

# Size congruity effects with two-digit numbers: Expanding the number line?

DANIEL FITOUSHI and DANIEL ALGOM  
Tel Aviv University, Tel Aviv, Israel

The *size congruity effect* is the interesting result that comparisons of the sizes of the physical formats in which numerals appear are affected by the numerical magnitudes of the respective numerals. We demonstrated that separating the physical and the numerical attributes in space leaves the effect unchanged. We then applied the spatially separated version to two-digit numerals and showed the effect to be comparable to that obtained with single numerals. We showed further that the effect is sensitive to the relative salience of the numeric and physical dimensions, to the extent that when the latter is the more salient dimension, a *reverse* effect obtains by which physical size interferes with number comparison. The results can be explained by a relative speed of processing account, but they are also compatible with an attention account that does not appeal to the notion of automaticity.

Is 7 greater than 5? This appealingly simple question prompts a rapid and accurate answer with virtually all people. The ease with which people respond conceals the fact that complex cognitive operations are involved. To respond, people must retrieve the underlying magnitudes; note that, as mere graphic signs, 7 is neither larger nor smaller than 5. This fast (automatic?) retrieval of numerical information is particularly impressive when numerical values are irrelevant to the task at hand and can actually hurt performance. Consider a task in which pairs of digits are compared with respect to *physical size*. This task focuses on a nonsemantic attribute of the digits, yet the now irrelevant numerical values do, nonetheless, facilitate (in congruent pairs, such as 8–6) or impair (in incongruent pairs, such as 8–6) performance. This size congruity effect (SCE) has evolved into an important marker of the potency of numerical activation (Algom, Dekel, & Pansky, 1996; Besner & Coltheart, 1979; Foltz, Poltrock, & Potts, 1984; Hatta, 1977; Henik & Tzelgov, 1982; Pansky & Algom, 1999, 2002; Schwarz & Heinze, 1998; Schwarz & Ischebeck, 2003; Takahashi & Green, 1983; Tzelgov, Meyer, & Henik, 1992; Vaid, 1985; Vaid & Corina, 1989).

Despite the importance of the SCE, studies probing the phenomenon have been confined to date to single-digit numbers. Application to two-digit numbers has been thwarted, due to methodological problems. This is unfortunate, because the similarity (or lack thereof) in the pro-

cessing of single- and two-digit numbers is consequential theoretically, yet relevant evidence is missing, due to the paucity of studies with two-digit numbers. Reynvoet and Brysbaert (1999) have noted that “although a lot of research on number representations has been done . . . there is still little or no direct evidence whether or not units and teens make access to a single number line. Most researchers avoid the problem by using unit numbers only” (p. 192). In this study, we bypassed the methodological problems and probed the SCE with single- and two-digit numbers within a common design. The results contribute to an understanding of two-digit number processing and bear on the general question of the uniformity of access to the number line across different orders of magnitude.

Are single- and two-digit numbers governed by the same rules of processing and common mechanisms? Available evidence is equivocal. Take the well-known *distance effect* (Moyer & Landauer, 1967) as a handy example. The effect documents the finding that the time taken to decide which of a pair of numerals is larger numerically is inversely related to the numerical distance separating the members of the pair. The overwhelming majority of studies concerned single digits, with the effect invariably governing performance with such digits. In contrast, the results with two-digit numbers have been variable. Brysbaert (1995) recorded the typical form of the distance effect only when the two-digit numbers were confined to the same decade. Dehaene and Akhavein (1995) did not detect a distance effect with a *same-different* judgment task (although they did find the effect with single digits, using the same task). Although Dehaene, Dupoux, and Mehler (1990) and Hinrichs, Yurko, and Hu (1981) did find full-blown distance effects with two-digit numbers, large discontinuities (across decades) characterized the sets of data. Thus, comparing 50 with the standard of 55 took disproportionately longer than comparing 49 with the same standard (despite the fact that 49 and 50 are close in magnitude). Noticeable increases in reaction time (RT)

---

Portions of this research were presented at the 19th Annual Meeting of the International Society for Psychophysics, Larnaca, Cyprus, and at the 44th Annual Meeting of the Psychonomic Society, Vancouver. The research was based partly on a doctoral dissertation at the Department of Psychology, Tel Aviv University, by D.F., under the supervision of D.A. We thank Marc Brysbaert and two anonymous reviewers for their helpful comments on an earlier version of this article. Correspondence concerning this article should be addressed to D. Fitousi, Department of Psychology, Tel Aviv University, Ramat Aviv 69978, Israel (e-mail: dannyfit@post.tau.ac.il).

were recorded with numbers ending with zero, too (the zero unit effect). Thus, comparing 40 with the standard of 55 was as fast as comparing 45 with the same standard, despite the difference in numerical distance.

A strictly monotonic distance effect with two-digit numbers requires that each two-digit number has a unique *holistic* representation (i.e., the component unit and teen digits coalesce into an integral, unitary representation). Each single-digit number is likely represented by such a unique mental referent. Recent research by Fias, Reynvoet, and Brysbaert (2001), Damian (2004), and Ischebeck (2003) strongly suggests that single (Arabic) digits are processed in a holistic fashion, much like pictures. This means that each individual digit is associated with a unique representation and this representation is typically accessed by a semantic route in the same way pictures are. It is moot whether each two-digit number is associated with a unique representation in tandem with a semantic mode of processing. Damian surmised that “the . . . similarity between digits and pictorial stimuli probably breaks down outside the realm of single-digit numbers” (p. 170).

Violations of strict monotonicity with two-digit numbers support Damian’s (2004) conjecture about differential processing of single- and two-digit numbers. Nuerk, Weger, and Willmes (2001) reported such violations in their study of comparative judgments with two-digit numerals. In their study, the pair 42–58 was compatible because  $4 < 5$  and  $2 < 8$ , whereas the pair 29–63 was incompatible because  $2 < 6$  but  $9 > 3$ . Decision time decreased and accuracy improved when the decades and the units were compatible. The presence of compatibility effects implies lexical and combinatorial operations—that is, separate processing of teens and units. Nuerk et al. concluded that a “single number line for one- and two-digit numbers . . . cannot account” for their results (p. 33, emphasis added). Similar conclusions were reached by Fias, Lammertyn, Reynvoet, Dupont, and Orban (2003) in another study of numerical comparison with two-digit numbers. The authors concluded that their set of results “provides behavioral demonstration that two-digit numbers are *not* processed exclusively holistically as has been generally assumed” (p. 53, emphasis added), in the way that has been demonstrated with single-digit numbers.

In yet another contrast, Brysbaert (1995) produced data that do suggest the presence of unique representations for each individual number in the 1–99 range. Brysbaert found that viewing durations for visually presented numerals were logarithmic functions of their magnitudes. Brysbaert concluded that “at least for the integers from 1 to 99, the representation is [a] number line on which each integer has its unique representation” (p. 451).

In summary, research with single- and two-digit numbers has revealed potentially conflicting features, partly compatible with a uniform pattern of processing, partly incompatible with a common pattern of processing. In this study, we attempted to disambiguate the modes of processing by probing one of the most popular effects in the literature on numerical cognition, the SCE. If two-digit numbers are susceptible to the SCE in the same way that single-digit

numbers are, strong evidence will have been accrued supporting the uniformity of the *number line* concept.

We also tested the sensitivity of SCE to a major factor of context, *dimensional discriminability*, already shown to affect processing with single digits (Algom et al., 1996; Pansky & Algom, 1999, 2002; Schwarz & Ischebeck, 2003). Discriminability specifies the size of the psychological differences separating stimulus values along a dimension. It is matched when the numerical values are as readily discriminable from one another as are the values of physical size. Algom et al. found that numerical magnitude interfered with judgments of physical size (the SCE) when the numerical dimension was more discriminable than the physical dimension. When physical size and numerical magnitude were matched on salience, the interference largely disappeared (no SCE was recorded). Strikingly, when physical size was made more salient than numerical size, a reversed pattern of interference emerged by which numerical judgments were disrupted by physical size (reverse SCE, or RSCE). Similar results were obtained by Pansky and Algom and by Schwarz and Ischebeck. In the present study, we examined the influence of relative dimensional salience on the processing of two-digit numbers.

Why have studies of the SCE considered only single-digit numbers? We suspect that two-digit numbers have been avoided due to the difficulty of engineering the Stroop conflict (Stroop, 1935) with two-digit numbers. How should one define physical size with a two-digit number? Should one present the entire 53 in larger font than the entire 48? If so, a holistic representation is already assumed. Moreover, accessing such a representation is complicated by the fact that the same numerals appear in different physical sizes in various two-digit numbers (e.g., 8 would differ in size in 18 and 81, as would 1). The variation in size of the composite digits likely affects performance and renders the interpretation of any ensuing SCE difficult. Should one assign different physical sizes to the composite digits (e.g., 53)? This format virtually dictates discrete, combinatorial processing. Physical size as a unitary attribute might be compromised. It is difficult to see how one can effect an SCE with such stimuli. Given the problems, the Stroop task has not been used to probe the nature of numerical processing with two-digit numbers. Bypassing these problems, we tested comparisons with two-digit numbers in this study. To foreshadow the outcome, the present results support the uniformity of numerical representations, expanding the applicability of the *number line* metaphor to the realm of two-digit numbers.

We applied the SCE to two-digit numbers by using the spatially separated version of the Stroop task (Dishon-Berkovits & Algom, 2000; Glaser & Glaser, 1989; Kahneman & Chajczyk, 1983; MacLeod & Hodder, 1998). In the original version of the Stroop test, word and color (or numerical value and physical size) are integrated into the same stimulus. In the spatially separated version, in contrast, word and color occupy different positions in space. In Kahneman and Chajczyk’s study, a color word (in black) was presented above or beneath a color patch. The

color word could name the color of the patch (congruent combination) or could name a color other than that of the patch (incongruent combination). The task was that used in the standard Stroop arrangement: to name the color of the patch speedily and accurately. In this study, we applied the separated version of the Stroop test in the number domain. The dimension of number was naturally defined by the pertinent numerical value, whereas the dimension of physical size was defined by the length of a horizontal bar drawn below the numeral. This method is applicable in the same way to the study of single- and multidigit numbers, bypassing the aforementioned methodological problems.

### Overview of the Experiments

The goal of the first pair of experiments was to validate the separate version of the numerical Stroop task. In Experiment 1, we presented the traditional, spatially integrated version of the Stroop task, with the numerals varying in numerical magnitude, as well as in physical size. The participants compared pairs of numerals with respect to physical size in one session and with respect to numerical magnitude in another session. In Experiment 2, we used the same values of numerical magnitude and physical size in the spatially separated version of the Stroop task. Numerical magnitude was represented, as before, by numerical value. However, all of the numerals appeared in the same size now. Physical size was represented by the length of the bar appearing below each numeral. The participants compared numerical magnitude in one session and the physical size of the bars in another session. Comparable performance across the two experiments would confer validation support for the spatially separated version of the Stroop task in the numerical domain.

In the remaining four experiments, we used the spatially separated version of the comparison task with single- and with two-digit numerals. In these experiments, we manipulated relative dimensional discriminability. In two experiments (Experiments 3 and 5), we made the numerical dimension more salient than physical size; in the other pair of experiments (Experiments 4 and 6), we reversed relative salience and rendered physical size more discriminable than numerical magnitude. Comparable sensitivity in the ranges of units and tens to relative dimensional salience would further support the uniformity of numerical processing in the 1–99 range. Looked at from an even broader perspective, the malleability of the SCE in response to relative dimensional salience can serve to elucidate the nature of numerical processing, an issue that we will discuss after the presentation of the empirical results.

## EXPERIMENT 1

The purpose of the first experiment was to reproduce the results typically obtained with single digits in the numerical Stroop paradigm (e.g., Henik & Tzelgov, 1982; Pansky & Algom, 1999, 2002; Schwarz & Ischebeck, 2003). Presented with pairs of numerals varying in physical size, as well as in numerical magnitude, the participants compared the numerals on physical size in one block

and on numerical magnitude in another block. To replicate previous results, we purposely made the number dimension more salient than the size dimension.<sup>1</sup> As a result, we expected the routine SCE to obtain in this experiment.

### Method

**Participants.** Twelve Tel Aviv University undergraduates participated in partial fulfillment of course requirements.

**Stimuli and Apparatus.** For numerical magnitude, we used the following six digits: 2, 3, 4, 6, 7, and 8. For physical size, we used the following six values:  $4.5 \times 3$  mm (a),  $5 \times 3.5$  mm (b),  $5.25 \times 3$  mm (c),  $5.5 \times 3.5$  mm (d),  $6 \times 4.5$  mm (e), and  $7 \times 7$  mm (f). The numbers were associated with the respective sizes, so that 2 appeared in size a, 3 in size b, 4 in size c, 6 in size d, 7 in size e, and 8 in size f. Of the possible pairs, the following six were presented: 2–6, 2–7, 2–8, 3–7, 3–8, and 4–8. Each pair appeared in congruent and incongruent combinations. For the congruent combination of the pair 2–6, 2 appeared in physical size a and 6 in physical size d. For the incongruent combination of the same pair, the sizes were reversed, so that 2 now appeared in physical size d and 6 in physical size a. The same manipulation was used repeatedly to create the congruent and incongruent combinations for all the other pairs. For a given pair in a given combination of number and size, each member appeared once in the left and once in the right position. Therefore, each pair was presented in four different arrangements in all. Presenting each individual arrangement four times resulted in a block of 96 trials. Order of stimulus presentation was random.

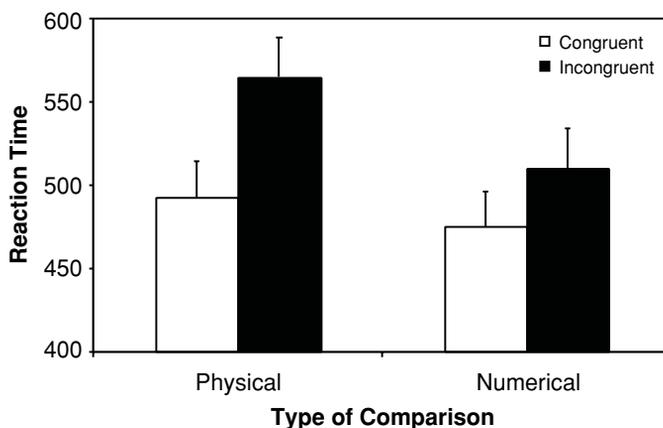
The stimuli were generated by an IBM-compatible microcomputer (PC 486) and were displayed on a  $600 \times 800$  pixel VGA color monitor. The numbers appeared black over a white background at the centers of the left and right hemifields. They were separated by approximately 15 mm. To avoid adaptation, we introduced a trial-to-trial spatial uncertainty of up to 15 pixels around the target locations. The viewing distance was 60 cm, so that the stimuli appeared at  $0.71^\circ$  of visual angle to the right and left of fixation.

**Procedure.** The participants were tested individually in a dimly lit room. Each individual performed in the numerical and the physical comparison tasks. Half of the participants first performed in the physical task, and half first performed in the numerical task. The participants were instructed to attend to the relevant dimension (physical size or numerical magnitude) and to select the larger member of the pair on that dimension. They were instructed to press either a right- or a left-hand key on the computer keyboard, according to the location of the larger stimulus. Twenty practice trials preceded the experimental trials in each task (unbeknownst to the participant). The stimuli were response terminated. A new pair of stimuli was presented following a 500-msec interval. Trials in which the participant's RT was either greater than 1,000 msec or smaller than 100 msec were excluded from the analyses.

### Results

Figure 1 shows the mean RTs for congruent and incongruent pairs in the two tasks. First, consider the overall RTs. The respective means were 492 msec for the numerical comparisons and 529 msec for the physical comparisons. The 37-msec advantage favoring number [ $F(1, 11) = 5.46$ ,  $MS_e = 2,988$ ,  $p < .05$ ] documents the success of our manipulation that made the comparison of numbers easier than that of physical sizes. Error rates for both tasks were relatively low: 1.13% for the physical comparison and 1.05% for the numerical comparison ( $F < 1$ ).

Next, consider the quality of selective attention in each task. Recall that if performance is free of intrusions from the task-irrelevant dimension, no Stroop effects (SCE or RSCE) should be observed. When the participants com-



**Figure 1. Experiment 1: mean reaction times (in milliseconds) for comparisons of numerical magnitudes and physical sizes of numerals. Data are averaged across congruent or incongruent combinations of numerical magnitude and size. Bars depict one standard error of the mean.**

pared the digits with respect to physical size, they suffered intrusions from the irrelevant numerical magnitudes. Responses on the congruent trials (mean of 493 msec) were faster than those on the incongruent trials (mean of 565 msec), the difference amounting to an appreciable SCE of 72 msec [ $F(1,11) = 48.17, MS_e = 653, p < .005$ ]. Error rates averaged 0.35% for the congruent pairs and 1.91% for the incongruent pairs, producing an SCE for errors, too [ $F(1,11) = 7.3, MS_e = 2.01, p < .05$ ].

For comparisons of numerical magnitude, performance was affected by the irrelevant dimension of physical size. The congruent pairs (mean of 475 msec) were compared more quickly than the incongruent pairs (mean of 510 msec), the difference amounting to an RSCE of 35 msec [ $F(1,11) = 30.5, MS_e = 238.4, p < .00$ ]. Error rates were 0.35% for the congruent pairs and 1.75% for the incongruent pairs [ $F(1,11) = 4, MS_e = 2.94, p = .07$ ].

In an overall ANOVA on RT, the main effect of congruity [ $F(1,11) = 53.9, MS_e = 640, p < .00$ ] confirmed the presence of Stroop effects, routine and reversed. The interaction of task and congruity [ $F(1,11) = 16.85, MS_e = 251, p < .05$ ] confirmed that the SCE was larger than the RSCE. In the parallel ANOVA on error, only the congruity effect was found reliable [ $F(1,11) = 5.92, MS_e = 4.45, p < .05$ ].

## Discussion

The typical result documented in the literature obtained: Comparisons of size were affected by irrelevant number to a greater extent than were comparisons of number by irrelevant physical size. The asymmetry in interference was planned in advance by judiciously making numbers more salient than physical sizes. We submit that this asymmetry is neither fixed nor dictated by theory. As we will subsequently show (Experiments 4 and 6), making physical size more salient than numerical magnitude reverses the interference, yielding a pattern in which the RSCE is

larger than the SCE. Before pursuing this issue further, we first sought to reproduce the standard observations in Experiment 1 with the spatially separated version of the numerical Stroop task.

## EXPERIMENT 2

In Experiment 2, we used the *same* values as those in Experiment 1, but in a spatially separated preparation. The numbers appeared in a constant physical size. The values of physical size were specified by the length of the line drawn beneath each digit. The purpose of the second experiment was to reproduce the results of Experiment 1 with such spatially separated components (Kahneman & Chajczyk, 1983). The participants compared the pairs of stimuli with respect to numerical magnitude in one block and with respect to line length in another block.

## Method

**Participants.** Eighteen Tel Aviv University undergraduates who did not participate in Experiment 1 took part in the experiment in partial fulfillment of course requirements.

**Stimuli and Apparatus.** For numerical value, we again used the digits 2, 3, 4, 6, 7, and 8. For line length, we used the six values of 4.5 mm (a), 5 mm (b), 5.25 mm (c), 5.5 mm (d), 6 mm (e), and 7 mm (f). Note that these are the values of digit height used in Experiment 1. The numbers were associated with the respective line lengths, so that 2 appeared along with line a, 3 with line b, 4 with line c, 6 with line d, 7 with line e, and 8 with line f. The digits were presented in the *constant* physical size of  $6.5 \times 4$  mm. Each line was 1 mm in width, printed in black. It was centered 1.5 mm beneath its corresponding digit. The following six pairs were again presented: 2–6, 2–7, 2–8, 3–7, 3–8, and 4–8. Each pair appeared in a congruent and an incongruent combination. For the congruent combination of the pair 2–6, 2 appeared underlined by line a and 6 by line d. For the incongruent combination, the lines were changed, so that 2 now appeared underlined by line d and 6 underlined by line a. The same manipulation was used to create congruent and incongruent combinations for all the other pairs. All the other details of the stimuli and presentation replicated those in Experiment 1.

**Procedure.** The procedure was the same as that in Experiment 1.

## Results

Mean RTs for the congruent and incongruent pairs in the two tasks are represented in Figure 2. First, consider the overall RTs. The respective means were 485 msec for the comparisons of numerical magnitude and 632 msec for the comparisons of line length. The 147-msec advantage favoring number [ $F(1,17) = 39.3$ ,  $MS_e = 9,828$ ,  $p < .00$ ] reflects the superior salience of numerical values over line length values. Error rates in both tasks were relatively low: 1.21% for the physical comparisons and 1.28% for the numerical comparisons ( $F < 1$ ).

The comparisons of physical size (line length) were affected by irrelevant numerical magnitude. Responses on the congruent trials (mean of 610 msec) were faster than those on the incongruent trials (mean of 654 msec), the difference amounting to an SCE of 44 msec [ $F(1,17) = 7.5$ ,  $MS_e = 3.61$ ,  $p < .05$ ]. Error rates averaged 0.35% for the congruent pairs and 2.08% for the incongruent pairs [ $F(1,17) = 7.51$ ,  $MS_e = 3.61$ ,  $p < .05$ ]. In contrast, comparisons of numerical magnitude were almost free of intrusions from the irrelevant dimension of line length. The congruent pairs were compared a bit more quickly than the incongruent pairs (means of 481 and 490 msec), but the difference was not reliable [ $F(1,17) = 2.07$ ,  $MS_e = 345.7$ ,  $p = .16$ ]. Error rates were 0.9% for the congruent pairs and 1.62% for the incongruent pairs [ $F(1,17) = 1.32$ ,  $MS_e = 3.38$ ,  $p = .3$ ].

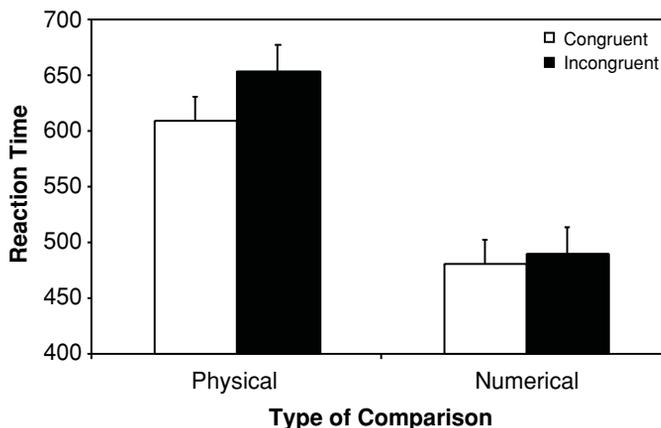
In an overall ANOVA on RT, the main effect of congruency [ $F(1,17) = 38.79$ ,  $MS_e = 327$ ,  $p < .00$ ] confirmed the presence of an SCE. The interaction of task and congruency [ $F(1,17) = 26.7$ ,  $MS_e = 209$ ,  $p < .00$ ] confirmed that the SCE was larger than the RSCE (which was not reliable). In a parallel ANOVA on error, a congruency effect [ $F(1,17) = 4.24$ ,  $MS_e = 6.27$ ,  $p = .054$ ] and task  $\times$  congruency interaction [ $F(1,17) = 4.16$ ,  $MS_e = 1.17$ ,  $p = .057$ ] replicated those observed for RT.

## Discussion

Comparisons of pairs of lines aligned spatially with pairs of numerals were affected by the numerical value of those numerals. Comparisons of the values of the numerals themselves were relatively free of intrusions from the corresponding lines. This pattern replicated that observed in Experiment 1, as well as those obtained in the voluminous literature on numerical Stroop effects. The pertinent studies involved only spatially integrated stimuli (i.e., various numbers presented in various physical sizes). The novel feature of Experiment 2 was the demonstration that the SCE obtains in a spatially separated preparation in which number and size appear in different locations in space.

In an ANOVA performed on the collective data from Experiments 1 and 2, numerical comparisons were performed more quickly than physical ones [ $F(1,28) = 33.96$ ,  $MS_e = 7,132$ ,  $p < .005$ ] and responses on congruent trials were faster than those on incongruent ones [ $F(1,28) = 102.36$ ,  $MS_e = 452.83$ ,  $p < .00$ ]. The absence of an experiment  $\times$  task  $\times$  congruency interaction ( $F < 1$ ) confirmed that the SCE was larger than the RSCE in both preparations. Therefore, separating the representations of physical size and numerical value in space left the pattern of interference invariant. Given the values used for number and size in Experiments 1 and 2 (more salient for the former dimension), comparisons of size were affected by irrelevant number more than vice versa, regardless of the way in which the dimensions were depicted.

As we recounted, the spatially integrated version is impractical for examining processing with two-digit numbers, because the representation of size is ambiguous. By administering the spatially separated version, in contrast, one can apply the Stroop task unchanged across single- and two-digit numerals. In the remaining four experiments, we used the new preparation to study comparative processing with two-digit numerals. In the great bulk of existing studies, the values used for number were more



**Figure 2.** Experiment 2: mean reaction times (in milliseconds) for comparisons of numerical magnitudes and values of line length. Data are averaged across congruent or incongruent combinations of numerical magnitude and length. Bars depict one standard error of the mean.

salient than the values used for physical size. In fact, many studies pitted nine values for number (the digits between 1 and 9) against merely two values for size (small or large). If the source of the asymmetry in interference is mismatched discriminability, then the pattern of interference is not fixed but is reversible. This prediction was tested in the following experiments.

### EXPERIMENT 3

In this experiment, we presented pairs of two-digit numbers, each of which was underlined with a corresponding bar. As in Experiments 1 and 2, the numerical dimension was made more salient than the physical dimension. Therefore, comparisons of lines were expected to be affected by the irrelevant numbers (SCE) to a larger degree than vice versa (RSCE).

#### Method

**Participants.** Sixteen Tel Aviv University undergraduates, who did not participate in any of the previous experiments, performed in partial fulfillment of course requirements.

**Stimuli and Apparatus.** For numerical value, we used the two-digit numbers 17, 33, 51, 65, 84, and 98. For line length, we used the values of 4.75 mm (a), 5.75 mm (b), 6.75 mm (c), 8 mm (d), 9 mm (e), and 10.75 mm (f). Expressed in pixels, the respective lengths were 12, 15, 18, 22, 25, and 30. The digits appeared in the constant physical size of  $5 \times 7.5$  mm. The lines presented beneath were 1 mm in width. They appeared black over a gray background. Each line was centered 1 mm beneath its corresponding two-digit number. The following 10 pairs were presented: 17–33, 33–51, 51–65, 65–84, 84–98, 17–51, 51–84, 65–98, 33–84, and 17–98. The numbers were associated with line lengths, such that 17 appeared along with line a, 33 with line b, 51 with line c, 65 with line d, 84 with line e, and 98 with line f. Each pair appeared in congruent and incongruent combinations. For the congruent combination of the pair 17–98, 17 appeared underlined by line a, and 98 by line f. For the incongruent combination, the lines were reversed, so that 17 now appeared underlined by line f and 98 by line a. The same manipulation was used to create congruent and incongruent combinations for all the other pairs. For a given pair in a given combination of number and size,

each member appeared once in the left and once in the right position. Therefore, each pair was presented in four different arrangements in all. Presenting each individual arrangement four times resulted in a block of 160 trials. Order of stimulus presentation was random.

**Procedure.** The procedure was the same as that in Experiment 2.

#### Results

In Figure 3, we present the mean RTs for the congruent and incongruent pairs in the two tasks. For the overall RTs, the respective means were 592 msec for the comparisons of numerical magnitude and 668 msec for the comparisons of line length. The 76-msec advantage favoring the former [ $F(1,15) = 19.79$ ,  $MS_e = 4,662$ ,  $p < .005$ ] reflects the better discriminability of number. Error rates were 3.7% for the numerical comparisons and 4.2% for the physical comparisons ( $F < 1$ ).

For the comparisons of physical size (line length), the judgments were influenced by irrelevant numerical magnitude. Responses to the congruent trials (mean of 651 msec) were faster than those to the incongruent trials (mean of 686 msec), the difference amounting to an SCE of 35 msec [ $F(1,15) = 50.3$ ,  $MS_e = 193.162$ ,  $p < .00$ ]. Error rates averaged 3.28% for the congruent pairs and 5.15% for the incongruent pairs [ $F(1,15) = 6.27$ ,  $MS_e = 4.47$ ,  $p < .05$ ], producing a reliable SCE for error, too. In contrast, the comparisons of numerical magnitude were almost free of interference from the now irrelevant line. The congruent pairs were compared slightly more quickly than the incongruent pairs (means of 590 and 594 msec); the small difference of 4 msec was not reliable [ $F(1,15) = 1.69$ ,  $MS_e = 60.6$ ,  $p = .2$ ]. Error rates were 3.28% for the congruent pairs and 4.29% for the incongruent pairs [ $F(1,15) = 3.56$ ,  $MS_e = 2.31$ ,  $p = .07$ ].

In an overall ANOVA on the RTs, the effect for congruity [ $F(1,15) = 36.39$ ,  $MS_e = 162$ ,  $p < .0005$ ] confirmed the presence of an SCE. The interaction of task and congruity [ $F(1,15) = 42.81$ ,  $MS_e = 91$ ,  $p < .00$ ] confirmed that the SCE was larger than the RSCE. In a parallel ANOVA on

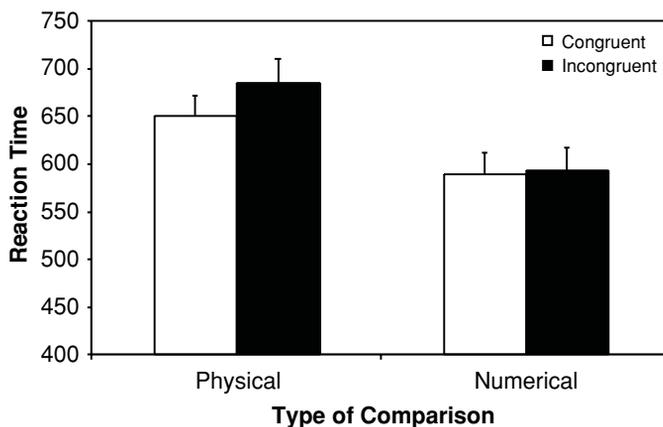


Figure 3. Experiment 3: mean reaction times (in milliseconds) for comparisons of numerical magnitudes and values of line length. Data are averaged across congruent or incongruent combinations of numerical magnitude and length. Bars depict one standard error of the mean.

errors, a congruity effect [ $F(1,15) = 9.61$ ,  $MS_e = 3.47$ ,  $p < .01$ ] mirrored that observed for RT.

## Discussion

Comparisons of pairs of lines were disrupted by the incongruent numerical magnitudes of two-digit numbers positioned above the lines. Comparisons of the numbers themselves were free of intrusions from the lines. The novel feature of Experiment 3 was the demonstration that the asymmetry applied to two-digit numbers in the same way that it did to single-digit numbers. To decide whether the number–size asymmetry derived from the better discriminability of the number dimension, we made physical size the more salient dimension in Experiment 4.

## EXPERIMENT 4

### Method

**Participants.** Sixteen Tel Aviv University undergraduates, who had not participated in any of the previous experiments, performed in partial fulfillment of course requirements.

**Stimuli and Apparatus.** The same two-digit numerals as those in Experiment 3 were used, but we modified the values of line length to make this the more discriminable dimension. Therefore, the respective values of line length presented in Experiment 4 were 14 mm (a), 19 mm (b), 25 mm (c), 30 mm (d), 37 mm (e), and 41 mm (f). Expressed in pixels, the respective lengths were 17, 33, 51, 65, 84, and 98. (Note that these lengths reproduce the respective numerical values.) The digits appeared in the constant size of  $5 \times 2.5$  mm. Each line was 8 mm in width, printed black over a white background. It was centered 1 mm beneath its corresponding two-digit number. As in Experiment 3, the following 10 pairs were presented: 17–33, 33–51, 51–65, 65–84, 84–98, 17–51, 51–84, 65–98, 33–84, and 17–98. The numbers were associated with the respective line lengths, so that 17 appeared along with line a, 33 with line b, 51 with line c, 65 with line d, 84 with line e, and 98 with line f. Each pair appeared in congruent and incongruent combinations. For the congruent combination of the pair 84–98, 84 appeared underlined by line e, and 98 was underlined by line f. For the incongruent combination, the lines were reversed, so that 84 now appeared underlined by line f and 98 underlined by line e. The same manipulation was used to create congruent

and incongruent combinations for all the other pairs. For a given pair in a given combination of number and size, each member appeared once in the left and once in the right position. Therefore, each pair was presented in four different arrangements in all. Presenting each individual arrangement four times resulted in a block of 160 trials. Order of stimulus presentation was random.

**Procedure.** The procedure was the same as that in Experiment 3.

### Results

Figure 4 shows the mean RTs for the congruent and incongruent pairs in the two tasks. The respective means were 709 msec for the comparisons of numerical magnitude and 532 msec for the comparisons of line length.<sup>2</sup> The 177-msec advantage favoring the latter [ $F(1,15) = 111.07$ ,  $MS_e = 448$ ,  $p < .00$ ] confirms that line was more salient than number. Error rates were 2.18% for the physical comparisons and 3.24% for the numerical comparisons [ $F(1,15) = 17.79$ ,  $MS_e = 3.21$ ,  $p = .032$ ].

For physical size (line length), the comparisons were not affected by irrelevant numerical magnitude (i.e., SCE was absent). Responses to the congruent trials (mean of 533 msec) were as fast as those to the incongruent trials (mean of 532 msec). Error rates averaged 2.34% for the congruent pairs and 2.03% for the incongruent pairs. In contrast, comparisons of numerical magnitude were affected by the irrelevant dimension of line length. The congruent pairs (695 msec) were compared more quickly than the incongruent pairs (723 msec) [ $F(1,15) = 21.98$ ,  $MS_e = 280$ ,  $p < .005$ ], yielding an RSCE of 28 msec. Error rates were 2.65% for the congruent pairs and 3.8% for the incongruent pairs [ $F(1,15) = 1.04$ ,  $MS_e = 10.46$ ,  $p = .32$ ].

In an overall ANOVA on RT, the main effect for congruity [ $F(1,15) = 16.29$ ,  $MS_e = 4,484$ ,  $p < .05$ ] confirmed the presence of an RSCE. The interaction of task and congruity [ $F(1,15) = 13.71$ ,  $MS_e = 238$ ,  $p < .05$ ] confirmed that the RSCE was larger than the SCE. In a parallel ANOVA on error, only task was significant [ $F(1,15) =$

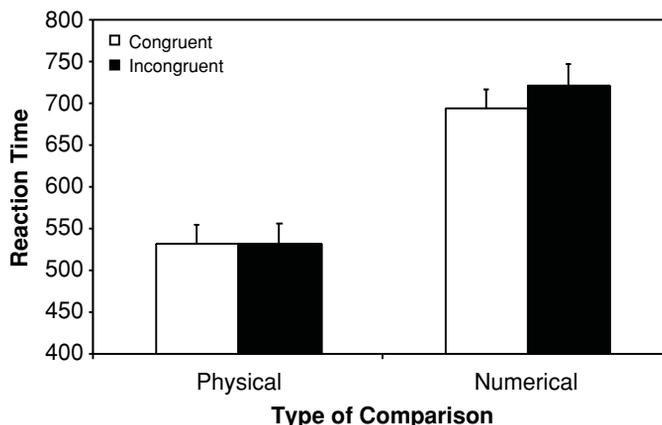


Figure 4. Experiment 4: mean reaction times (in milliseconds) for comparisons of numerical magnitudes and values of line length. Data are averaged across congruent or incongruent combinations of numerical magnitude and length. Bars depict one standard error of the mean.

5.53,  $MS_e = 3.21$ ,  $p < .05$ ], mirroring the overall advantage for line over number in comparisons.

## Discussion

The comparisons of two-digit numbers were affected by the sizes of the lines presented beneath each number. The intrusion of physical size on judgments of numerical magnitude amounted to an appreciable RSCE. In contrast, the comparisons of lines were free of intrusions from the corresponding numbers. The pattern of results is a mirror image of that observed in Experiment 3, documenting the effect of relative dimensional discriminability. We attempted to replicate these results with mixed sets of single- and two-digit numbers.

## EXPERIMENT 5

The unique feature of Experiment 5 was the examination of single- and two-digit numbers in the same experimental setup. The spatially separated version of the Stroop task allows such joint scrutiny. In Experiment 5, we made number the more salient dimension.

## Method

**Participants.** Sixteen Tel Aviv University undergraduates, who had not participated in the previous experiments, performed in partial fulfillment of course requirements.

**Stimuli and Apparatus.** For numerical value, we used the numbers 3, 7, 18, 41, 65, and 89. For line length, we used the values of 2.25 mm (a), 3.25 mm (b), 5 mm (c), 6.25 mm (d), 7 mm (e), and 7.75 mm (f). Expressed in pixels, the respective values were 5, 8, 13, 16, 18, and 20. Each line was 1 mm in width, printed black over a white background. It was centered 1 mm beneath its corresponding number. Each digit appeared in a constant physical size of  $7 \times 4.5$  mm. For two-digit numerals, 0.5 mm separated the component digits. The following 15 pairs were presented: 3–7, 3–18, 3–41, 3–65, 3–89, 7–18, 7–41, 7–65, 7–89, 18–41, 18–65, 18–89, 41–65, 41–89, and 65–89. The numbers were associated with the respective line lengths, so that 3 appeared along with line a, 7 with line b, 18 with

line c, 41 with line d, 65 with line e, and 89 with line f. Each pair appeared in congruent and incongruent combinations. Thus, for the congruent combination of the pair 3–89, 3 appeared underlined by line a, and 89 was underlined by line f. For the incongruent combination, the lines were reversed, so that 3 now appeared underlined by line f and 89 underlined by line a. The same manipulation was used to create congruent and incongruent combinations for all the other pairs. For a given pair in a given combination of number and size, each member appeared once in the left and once in the right position. Therefore, each pair was presented in four different arrangements in all. Presenting each individual arrangement four times resulted in a block of 240 trials. Order of stimulus presentation was random.

**Procedure.** The procedure was the same as that in Experiments 3 and 4.

## Results

Figure 5 shows the mean RTs for the congruent and incongruent pairs in the two tasks. The respective means were 514 msec for the comparisons of numerical magnitude and 585 msec for the comparisons of line length. The 71-msec advantage favoring number [ $F(1,15) = 22.26$ ,  $MS_e = 3,604$ ,  $p < .005$ ] showed that number was indeed more discriminable than line. The error rates were 4.29% for the physical comparisons and 2.99% for the numerical comparisons [ $F(1,15) = 3.66$ ,  $MS_e = 7.4$ ,  $p = .074$ ].

The comparisons of physical size (line length) were affected by irrelevant numerical magnitude. Responses on the congruent trials (mean of 572 msec) were faster than those on the incongruent trials (mean of 592 msec), the difference amounting to an SCE of 20 msec [ $F(1,15) = 15.06$ ,  $MS_e = 215.46$ ,  $p < .05$ ]. Error rates averaged 2.91% for the congruent pairs and 5.67% for the incongruent pairs [ $F(1,15) = 11.57$ ,  $MS_e = 5.26$ ,  $p < .05$ ]. The comparisons of numerical magnitude were free of intrusions from the irrelevant dimension of line length. The congruent pairs were compared as quickly as the incongruent pairs (means of 528 msec for both types of pairs). Error rates were 2.70% for the congruent pairs and 3.28% for the incongruent pairs [ $F(1,15) = 2.49$ ,  $MS_e = 1.05$ ,  $p = .13$ ].

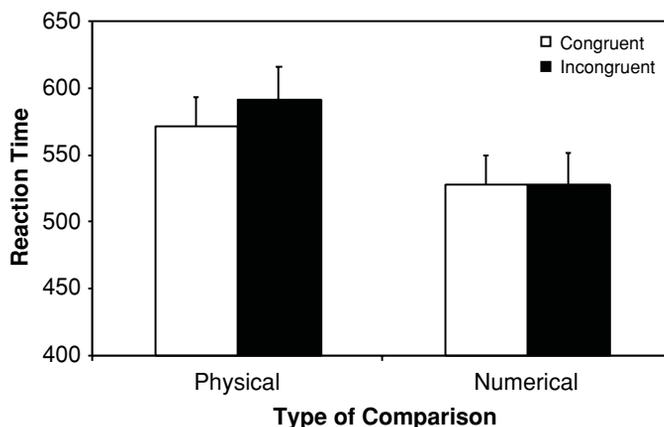


Figure 5. Experiment 5: mean reaction times (in milliseconds) for comparisons of numerical magnitudes and values of line length. Data are averaged across congruent or incongruent combinations of numerical magnitude and length. Bars depict one standard error of the mean.

In an overall ANOVA on the RTs, the main effect of congruity [ $F(1,15) = 9.9, MS_e = 161, p < .00$ ] confirmed the presence of on SCE. The interaction of task and congruity [ $F(1,15) = 42.81, MS_e = 91, p < .05$ ] confirmed that the SCE was larger than the RSCE. In a parallel ANOVA on error, the congruity effect [ $F(1,15) = 13.33, MS_e = 3.33, p < .01$ ] and a task  $\times$  congruity interaction [ $F(1,15) = 6.41, MS_e = 2.98, p < .05$ ] replicated those observed for RT.

## Discussion

In this experiment, the stimulus pairs involved single- and two-digit numbers. The results extended those obtained in Experiment 3 for two-digit numbers. Including single- and two-digit numerals in the same block of trials did not influence performance. We did not detect a difference between the units and the teens in the pattern of interference. In the next experiment, we reversed the relative discriminability of the constituent dimensions. Lines were made more salient than numbers. We expected to replicate the results obtained in Experiment 4—namely, recording an RSCE larger than the SCE.

## EXPERIMENT 6

### Method

**Participants.** Sixteen Tel Aviv University undergraduates, who had not participated in any of the previous experiments, performed in partial fulfillment of course requirements.

**Stimuli and Apparatus.** For numerical value, we used the same single- and two-digit numbers as those in Experiment 5. However, we altered the values of length as follows. The lines were 9 mm (a), 11 mm (b), 14 mm (c), 22 mm (d), 30 mm (e), and 38 mm (f). Expressed in pixels, the respective values were 3, 7, 18, 41, 65, and 89. (Note that these values reproduce the numerical values.) Each line was 8 mm in width, printed black over white background. It was centered 1 mm beneath its corresponding number. Each digit appeared in a constant physical size of  $4 \times 4.5$  mm. The following 15 pairs were presented: 3–7, 3–18, 3–41, 3–65, 3–89, 7–18, 7–41, 7–65, 7–89, 18–41, 18–65, 18–89, 41–65, 41–89, and 65–89. The numbers were associated with the respective line lengths, so that 3 appeared

along with line a, 7 with line b, 18 with line c, 41 with line d, 65 with line e, and 89 with line f. Each pair appeared in congruent and incongruent combinations. Thus, for the congruent combination of the pair 3–89, 3 appeared underlined by line a, and 89 underlined by line f. For the incongruent combination, the lines were reversed, so that 3 now appeared underlined by line f, and 89 underlined by line a. The same manipulation was used to create congruent and incongruent combinations for all the other pairs. For a given pair in a given combination of number and size, each member appeared once in the left and once in the right position. Therefore, each pair was presented in four different arrangements in all. Presenting each individual arrangement four times resulted in a block of 240 trials. Order of stimulus presentation was random.

**Procedure.** The procedure was the same as that in Experiment 5.

### Results

Figure 6 shows the mean RTs for the congruent and incongruent pairs in the two tasks. The respective means were 564 msec for the comparisons of numerical magnitude and 460 msec for the comparisons of line length. The 104-msec advantage favoring line length [ $F(1,15) = 50.3, MS_e = 3,463, p < .00$ ] showed that the lines were more discriminable than the numbers. The error rates were 1.82% for the physical comparisons and 3.88% for the numerical comparisons [ $F(1,15) = 18.23, MS_e = 3.71, p < .005$ ].

The comparisons of physical size (line length) were not affected by irrelevant numerical magnitude. Responses on the congruent trials (mean of 458 msec) were as fast as those on the incongruent trials (mean of 462 msec) [ $F(1,15) = 1.18, MS_e = 100.24, p = .29$ ]. Error rates averaged 1.84% for the congruent pairs and 3.85% for the incongruent pairs; the difference was reliable [ $F(1,15) = 10.93, MS_e = 5.88, p < .05$ ]. By contrast, the comparisons of numerical magnitude were affected by the irrelevant lines. The congruent pairs (551 msec) were compared more quickly than the incongruent pairs (578 msec) [ $F(1,15) = 48.36, MS_e = 115.2, p < .00$ ], yielding an RSCE of 27 msec. Error rates were 2.55% for the congru-

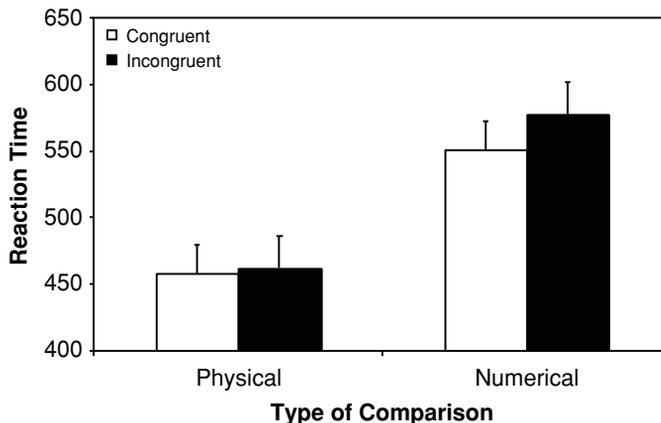


Figure 6. Experiment 6: mean reaction times (in milliseconds) for comparisons of numerical magnitudes and values of line length. Data are averaged across congruent or incongruent combinations of numerical magnitude and length. Bars depict one standard error of the mean.

ent pairs and 5.20% for the incongruent pairs [ $F(1,15) = 9.15$ ,  $MS_e = 6.16$ ,  $p < .05$ ].

In an overall ANOVA on RT, the main effect of congruity [ $F(1,15) = 49.13$ ,  $MS_e = 74$ ,  $p < .00$ ] confirmed the presence of an RSCE. The interaction of task and congruity [ $F(1,15) = 14.41$ ,  $MS_e = 141$ ,  $p < .05$ ] confirmed that the RSCE was larger than the SCE (which itself was not reliable). In a parallel ANOVA on error, we recorded task [ $F(1,15) = 18.23$ ,  $MS_e = 3.71$ ,  $p < .005$ ] and congruity [ $F(1,15) = 10.93$ ,  $MS_e = 5.58$ ,  $p < .05$ ] effects.

## Discussion

We presented the numbers used in Experiment 5 but changed the lines to make them more discriminable than the numbers. Creating a dimensional imbalance favoring the physical dimension of line altered the pattern of interference. Comparisons of numerical magnitude now suffered interference from the incongruent lines, whereas line judgments were free of intrusion from the incongruent numbers. These results replicate and extend those observed in Experiment 4 with two-digit numbers. Presenting single- and two-digit numerals in a single experimental setup did not alter performance. The results did not show any discernible discontinuity between units and teens. The same mode of processing seemed to govern single- and two-digit numbers in the present tasks.

## GENERAL DISCUSSION

The present work was the first study to test the SCE with two-digit numbers. Research on numerical cognition has been notoriously uneven, with a disproportionate majority of the pertinent studies concerned with the units range. There have been few studies in the teens range of such basic phenomena as the distance effect or the effect of spatial-numerical association of response codes (the SNARC effect; Dehaene, Bossini, & Giroux, 1993), and there has been none of the SCE. As a result, there exists "no direct evidence that units and teens are represented on the same number line" (Reynvoet & Brysbaert, 1999, p. 191). Reynvoet and Brysbaert went on to produce data suggesting that units and teens *are* parts of a continuous number line. However, recent studies involving neural network modeling (Verguts, Fias, & Stevens, 2005) and electrophysiological recording (Fias et al., 2001; Göbel, Walsh, & Rushworth, 2001a, 2001b; Nuerk et al., 2001; Whalen & Morelli, 2002) have called into question the psychological reality of a single number line. According to Verguts et al., for example, there are "two numerical systems." One system represents small numbers (and approximate large numbers), the other system multidigit numbers. This second system "decomposes multi-digit numbers . . . on *different* number lines" (p. 38, emphasis added). Thus, the literature allows multiple predictions about the uniformity of the number line concept. We addressed the issue by probing the presence and form of the SCE in the teens range.

We found an SCE with double-digit numbers. Moreover, the SCE with those numbers did not differ from the

SCE with single-digit numbers, a classic phenomenon richly documented in the literature. The most straightforward interpretation of these data is that offered by Dehaene (1997):

Our brain apprehends a two-digit numeral as a whole and transforms it mentally into an internal quantity or magnitude. At this stage, it forgets about the precise digits that led to this quantity. The comparison operation is concerned only with numerical quantities, not the symbols that convey them. (p. 76)

Indeed, extensive syntactic, lexical, and combinatorial operations with the composite digits would have precluded the emergence of SCEs with multidigit numbers. Their presence in this study therefore taps holistic processing of the two-digit numbers (at least at an early stage). Our results are also consistent with those reported by Brysbaert (1995) on the processing times for numerals in the 1–99 range. Processing times were a logarithmic function of the numerical magnitude of the numerals, implying holistic representations of the individual stimuli. Commensurate with Brysbaert's conclusions, our results imply powerful magnitude representations for numerals in the 1–99 range.

The potency of numerical activation tapped in this study is actually stronger than that recorded in existing research with single numerals. To enable SCEs with two-digit numbers, we disjoined semantic and nonsemantic attributes, with the former represented by the numerical values of constant-size digits and the latter by the lengths of bars located beneath the digits. Despite the severance of components, the numbers affected comparisons of the lines, and the lines affected comparisons of the numbers. The SCE thus transcended object- and space-based constraints. The fact that numerical values affected comparisons of lines appearing in separate locations in space (and vice versa) betrays the presence of deep associations connecting the two dimensions.

What is the level of cross talk between numerical magnitude and physical size? With the dimensions integrated in the same stimulus, the cross talk may occur (partly) at the sensory/perceptual level of processing. The SCE can result from the attentional effort required to analyze each stimulus whole into its component attributes. At the sensory/perceptual level, the separate dimensions have not (yet) been well articulated psychologically; the mutual influence of the two channels can easily engender (an early portion of) the SCE. With the dimensions disintegrated into different stimuli in space, the cross talk occurs entirely at a semantic level. As purely visual stimuli, there is little in common between pairs of idiographic signs (which people call "seven," "forty-six," or "numbers" in general) and pairs of lines. With symbolic meaning attached to the former, there is still little room for early interference, because the two stimuli are separated in space, each processed through different receptors. Therefore, spatially separated SCEs most clearly reflect dimensional cross talk at a semantic level of processing.

The properties of this cross talk were elucidated through manipulations of relative salience. The SCE was present

when number was more discriminable than size, but it was absent and a reverse SCE appeared when size was more discriminable than number. This pattern is readily accommodated by a relative speed of processing account: The speedier attribute (by virtue of enhanced discriminability) interfered with the slower attribute more than vice versa. Number and size were always activated, but the processing of the less salient attribute was too sluggish to appreciably influence the response. Consequently, the absence of the SCE (in Experiments 4 and 6) should not be interpreted as reflecting on the absence of automatic activation of numerical information.

The last point is notable. It is not always recognized that the relative speed of processing account of the SCE is contingent on the prior assumption of automatic activation of numerical value. Schwarz and Ischebeck (2003) are commendably explicit on this point, earmarking the "automaticity of processing the various attributes of a stimulus" as one of the "basic assumptions" of the relative speed account (p. 508). Automaticity indeed plays a prominent role in their quantitative, evidence-accrual model of the SCE. Because the "information about the irrelevant attribute cannot be completely ignored" (Schwarz & Ischebeck, 2003, p. 510), the rate at which evidence is accumulated (eventuating in the experimental RT) is jointly determined by the task-relevant attribute and the task-irrelevant attribute. The extent of the influence of the irrelevant attribute is specified by an *automaticity* parameter that varies between zero (no activation) and unity (maximum activation). The SCE results from the difference in rates of evidence accrual on congruent and incongruent trials (provided that the automaticity parameter differs from zero). On the former trials, evidence accrual along the irrelevant attribute augments that on the relevant attribute. On the latter, evidence accrual along the irrelevant attribute detracts from that along the relevant attribute. Finally, the SCE decreases with large distance along the relevant dimension but increases with large distance along the irrelevant dimension. Clearly, the irrelevant attribute has greater opportunity to influence processing in the latter than in the former case (see Roelofs, 2006, for a recent chronometric examination of the SCE and discriminability with further numerical dimensions).

Our results are conveniently accommodated by the Schwarz and Ischebeck (2003) model. An SCE is produced when the generic drift rate for number is larger than that for size. A reverse SCE is produced when the respective rates are reversed. The model predicts error rates that vary in tandem with the respective RTs, and this prediction is also borne out by the present data. In general, the model treats number and size in a symmetric fashion, a feature shared with the present approach. This much granted, dimensional speed, on the one hand, and interference, on the other hand, are *independent* in the model. Relevant Attribute A can be much slower than irrelevant Attribute B, yet the responses to A can be completely free of intrusions from B. A generally slower attribute can interfere with a generally speedier attribute more than vice versa. Separating dimensional drift rate from dimensional

automaticity greatly increases the flexibility of the Schwarz and Ischebeck model, but this maneuver also removes the model a bit from mainstream relative speed accounts, as well as from our contextual approach (cf. Melara & Algom, 2003). In our approach, a simple rule governs performance: The more discriminable dimension intrudes on the less discriminable dimension more than vice versa.<sup>3</sup>

In this study, we did not test the dependence of the SCE on distance, because a given pair of numbers always appeared with a corresponding pair of size values (in a congruent or an incongruent arrangement). Our goal in this study was different: to ensure that precisely the same discriminability (time relation) applied to *all* the stimuli in a given experiment. We emulated the item-specific method developed in the domain of the classic Stroop effect by Jacoby and his associates (e.g., Jacoby, Lindsay, & Hesses, 2003). Consider the word *red* appearing only in two possible colors, red (congruent) and blue (incongruent). The difference in color-naming performance gives the item-specific Stroop effect with that stimulus (the word *red*). The same calculation is performed with other items carefully selected to satisfy matched presentation. Applying this method, the listwide effect (Stroop or the SCE) is the sum of the item-specific effects observed with the individual items. We introduced Jacoby et al.'s method into the numerical realm in order to achieve as rigorous a control of relative dimensional discriminability as is possibly allowed by the dimensions in question. A main goal of this study was to test the presence and magnitude of the SCE as a sole function of relative dimensional discriminability.

Finally, dimensional discriminability itself deserves comment, because this notion has been fraught with inconsistency and confusion in the literature. First, dimensional discriminability is different from physical or symbolic distance. Given a pair of the stimuli presented in this study (i.e., ones with constant distance), one can still make the stimuli more or less discriminable by modifying such attributes as color, location, or font. Discriminability specifies *psychological* distance; the latter is obviously affected by physical or symbolic distance along the tested dimensions but is not limited to those manipulations. Second, discriminability is not identical with speed of processing. It is an endogenous structure based on the physical makeup of the presented stimuli *and* the individual's long-term experience with the dimension in question. It is recovered from model-estimated rates of activation (within tectonic theory [Melara & Algom, 2003] or within a model such as that developed by Schwarz & Ischebeck, 2003). Processing speed serves as a useful surrogate of discriminability. Third, manipulations of discriminability do not render the pertinent stimuli illegible or difficult to perceive. Detectability and recognition remain unaltered, as does accuracy (nearly perfect with the less discriminable dimension, too). What manipulations of discriminability do accomplish is making one or the other attribute more salient. That attribute commands priority in attention at the expense of the less salient attribute.

Most important, relative dimensional discriminability is willy-nilly manipulated in every experiment. Creating

the to-be-presented stimuli, the experimenter determines (unwittingly perhaps) the relative salience of the constituent attributes. It was Garner's (1962, 1974) insight that made explicit the influence on the results of the routine act of fixing the stimulus values (at whatever level). The experimenter must select those values; yet, by virtue of the selection, the experimenter also determines to a large extent the attentional outcome. A unique feature of our study was the judicious manipulation (and monitoring) of relative dimensional discriminability.

Garner's (1962, 1974) insight, augmented by the role of relative dimensional discriminability, enables an alternative framework for the present results that does not appeal to the concept of automaticity. According to the attention view, the selectivity of attention to the relevant dimension can succeed (resulting in the absence of the SCE) or fail (indexed by the presence of the SCE), depending on factors of context (discriminability, in this study). If the former, the irrelevant dimension has been successfully ignored and minimally processed; if the latter, the irrelevant dimension has been fully (semantically) processed. On this view, meaning is not activated in an automatic fashion upon the presentation of a numeral just for any purpose. As Stevens (1951) has noted, a numeral is sometimes an ink mark on a piece of paper. This account can explain the robustness of the SCE by the superior discriminability of the number dimension in the great majority of existing studies. Without precautions, the numerical values presented are typically easier to tell apart than the values of physical size. On the attention view, therefore, the SCE is not a mandatory outcome of numerical processing.

Let us emphasize that our results are fully consistent with a relative speed account and that they certainly do not rule out automatic processing. However, the attention view does provide a parsimonious explanation. The complete symmetry of number (a semantic variable) and physical size in response to the experimental manipulations is a natural prediction of the attention view. An asymmetry that typically favors the semantic attribute is the signature of the automaticity approach. Looked at from a broad perspective, what is remarkable about the present results is how little stimulus alchemy suffices to undo the effects of long-term learning, excessive practice, and automatic tendencies in given experimental contexts.

## REFERENCES

- ALGOM, D., DEKEL, A., & PANSKY, A. (1996). The perception of number from the separability of the stimulus: The Stroop effect revisited. *Memory & Cognition*, *24*, 557-572.
- BARGH, J. A. (1992). The ecology of automaticity: Toward establishing the conditions needed to produce automatic processing effects. *American Journal of Psychology*, *105*, 181-199.
- BESNER, D., & COLTHEART, M. (1979). Ideographic and alphabetic processing in skilled reading of English. *Neuropsychologia*, *17*, 467-472.
- BRYLSBAERT, M. (1995). Arabic number reading: On the nature of numerical scale and the origin of phonological recoding. *Journal of Experimental Psychology: General*, *124*, 434-452.
- DAMIAN, M. F. (2004). Asymmetries in the processing of Arabic digits and number words. *Memory & Cognition*, *32*, 164-171.
- DEHAENE, S. (1997). *The number sense: How the mind creates mathematics*. New York: Oxford University Press.
- DEHAENE, S., & AKHAVEIN, R. (1995). Attention, automaticity, and levels of representation in number processing. *Journal of Experimental Psychology: Learning, Memory, & 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., DUPOUX, E., & MEHLER, J. (1990). Is numerical comparison digital? Analogical and symbolic effects in two-digit number comparison. *Journal of Experimental Psychology: Human Perception & Performance*, *16*, 626-641.
- DISHON-BERKOVITS, M., & ALGOM, D. (2000). The Stroop effect: It is not the robust phenomenon that you have thought it to be. *Memory & Cognition*, *28*, 1437-1449.
- FIAS, W., LAMMERTYN, J., REYNVOET, B., DUPONT, P., & ORBAN, G. A. (2003). Parietal representation of symbolic and nonsymbolic magnitude. *Journal of Cognitive Neuroscience*, *15*, 47-56.
- FIAS, W., REYNVOET, B., & BRYLSBAERT, M. (2001). Are Arabic numerals processed as pictures in a Stroop interference task? *Psychological Research*, *65*, 242-249.
- FOLTZ, G. S., POLTROCK, S. E., & POTTS, G. R. (1984). Mental comparison of size and magnitude: Size congruity effects. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *10*, 442-453.
- GARNER, W. R. (1962). *Uncertainty and structure as psychological concepts*. New York: Wiley.
- GARNER, W. R. (1974). *The processing of information and structure*. Hillsdale, NJ: Erlbaum.
- GLASER, W. R., & GLASER, M. O. (1989). Context effects in Stroop-like word and picture processing. *Journal of Experimental Psychology: General*, *118*, 13-42.
- GÖBEL, S., WALSH, V., & RUSHWORTH, M. F. S. (2001a). The mental number line and the human angular gyrus. *NeuroImage*, *14*, 1278-1289.
- GÖBEL, S., WALSH, V., & RUSHWORTH, M. F. S. (2001b). rTMS disrupts the representation of small numbers in supramarginal gyrus. *NeuroImage*, *3*, 409.
- HATTA, T. (1977). Recognition of Japanese kanji in the left and right visual field. *Neuropsychologia*, *15*, 685-688.
- HENIK, A., & TZELGOV, J. (1982). Is three greater than five: The relation between physical and semantic size in comparison tasks. *Memory & Cognition*, *10*, 389-395.
- HINRICH, J. V., YURKO, D. S., & HU, J. M. (1981). Two-digit number comparison: Use of place information. *Journal of Experimental Psychology: Human Perception & Performance*, *7*, 890-901.
- ISCHEBECK, A. (2003). Differences between digit naming and number word reading in a flanker task. *Memory & Cognition*, *31*, 529-537.
- JACOBY, L. L., LINDSAY, D. S., & HESSELS, S. (2003). Item-specific control of automatic processes: Stroop process dissociations. *Psychonomic Bulletin & Review*, *10*, 638-644.
- KAHNEMAN, D., & CHAJCZYK, D. (1983). Tests of automaticity of reading: Dilution of Stroop effects by color-irrelevant stimuli. *Journal of Experimental Psychology: Human Perception & Performance*, *9*, 497-509.
- LOGAN, G. D. (1980). Attention and automaticity in Stroop and priming tasks: Theory and data. *Cognitive Psychology*, *12*, 523-553.
- LOGAN, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*, 492-527.
- MACLEOD, C. M., & HODDER, S. L. (1998). Presenting two incongruent color words on a single trial does not alter Stroop interference. *Memory & Cognition*, *26*, 212-219.
- MELARA, R. D., & ALGOM, D. (2003). Driven by information: A tectonic theory of Stroop effects. *Psychological Review*, *110*, 422-471.
- MOYER, R. S., & LANDAUER, T. K. (1967). The time required for judgments of numerical inequality. *Nature*, *215*, 1519-1520.
- NUERK, H.-C., WEGER, U., & WILLMES, K. (2001). Decade breaks in the mental number line? Putting the tens and units back in different bins. *Cognition*, *82*, B25-B33.
- PANSKY, A., & ALGOM, D. (1999). Stroop and Garner effects in comparative judgments of numerals: The role of attention. *Journal of Experimental Psychology: Human Perception & Performance*, *25*, 39-58.
- PANSKY, A., & ALGOM, D. (2002). Comparative judgment of numerosity and numerical magnitude: Attention preempts automaticity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *28*, 259-274.

- REYNVOET, B., & BRYLSBAERT, M. (1999). Single-digit and two-digit Arabic numerals address the same semantic number line. *Cognition*, **72**, 191-201.
- ROELOFS, A. (2006). Functional architecture of naming dice, digits, and number words. *Language & Cognitive Processes*, **21**, 78-111.
- SCHWARZ, W., & HEINZE, H. J. (1998). On the interaction of numerical and size information in digit comparison: A behavioral and event-related potential study. *Neuropsychologia*, **36**, 1167-1179.
- SCHWARZ, W., & ISCHEBECK, A. (2003). On the relative speed account of number-size interference in comparative judgments of numerals. *Journal of Experimental Psychology: Human Perception & Performance*, **29**, 507-522.
- STEVENS, S. S. (1951). Mathematics, measurement, and psychophysics. In S. S. Stevens (Ed.), *Handbook of experimental psychology* (pp. 1-49). New York: Wiley.
- STROOP, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, **18**, 643-662.
- TAKAHASHI, A., & GREEN, D. (1983). Numerical judgments with Kanji and Kana. *Neuropsychologia*, **21**, 259-263.
- TZELGOV, J., MEYER, J., & HENIK, A. (1992). Automatic and intentional processing of numerical information. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 166-179.
- TZELGOV, J., YEHENE, V., KOTLER, L., & ALON, A. (2000). Automatic comparisons of artificial digits never compared: Learning linear ordering relations. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **26**, 103-120.
- VAID, J. (1985). Numerical size comparisons in a phonologically transparent script. *Perception & Psychophysics*, **37**, 592-595.
- VAID, J., & CORINA, D. (1989). Visual field asymmetries in numerical size comparisons of digits, words, and signs. *Brain & Language*, **36**, 117-126.
- VERGUTS, T., FIAS, W., & STEVENS, M. [A.] (2005). A model of exact small-number representation. *Psychonomic Bulletin & Review*, **12**, 66-80.
- WHALEN, J., & MORELLI, F. (2002, November). *How the brain represents numerical quantity*. Poster presented at the 43rd Annual Meeting of the Psychonomic Society, Kansas City, MO.

## NOTES

1. Extensive pilot testing preceded this and all of the subsequent experiments, in order to effect the planned pattern of relative dimensional salience.

2. The same numbers were presented in this and in the previous experiment, yet the overall RTs differed by more than 100 msec (means of 592 and 709 msec, respectively, for Experiments 3 and 4). The size of this difference is a bit surprising, yet the following points should be considered. First, independent groups of participants performed in the two experiments. Second, responses to the numbers were free of interference from the irrelevant lines in Experiment 3 but suffered such interference in Experiment 4.

3. Severing dimensional speed from dimensional interference renders the notion of automatic activation somewhat ambiguous. In standard automaticity theory (e.g., Bargh, 1992; Logan, 1980, 1988; Tzelgov, Yehene, Kotler, & Alon, 2000), speed of processing and propensity to interfere with concurrent activity are part and parcel of the same phenomenon. In the Schwarz and Ischebeck model, by contrast, dimensions can be automatic in the first sense (speed), in the second sense (interference), or in both senses or can display a trade-off of speed and interference (see Schwarz & Ischebeck, 2003, Experiment 1). From an empirical point of view, it is unclear how one can *selectively* influence generic drift rate and automaticity parameters in a given experiment.

(Manuscript received June 15, 2004;  
revision accepted for publication March 10, 2005.)