays with more than six items, the neutral condition was slower than the incongruent conditions, suggesting that digit identity interfered less than letters at longer stimulus exposures.

The effect of enumeration strategy

The analysis of numerosity effects showed a different pattern of results in the two experiments. In Exp. 1, latencies and error rate increased as a function of numerosity, whereas in Exp. 2 latencies and error rate did not show a linear increase with numerosity, but rather seemed to depend on the ease of discrimination between adjacent numerosities. In Exp. 2, the central numerosity, five, which had to be discriminated from two adjacent numerosities (four and seven), was the slowest and the least accurate in the experiment, whereas the numerosities four and seven, which had to be discriminated only from five, were responded to faster and more accurately. This result suggests that the enumeration strategy used in Exp. 2 was different from the one used in Exp. 1. In the first experiment, participants probably used a subitizing strategy, as suggested by the small but reliable increase in RTs with numerosity. In Exp. 2, however, subitizing was not an adequate enumeration strategy because of the presence of larger numerosities. Therefore, it is likely that participants used a different strategy that took advantage of the small response set and used a process of numerosity discrimination.

Further evidence suggesting that subjects were using a different enumeration process in Exp. 2 comes from an examination of the overall error rate in the two experiments. The errors (across the three numerosities) were 2.3 and 13.2 %, respectively, for Exps. 1 and 2. Thus, enumeration in Exp. 2 was extremely less accurate than in Exp. 1, as one should expect if participants were using subitizing in the first experiment and estimation in the second.

A final consideration supports the hypothesis that two different enumeration processes were used in the two experiments. In both experiments, participants responded to displays with 4 red items. If we assume that (1) subitizing only depends on the current number of items to be enumerated, and not on the numerosity set tested, and (2) four is within the subitizing range (e.g., Trick & Pylyshyn, 1994), we should expect the numerosity four to be always enumerated in the same way—that is, subitized. More specifically, we should expect similar latencies and accuracy levels in the responses to four-item displays in the two experiments. Our results suggest that four-item displays were enumerated in a similar way by the two groups in Exp. 1, but in a different way in Exp. 2. In Exp. 1, enumeration responses were slightly longer in the "2-4-5" group than in the "1-3-4" group, but were also more accurate (see Table 1), suggesting that the presence of five-item display in the "2-4-5" group had the effect of increasing the accuracy level at the expense of speed, without a major change in enumeration strategy. In Exp. 2, however, both speed and accuracy of four-item display responses were much lower than in Exp. 1. Four-item displays were enumerated slower in Exp. 2 than in Exp. 1 (600 and 516 ms, respectively; $t[47] = -3.48$, $p = .001$) and were also less accurate (the error rate was 10.8 and 3.3%, respectively; $t[47] = -5.78$, $p < .0001$).

Even though the results strongly suggest that different enumeration strategies were used in the two experiments, we did not find any reliable difference in the SDE. This result confirms that the SDE depends neither on a particular experimental setting nor on a particular processing strategy to perform the enumeration task. More generally, this finding is consistent with a number of studies that investigated the SDE in a variety of paradigms. As already mentioned, the SDE has been found originally in comparison tasks, but also in naming tasks and same-different judgements. Therefore, SDE seems to be a property of numerical stimuli rather than of a particular task (Dehaene et al., 1993).

Do incongruent digits interfere with enumeration?

These experiments provide some evidence that incongruent digits that are symbolically far from the enumeration response produce an interference effect greater than that produced by letters. Pavese and Umiltà (1998) did not find greater interference from incongruent far digits than from letters. In that experiment, however, variation in interference patterns might have been caused by the increase of exposure time as a function of numerosity. The authors reported that the neutral condition tended to become increasingly slower, relatively to the other conditions, with larger numerosities.

In the present experiments the exposure time was fixed at 200 ms and the neutral condition proved to be a reliable baseline. In the two experiments, the usual congruency effect was found: The neutral condition was slower than the congruent condition and faster than the incongruent condition. Interference from incongruent far digits was found in Exp. 2 and in the omnibus analysis, suggesting that a residual interference effect may still be present in digits that are far from the enumeration response, at least for brief exposure times.

Conclusions

This study replicates and generalises the results found in Pavese and Umiltà (in press) to a paradigm in which the stimulus is only briefly presented and masked. In an enumeration task, interference from digit identity is modulated by the symbolic distance between numerosity and identity. Symbolically close digits produce more interference than symbolically far digits. The results of this study suggest two conclusions: (1) the activation of magnitude representation is rapid and automatic and (2) the SDE effect is independent of enumeration strategy.

Aknowledgements This study was carried out when the first author was at the Department of Psychology of the University of Padua. We want to thank Barbara Montanari and Anna Dallavedova for help in data collection. The authors were supported by grants from the Consiglio Nazionale delle Ricerche and from the Ministero della Università e della Ricerca Scientifica e Tecnologica.

References

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.

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.

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

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.

Fias, W., Brysbaert, M., Geypens, F., & d'Ydewalle, G. (1996). The importance of magnitude information in numerical processing: Evidence from the SNARC effect. *Mathematical Cognition, 2,* 95-110.

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.

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. (1979). Perceptual selectivity is task dependent: The pop-out effect poops out. *Perception & Psychophysics, 25*, 99-110.

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

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 color-naming. *American Journal of Psychology, 77,* 576-588.

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

La Heij, W., van der Heijden, A. H. C., & Schreuder, R. (1985). Semantic priming and Stroop-like interference in word-naming tasks. *Journal of Experimental psychology: Human Perception and Performance, 11*, 62-80.

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.

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., & 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.

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

Pavese, A., & Umiltà, C. (1998). Symbolic distance between numerosity and identity modulates Stroop interference. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 1535-1545.

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

Schneider, W. (1988). Micro-Experimental Laboratory: An integrated system for IBM PC compatibles. *Behavior Research Methods, Instruments, and Computers, 20*, 84-99.

Shepard, R. N., Kilpatric, D. W., & Cunnigham, J. P. (1975). The internal representation of numbers. *Cognitive Psychology, 7*, 82-138.

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.

Trick, L. M., & Pylyshyn, Z. W. (1993). What enumeration studies can show us 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.

Washburn, D. A. (1994). Stroop-like effects for monkeys and humans: Processing speed or strength of association? *Psychological Science, 5*, 375-379.

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.