

A memory-based account of automatic numerosity processing

JESSICA M. CHOPLIN and GORDON D. LOGAN
Vanderbilt University, Nashville, Tennessee

We investigated the mechanisms responsible for the automatic processing of the numerosities represented by digits in the size congruity effect (Henik & Tzelgov, 1982). The algorithmic model assumes that relational comparisons of digit magnitudes (e.g., *larger than* {8,2}) create this effect. If so, congruity effects ought to require two digits. Memory-based models assume that associations between individual digits and the attributes “small” and “large” create this effect. If so, congruity effects ought only to require one digit. Contrary to the algorithmic model and consistent with memory-based models, congruity effects were just as large when subjects judged the relative physical sizes of small digits paired with letters as when they judged the relative physical sizes of two digits. This finding suggests that size congruity effects can be produced without comparison algorithms.

In the size congruity effect, subjects observe two digits, one of which is presented in a larger font size than the other. Subjects then judge which of the two digits is presented in the larger (or smaller) font size. Only the physical sizes of the digits are relevant for this task; the numerosities represented by the digits are irrelevant. Nevertheless, a size congruity effect is typically observed such that the time needed to identify the relative sizes of the digits is faster when the difference in the numerosities represented by the digits is congruent with the difference in font sizes (e.g., {2 8}) than when it is incongruent with the difference in font sizes (e.g., {2 8}). These size congruity effects also tend to interact with the distance between the numerosities represented by the digits such that congruity effects are larger for pairs that are far away from each other (e.g., {2 8}) than for pairs that are close together (e.g., {2 4}). These effects demonstrate that the numerosities represented by the digits interfere with size judgments even though people are trying to ignore them and pay exclusive attention to physical size. The purpose of the research reported here was to explore the mechanisms responsible for these size congruity effects.

Several views regarding the nature of the representations and processes involved in the size congruity effect have been proposed. One view—henceforth, called the *algorithmic model*—is that subjects map the numerosities represented by the digits to an analogue numerosity representation and then use the analogue representation to identify the digit that represents the larger (or smaller) numerosity. The result of this comparison process is thought

to interfere with judgments of physical size (Dehaene & Akhaverin, 1995; Schwarz & Ischebeck, 2003; Tzelgov, Yehene, Kotler, & Alon, 2000). The interaction between size congruity and distance is often cited as especially strong evidence for this view because the larger congruity effects for digit pairs whose members represent dissimilar numerosities are believed to reflect faster processing for values that are more discriminable (i.e., farther apart) on the analogue numerosity representation.

The algorithmic model assumes that processing of the numerosities represented by digits is accomplished in multiple stages. The numerosities represented by each of the to-be-compared digits, A and B, must be retrieved and mapped to the analogue numerosity representation. The values on the analogue representation must then be compared to produce a relational judgment of the form *larger than* [numerosity A, numerosity B] or *smaller than* [numerosity B, numerosity A]. This calculation is faster the greater the difference between the numerosities, purportedly because larger differences are more discriminable on the analogue numerosity representation. Finally, this relational judgment undergoes processing by which it facilitates or interferes with relational judgments of the form *larger than* [physical size B, physical size A] or *smaller than* [physical size A, physical size B]. According to this view, automatic processing of the numerosities represented by digits in the size congruity effect is a type of procedural memory wherein multiple stages of information processing become seamlessly linked together through practice to produce a result (Dehaene & Akhaverin, 1995; Schwarz & Ischebeck, 2003; Tzelgov et al., 2000).

An alternative to the algorithmic model is suggested by memory-based theories of automaticity (Logan, 1988). These theories assume that automaticity reflects a single-step retrieval of information from memory. Previous work has argued that memory-based automaticity is sufficient

This research was supported by National Science Foundation Grant BCS 0133202. We thank Joseph Tzelgov, Avishai Henik, Daniel Algom, and Stan Dehaene for helpful discussions. Correspondence concerning this article should be addressed to J. M. Choplin, DePaul University, Department of Psychology, 2219 North Kenmore Avenue, Chicago, IL 60614-3504 (e-mail: jchoplin@depaul.edu).

to explain the development of automaticity in tasks such as memory search (Strayer & Kramer, 1990), category search (Logan & Etherton, 1994), lexical decisions (Logan, 1988), and numerosity judgments (Lassaline & Logan, 1993; Logan & Zbrodoff, 2003; Palmeri, 1997). Postulating changes in the algorithms used to perform those tasks did not appear necessary to explain the development of automaticity in those tasks, and there is no a priori reason to think that the automatic processing of numerosity information in the size congruity effect should somehow be different. A memory-based account of size congruity effects would be more consistent with previous work on automatic processing (see Tzelgov, Meyer, & Henik, 1992, for an example of a memory-based account of the size congruity effect; see also Algom, Dekel, & Pansky, 1996, and Pansky & Algom, 1999, for a memory-based account that does not appeal to the concept of automaticity).¹

We present three memory-based models of the size congruity effect: the *laterality model*, the *graded associations model*, and the *retrieved instance model*. The basic idea behind all three memory-based models is that associations between individual digits and the attributes “small” and “large” produce size congruity effects. These models differ only in the hypothesized strength of the association between individual digits and the attributes “small” and “large.”

The laterality model was proposed by Tzelgov et al. (1992) as a supplement to the algorithmic model described above. The word *laterality* here refers to whether a digit is less than or greater than 5. In this model, digits representing values smaller than 5 (i.e., 1, 2, 3, and 4) activate the attribute “small,” and digits representing values larger than 5 (i.e., 6, 7, 8, and 9) activate the attribute “large.” In this model, facilitation occurs when physically small digits are associated with the attribute “small” and physically large digits are associated with the attribute “large.” Interference occurs when physically small digits are associated with the attribute “large” and physically large digits are associated with the attribute “small.” This model predicts that congruity effects only ought to be observed for bilateral pairs of digits (i.e., pairs in which one digit is less than 5 and the other digit is greater than 5). The size of the congruity effects ought to be the same for all bilateral pairs. Congruity effects ought not to be observed for unilateral pairs (i.e., pairs in which both digits are less than or greater than 5). Interactions between size congruity and distance only occur because bilateral pairs are more frequent among more distant pairs. Of the pairs that have a distance of one step between them, no pairs (0%) straddle 5. Of the pairs that have a distance of two steps between them, only one pair out of seven (14% of pairs) straddles 5 (i.e., the pair {4, 6} crosses 5, but {1, 3}, {2, 4}, {3, 5}, {5, 7}, {6, 8}, and {7, 9} do not). Of the pairs that have a distance of three steps between them, two pairs out of six (33% of pairs) straddle 5 (i.e., {3, 6} and {4, 7} cross 5, but {1, 4}, {2, 5}, {5, 8}, and {6, 9} do not), and so forth.

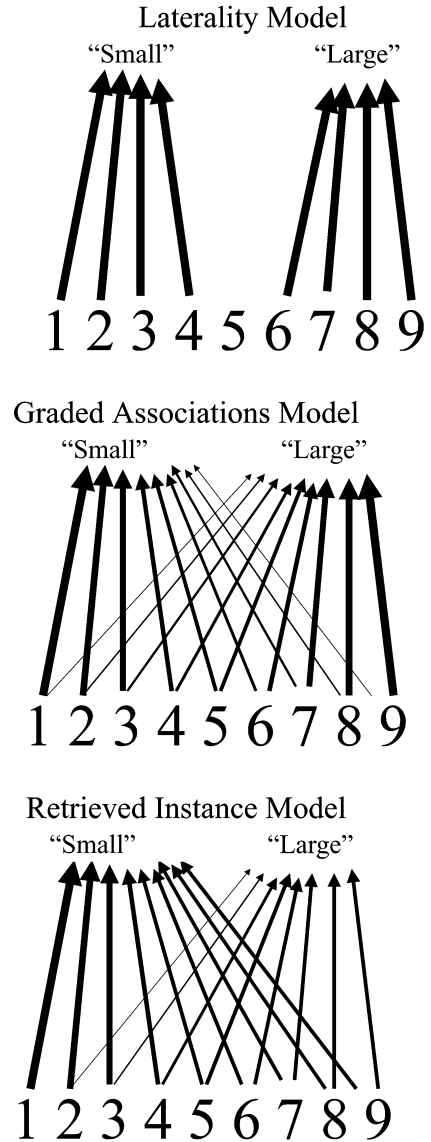


Figure 1. Associations between digits and the attributes “small” and “large” in the three memory-based models of the size congruity effect. Stronger associations are represented by thicker lines.

The graded associations model was inspired by Leth-Steenen and Marley (2000). They proposed a connectionist model of symbolic comparison that learned to identify which of two men was taller by positively associating greater heights with a response identifying the man as taller. The tallest man’s name had the strongest association with the response identifying him as the tallest, followed by the second-tallest man’s name, and so forth. Similar graded associations between digits and the attributes “small” and “large” could produce size congruity effects. In such a model, the digit 1 would have the strongest association to the attribute “small,” followed by the digit 2, and so forth. The digit 9 would have the strongest association to the attribute “large,” fol-

lowed by the digit 8, and so forth (see Figure 1, middle panel). In this model, facilitation occurs when the physically larger digit is more strongly associated with the attribute "large" than it is with the attribute "small" and the physically smaller digit is more strongly associated with the attribute "small" than it is with the attribute "large." For example, in the pair {2 8}, the digit 8 would be more strongly associated with the attribute "large" than it would be with the attribute "small." It would, therefore, be easy to identify the 8 as large. The digit 2 would be more strongly associated with the attribute "small" than it would be with the attribute "large." It would, therefore, be easy to identify the 2 as small. Interference would occur in this model when the physically larger digit is more strongly associated with the attribute "small" than it is with the attribute "large" and the physically smaller digit is more strongly associated with the attribute "large" than it is with the attribute "small." For example, in the pair {2 8}, the association between the digit 2 and the attribute "small" would make it difficult to identify the 2 as large and the association between the digit 8 and the attribute "large" would make it difficult to identify the 8 as small. Interactions between size congruity and distance would occur because the differences in the strengths of the associations would be greater for pairs whose members represent dissimilar numerosities than for pairs whose members represent similar numerosities. For example, the strengths of the associations between the digit 4 and the attributes "small" and "large" would be more similar to the strengths of the associations between the digit 2 and the attributes "small" and "large" than would be the strengths of the associations between the digit 8 and the attributes "small" and "large." Congruity effects would, therefore, be smaller for the pair {2 4} than for the pair {2 8}.

The retrieved instance model is suggested by the view that the strengths of the associations between digits and attributes would likely be determined by the instances in people's daily lives in which they have identified digits as small or large. This model could be implemented by either instance-based (Logan, 1988) or strength-based (Rickard, 1997, 1999) retrieval processes. The key point is that association strength depends on frequency of occurrence. If a digit were judged small in every instance in which it was apprehended, then retrieval of instances of that digit would always retrieve the attribute "small" and would never retrieve the attribute "large." The digit 1, for example, might rarely be judged large. By contrast, if a digit were judged small in half of the instances in which it was apprehended and judged large in the other half, then retrieval of instances of that digit would retrieve the attribute "small" half of the time and retrieve the attribute "large" half of the time. The digit 9, for example, might be judged "large" in some contexts (i.e., the context of the single digits 1–9) and "small" in other contexts (i.e., the context of numbers 1–99). The associations between the digit 9 and the attributes "small" and "large" might, therefore, be approximately equal.

Without a complete log of every instance in which subjects have judged digits small or large throughout their lives, we cannot know the precise strengths between particular digits and the attributes "small" and "large" for each subject a priori. Nevertheless, the general pattern of these strengths can be surmised from the observation that for comparisons between integers, small digits will tend to be judged small. The digit 1 will always be judged small. The digit 2 will be judged small unless it is compared to the digit 1, and so forth. The digit 1 should have the strongest association to the attribute "small" and the weakest association to the attribute "large," followed by the digit 2, then the digit 3, and so forth (see Figure 1, bottom panel). Large digits will be judged "large" in some contexts and "small" in others. The association strengths in the bottom panel of Figure 1 were set under the assumption that while the digits representing numerosities smaller than 5 are more strongly associated with the attribute "small" than they are associated with the attribute "large," the digits representing numerosities larger than 5 are approximately equally associated with the attributes "small" and "large."

To investigate this assumption, we estimated the approximate frequencies of digit comparisons in public discourse by searching the World Wide Web (using the search engine at Google.com).² For each number (1–9), we searched the Web for phrases identifying it as large, larger, and more than (e.g., "2 is large," "2 is larger," "2 is more than") and small, smaller, and less than (e.g., "2 is small," "2 is smaller," "2 is less than"). We did this both with the number in digit form (i.e., 1, 2, 3, etc.) and with the number in word form (i.e., one, two, three, etc.). The results are presented in Figure 2. Note that 1 is judged small much more often than it is judged large, followed by 2 and 3. The numbers 4 and 5 are judged small slightly more often than they are judged large. Larger numbers are judged small and large approximately equally often. These comparison frequencies are consistent with the association strengths assumed in the bottom panel of Figure 1.

The retrieved instance model, like the graded associations model, assumes that congruity effects and interactions between size congruity and distance occur due to differences in the respective strengths of the associations between each digit in a pair and the attributes "small" and "large." However, the pattern of predicted effects is different. While the graded associations model predicts that congruity effects always ought to increase with distance, the retrieved instance model does not. In particular, for pairs where the larger digit is equally associated with both of the attributes "small" and "large" and the smaller digit is not, the retrieved instance model predicts that the size of congruity effects ought to depend upon the small digit in the pair rather than distance per se. For example, assuming the association strengths shown in the bottom panel of Figure 1, the retrieved instance model would predict that the congruity effect for the pair {8 2} would likely be larger than the congruity effect for the pair {8 3}. The reason for this prediction is that in the bot-

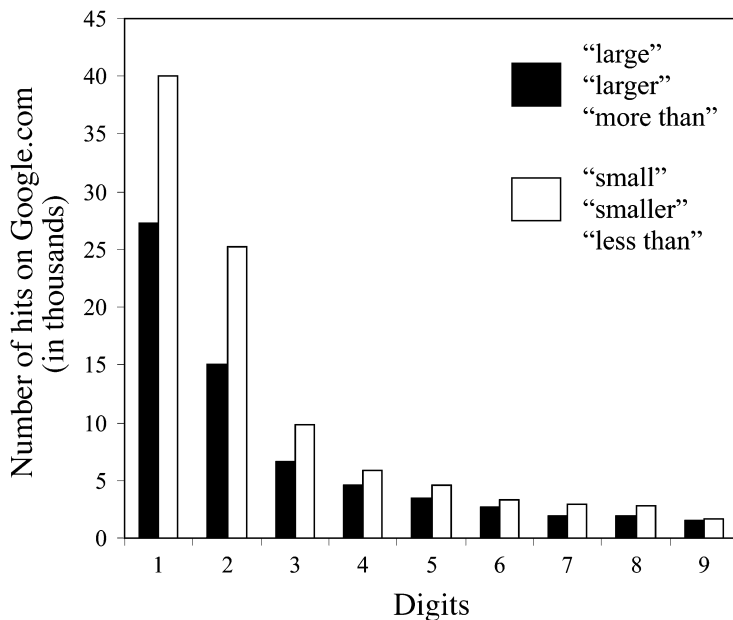


Figure 2. Number of hits on Google.com for phrases identifying numbers—in either digit or word form—as “large,” “larger,” or “more than” (e.g., “8 is larger”) and “small,” “smaller,” or “less than.” Unlike large numbers, small numbers are identified as “small,” “smaller,” or “less than” more frequently than they are identified as “large,” “larger,” or “more than.”

tom panel of Figure 1, the difference in the strength of the associations between the digit 2 and the attributes “small” and “large” is greater than is the difference in the strength of the associations between the digit 3 and the attributes “small” and “large.” By contrast, the congruity effect for the pair {2 8} would likely be approximately the same as the congruity effect for the pair {2 9}. In the bottom panel of Figure 1, the difference in the strength of the associations between the digit 8 and the attributes “small” and “large” is approximately the same as that between the digit 9 and the attributes “small” and “large.”

Our goal in pursuing the research reported here was to pit the assumption of the algorithmic model that relational judgments of the numerosities represented by digits (e.g., judgments like *larger than* [8,2]) are responsible for the size congruity effect against the assumption of the memory-based models that associations between digits and the attributes “small” and “large” produce size congruity effects. If relational judgments of numerosity produce size congruity effects, two digits will be required to produce size congruity effects because relations and the algorithms that compute them require two inputs. Size congruity effects ought not to be observed when the task only involves a single digit.³

To test this prediction, we performed an experiment in which the small digits (i.e., 1, 2, 3, and 4) and the large digits (i.e., 6, 7, 8, and 9) were paired with letters, in addition to being paired with each other. One character was physically larger than the other, and subjects identified the larger or the smaller member of the pair. If comparisons between the numerosities represented by the two digits are responsi-

ble for size congruity effects—as assumed by the algorithmic model—then size congruity effects ought not to occur when digits are paired with letters. By contrast, if attributes associated with individual digits are responsible for size congruity effects in judging the relative sizes of digits presented in pairs—as assumed by the memory-based models—then size congruity effects also ought to be observable in conditions in which digits are paired with letters.

The pattern of results among pairs of digits will also test the predictions of the four models. The laterality model predicts no interaction between congruity and distance in this experiment, because all pairs of digits were bilateral (i.e., contained one digit that was less than 5 and one digit that was more than 5). The algorithmic and graded associations models predict that congruity and distance always ought to interact, because discriminability or differences in the strengths of the associations between digits and the attributes “small” and “large” always ought to increase with distance. The retrieved instance model predicts that the size of congruity effects ought to be a function of the small digit and not distance per se, because the small digits 1, 2, 3, and 4 are assumed to differ greatly in the degree to which they are associated with the attributes “small” and “large,” whereas the large digits 6, 7, 8, and 9 are not (see Figure 1, bottom panel, and Figure 2).

METHOD

Subjects

Forty-eight undergraduate students with normal or corrected-to-normal vision participated in partial fulfillment of course requirements.

Apparatus and Stimuli

The subjects sat approximately 60 cm in front of a 30 × 40 cm computer screen. On each of 2,016 trials (plus 4 training trials), a fixation point was presented at the center of the screen for 500 msec, followed by two characters. Both characters were presented at the vertical center of the screen. One character was presented approximately 1.2 cm (1.15° visual angle) left of the horizontal center of the screen; the other character was presented approximately 1.2 cm (1.15° visual angle) right of the horizontal center of the screen. One character was presented in Courier 30-point type (approximately 1.0 × 1.4 cm, 0.95° × 1.34° visual angle), and its counterpart was presented in Courier 40-point type (approximately 1.4 × 2.0 cm, 1.34° × 1.91° visual angle). The assignment of characters to presentation on the left or right side of the screen, as well as presentation in 30- or 40-point type, was fully counterbalanced. These characters remained on the screen until the subject responded, after which the next trial began.

We created three different character-pair conditions. We created the first condition (henceforth, the *small+large condition*, for small digits paired with large digits) by taking all pairwise combinations involving one digit that was less than 5—that is, selected from the set {1, 2, 3, 4}, paired with one digit that was more than 5—namely, selected from the set {6, 7, 8, 9}. The frequency with which these pairs were presented equated for the distance between the pairs, such that a distance of 8 between the pairs (i.e., the pair {1, 9}) was presented just as frequently as a distance of 7 between the pairs (i.e., the pairs {1, 8} and {2, 9} together) and so forth. To assess the extent to which attributes associated with the small digits account for size congruity effects, we created a second condition by taking all of the pairs used in the small+large condition and substituting the letter H for the digit 6, the letter N for the digit 7, the letter P for the digit 8, and the letter T for the digit 9. This condition will be called the *small+letter condition* (for small digits paired with letters). To assess the extent to which attributes associated with the large digits account for size congruity effects, we created a third condition by taking all of the pairs used in the small+large condition and substituting the letter J for the digit 1, the letter L for the digit 2, the letter R for the digit 3, and the letter V for the digit 4. This condition will be called the *large+letter condition* (for large digits paired with letters).

Procedure

Twenty-four subjects identified which of the two characters was physically smaller, and the other 24 subjects identified which of the two characters was physically larger as accurately and quickly as they could. They did this by pressing the “S” key, which is on the left side of the keyboard, if the character on the left was the smaller (or larger) of the two or the “K” key, which is on the right side of the keyboard, if the character on the right was the smaller (or larger) of the two. Character pairs were presented in a random order in which the three conditions (small+large, small+letter, large+letter) were intermixed. After every 253 trials, the subjects were given a break. The experiment lasted approximately 55 min.

RESULTS AND DISCUSSION

To discriminate the algorithmic model from the memory-based models, the first analysis investigated reaction times and error rates for pairs in which each of the eight digits used in this experiment (i.e., 1–4 and 6–9) were paired with letters (i.e., the small+letter and large+letter conditions). The algorithmic model predicted no congruity effects in these conditions because pairs in these conditions had only one digit, whereas the memory-based models predicted congruity effects in these condi-

tions because they assume that there are associations between single digits and the attributes “small” and “large.” Reaction times from the small+large condition in which digits were paired with other digits were not included in this first analysis because the congruity effects in that condition might have been confounded by algorithmic processing (if it occurs) or the associations of the digits with which each digit was paired.

Reaction time results from this analysis are presented in Figure 3, and error rate results are presented in Figure 4. These data were analyzed using 2 (font size: digit in 30-point type, digit in 40-point type) × 8 (digits) × 2 (type of judgment: larger vs. smaller)⁴ mixed-factor analyses of variance (ANOVAs). All three memory-based models predicted interactions between the font size and the particular digit, and the algorithmic model did not. Contrary to the algorithmic model, font size interacted with digits for both reaction times [$F(7,322) = 21.45$, $MS_e = 360.4$, $p < .01$] and error rates [$F(7,322) = 16.80$, $MS_e = 0.06$, $p < .01$].

The laterality and graded associations models predicted different interactions than did the retrieved instance model. All three memory-based models predicted that performance in the small+letter condition would be better when digits were small and paired with large letters than when digits were large and paired with small letters. The models differ in their predictions regarding the large+letter condition. The laterality and graded associations models predicted that performance in the large+letter condition would be better when digits were large and paired with small letters than when digits were small and paired with large letters. The retrieved instance model, by contrast, predicted no effects due to font size in the large+letter condition.

To pit the predictions of the three memory-based models against each other, these interactions were further analyzed using planned contrasts on performance in the small+letter and large+letter conditions, separately. The coefficients for these contrasts are presented in Table 1. These contrast analyses pitted performance when digits were presented in 30-point type against performance when digits were presented in 40-point type. Consistent with the predictions of all three memory-based models (and contrary to algorithmic processing models), a contrast analysis pitting performance when digits were presented in 30-point type against performance when digits were presented in 40-point type on the results of the small+letter condition (first set of coefficients presented in Table 1) found that subjects were faster judging physical sizes when the small digits were presented in 30-point type ($M = 441$ msec) than when they were presented in 40-point type [$M = 475$ msec; $F(1,46) = 162.43$, $MS_e = 688.9$, $p < .01$]. A contrast analysis using the same coefficients calculated on error rates found that subjects made fewer errors judging physical sizes when the small digits were presented in 30-point type ($M = 1.27\%$ errors) than when they were presented in 40-point type [$M = 4.36\%$ errors; $F(1,46) = 51.20$, $MS_e = 0.2$, $p < .01$].

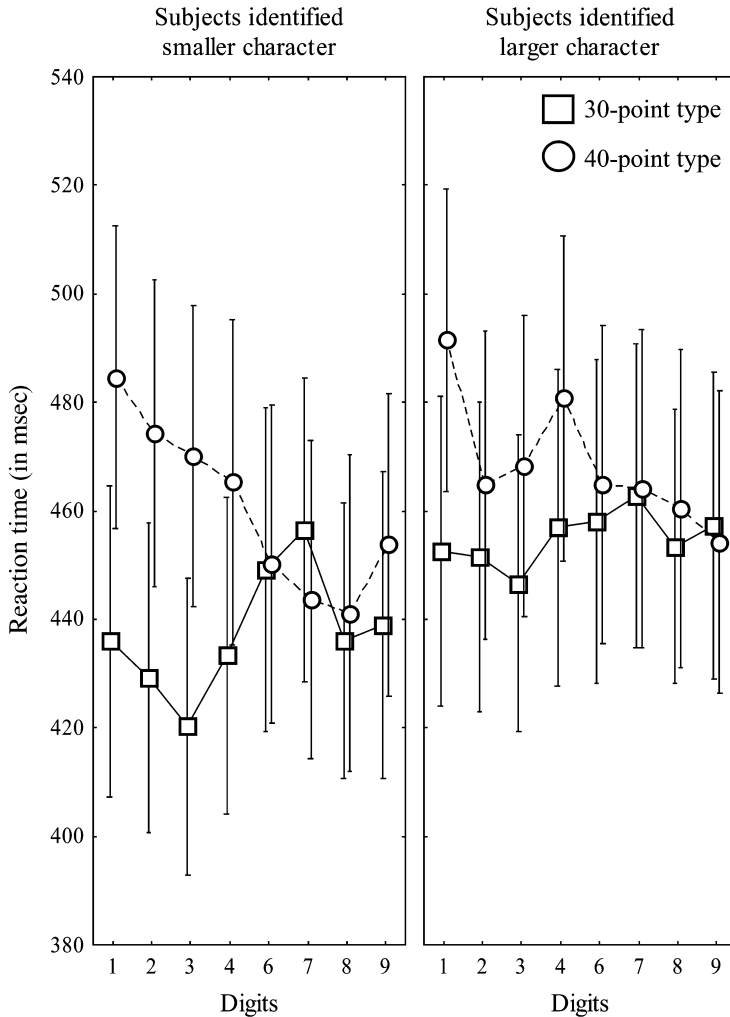


Figure 3. Reaction time results from the small+letter and large+letter conditions. Reaction times to identify the relative sizes of the characters when the digits were presented in 40-point type (and paired with 30-point letters) was greater than the time needed to identify the relative sizes of the characters when the digits were presented in 30-point type (and paired with 40-point letters) for the digits 1–4, but not for the digits 6–9. Vertical bars represent one standard error of the mean.

Consistent with the retrieved instance model and inconsistent with the laterality and graded associations models, these font size effects were only found for small digits, not for large digits. Contrast analyses pitting performance when digits were presented in 30-point type against performance when digits were presented in 40-point type in the large+letter condition using the second set of coefficients presented in Table 1 found that subjects were not faster when the large digits (i.e., 6–9) were presented in 40-point type ($M = 454$ msec) than when they were presented in 30-point type [$M = 452$ msec; $F(1,46) = 1.80$, $MS_e = 341.5$, $p > .05$], nor were they significantly more accurate when the large digits (6–9) were presented in 40-point type ($M = 1.76\%$ errors) than when these dig-

its were presented in 30-point type [$M = 2.01\%$ errors; $F(1,46) = 2.34$, $MS_e = 0.03$, $p > .05$]. Because font size only mattered for the small+letter condition, there was also a main effect of the font size in which the digit was presented such that subjects were on average faster judging the relative physical sizes of digits when they were presented in 30-point type ($M = 446$ msec) than when they were presented in 40-point type [$M = 464$ msec; $F(1,46) = 113.59$, $MS_e = 568.3$, $p < .01$]. An analogous analysis on error rates also found a main effect of the font size in which the digit was presented such that subjects were on average more accurate judging the relative physical sizes of digits when they were presented in 30-point type ($M = 1.64\%$ errors) than when they were

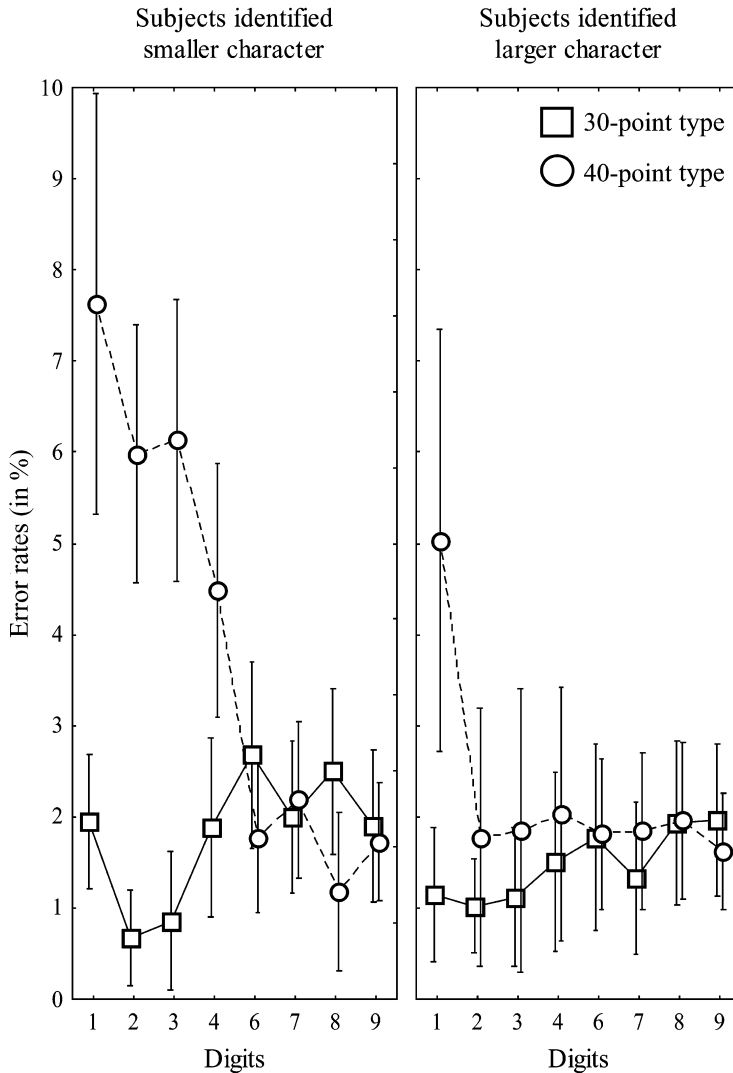


Figure 4. Error rate results from the small+letter and large+letter conditions. Error rates in identifying the relative sizes of the characters when the digits were presented in 40-point type (and paired with 30-point letters) were greater than the error rates in identifying the relative sizes of the characters when the digits were presented in 30-point type (and paired with 40-point letters) for the digits 1–4, but not for the digits 6–9. Vertical bars represent one standard error of the mean.

presented in 40-point type [$M = 3.06\%$ errors; $F(1,46) = 46.65$, $MS_e = 0.08$, $p < .01$].

Further planned contrasts pitting the predictions of the three models against each other revealed that all three memory-based models accounted for significant amounts of variance. A contrast analysis testing the pattern of results predicted by the laterality model in which performance was predicted to be better in the small+letter condition when digits were small than when digits were large and better in the large+letter condition when digits were large than when digits were small, and performance in response to different digits within each of these

conditions were not predicted to differ (using the third set of coefficients presented in Table 1) revealed highly reliable effects [$F(1,46) = 103.78$, $MS_e = 462.1$] for reaction times and [$F(1,46) = 44.23$, $MS_e = 0.1$] for error rates ($ps < .01$).

Similarly, a contrast analysis testing the pattern of results predicted by the graded associations model in which the predicted better performance in the small+letter condition when digits were small than when digits were large was predicted to be greatest for the smallest digit (i.e., 1), followed by the second smallest digit (i.e., 2), and so forth, and the predicted better performance in the

Table 1
Coefficients Used in Contrast Analyses on the Results of the Small+Letter and Large+Letter
Conditions Presented in Figures 3 and 4

	Digit Size	Digits							
		1	2	3	4	6	7	8	9
Font size effects in small+letter condition	Large	1	1	1	1	0	0	0	0
	Small	-1	-1	-1	-1	0	0	0	0
Font size effects in large+letter condition	Large	0	0	0	0	-1	-1	-1	-1
	Small	0	0	0	0	1	1	1	1
Predictions of the laterality model	Large	1	1	1	1	-1	-1	-1	-1
	Small	-1	-1	-1	-1	1	1	1	1
Predictions of the graded associations model	Large	4	3	2	1	-1	-2	-3	-4
	Small	-4	-3	-2	-1	1	2	3	4
Predictions of the retrieved instance model	Large	4	3	2	1	0	0	0	0
	Small	-4	-3	-2	-1	0	0	0	0

large+letter condition when digits were large than when digits were small was predicted to be greatest for the largest digit (i.e., 9), followed by the second largest digit (i.e., 8), and so forth (using the fourth set of coefficients presented in Table 1) also revealed highly reliable effects [$F(1,46) = 104.46$, $MS_e = 411.7$] for reaction times and [$F(1,46) = 51.05$, $MS_e = 0.1$] for error rates ($ps < .01$).

A contrast analysis testing the pattern of results predicted by the retrieved instance model in which performance in the small+letter condition was predicted to be better for the smallest digit (i.e., 1) when it was small than when it was large, followed by the second-smallest digit (i.e., 2), and so forth, but performance in the large+letter condition was not predicted to depend upon the font size of the digit (using the fifth set of coefficients presented in Table 1) also revealed highly reliable effects [$F(1,46) = 165.37$, $MS_e = 634.0$] for reaction times and [$F(1,46) = 58.46$, $MS_e = 0.2$] for error rates ($ps < .01$). Because these contrasts were not orthogonal, we cannot compare them statistically to see which model provides the best fit. To compare the amount of variance accounted for by each contrast, we used partial η^2 . The contrasts for the retrieved instance model had larger partial η^2 values (partial $\eta^2 = .78$ for reaction times and partial $\eta^2 = .56$ for error rates) than did either the contrasts for the laterality model (partial $\eta^2 = .69$ for reaction times and partial $\eta^2 = .49$ for error rates) or the contrasts for the graded associations model (partial $\eta^2 = .69$ for reaction times and partial $\eta^2 = .53$ for error rates). Although we cannot determine whether these differences in partial η^2 values are statistically significant, the retrieved instance model provides the best fit.

The four models (algorithmic, laterality, graded associations, and retrieved instance) also differed in the predictions they made for digits paired with other digits. The algorithmic and the graded associations models predicted that congruity effects always ought to increase with distance, but the laterality and retrieved instance models did not. The laterality model predicted that congruity effects ought to be the same for all bilateral pairs (i.e., the pairs used in this experiment in which one digit from the set of small digits {1, 2, 3, 4} was paired with

one digit from the set of large digits {6, 7, 8, 9}). The retrieved instance model predicted that for bilateral pairs, the size of congruity effects ought to depend upon the small digit in the pair rather than distance per se. To investigate these predictions, we analyzed the results for each of the eight digits used in this experiment separately.

The results for the small digits 1, 2, 3, and 4 when they were paired with the large digits 6, 7, 8, and 9 and letters are presented in Figure 5. The results for each digit were analyzed using 2 (font size: digit in 30-point type, digit in 40-point type) \times 5 (character types with which each digit was paired: 6, 7, 8, and 9, and letters) \times 2 (types of judgment: larger vs. smaller) mixed-factor ANOVAs. These ANOVAs revealed main effects of font size such that reaction times were longer and error rates were higher when the digits 1, 2, 3, and 4 were large and paired with small characters than when they were small and paired with large characters. Reaction times identifying the relative sizes of the digit 1 paired with characters, for example, were longer when the digit 1 was presented in 40-point type and paired with characters presented in 30-point type ($M = 474$ msec) than when it was presented in 30-point type and paired with characters presented in 40-point type [$M = 441$ msec; $F(1,46) = 68.86$, $MS_e = 1,922.7$, $p < .01$]. The same effect was found for the digit 2 [$F(1,46) = 112.62$, $MS_e = 1,305.8$, $p < .01$], the digit 3 [$F(1,46) = 40.86$, $MS_e = 2,056.4$, $p < .01$], and the digit 4 [$F(1,46) = 94.82$, $MS_e = 896.8$, $p < .01$]. Error rates were also greater when these small digits were large and paired with small characters than when they were small and paired with large characters. For example, subjects made more errors when the digit 1 was presented in 40-point type and paired with characters presented in 30-point type ($M = 5.08\%$ errors) than when it was presented in 30-point type and paired with characters presented in 40-point type [$M = 1.64\%$ errors; $F(1,46) = 39.99$, $MS_e = 0.4$, $p < .01$]. The same effect was found for the digit 2 [$F(1,46) = 39.32$, $MS_e = 0.2$, $p < .01$], the digit 3 [$F(1,46) = 19.85$, $MS_e = 0.3$, $p < .01$], and the digit 4 [$F(1,46) = 13.92$, $MS_e = 0.2$, $p < .01$].

Contrary to the algorithmic and graded associations models and consistent with the retrieved instance model,

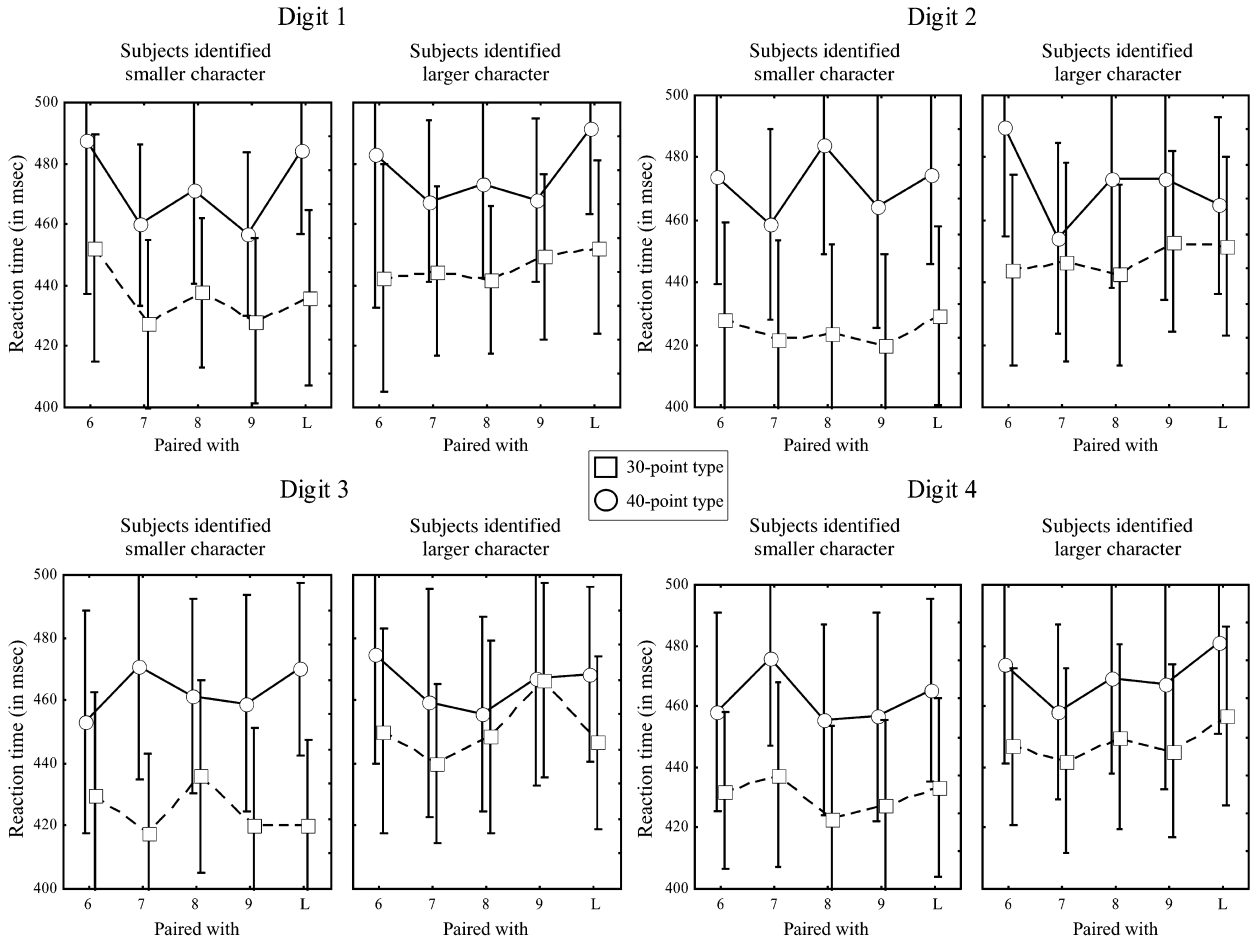


Figure 5. Results for each of the small digits 1, 2, 3, and 4 paired with the large digits 6, 7, 8, and 9 and letters. Interactions between font size and the digits with which each digit was paired were not reliable. Vertical bars represent one standard error of the mean.

these ANOVAs revealed no interactions between font size and characters with which these small digits (i.e., the digits 1–4) were paired either for reaction times [$F(4,184) = 1.22$, $MS_e = 1,300.6$ for the digit 1; $F(4,184) = 2.16$, $MS_e = 1,197.2$ for the digit 2; $F(4,184) = 1.36$, $MS_e = 1,502.1$ for the digit 3; and $F(4,184) = 0.02$, $MS_e = 1,137.0$ for the digit 4, respectively (all $ps > .05$), although the results for the digit 2 might be described as marginally reliable ($p < .1$) or for error rates [$F(4,184) = 2.34$, $MS_e = 0.1$ for the digit 1; $F(4,184) = 1.17$, $MS_e = 0.1$ for the digit 2; $F(4,184) = 1.32$, $MS_e = 0.1$ for the digit 3; and $F(4,184) = 1.12$, $MS_e = 0.1$ for the digit 4, respectively (all $ps > .05$), although the results for the digit 1 might be described as marginally reliable ($p < .1$)]. That is, font size (i.e., congruity) effects were not larger when small digits were paired with the digit 7 than when they were paired with the digit 6 and so forth (despite the fact that the digit 7 is farther away from these digits than the digit 6).

These results are problematic for the algorithmic model. Font size (i.e., congruity) effects were not even larger

when small digits were paired with digits than when small digits were paired with letters. This finding suggests that distance effects are most likely an effect of associations between digits and attributes, rather than an effect of distance per se. Post hoc least significant difference tests using the error terms from the interactions of font size \times characters with which digits were paired revealed that reaction times were significantly longer when these small digits were large and paired with small letters than when they were small and paired with large letters for the digits 1, 2, 3, and 4 among subjects who identified the smaller character and for the digits 1, 2, 3, and 4 among subjects who identified the larger character. Error rates were greater for the digits 1, 2, 3, and 4 presented in the larger font size among subjects who identified the smaller character and for the digit 1 presented in the larger font size among subjects who identified the larger character. This analysis confirms the findings of the analysis of the small+letter and large+letter conditions reported above and presented in Figures 3 and 4. In addition to being problematic for the algorithmic

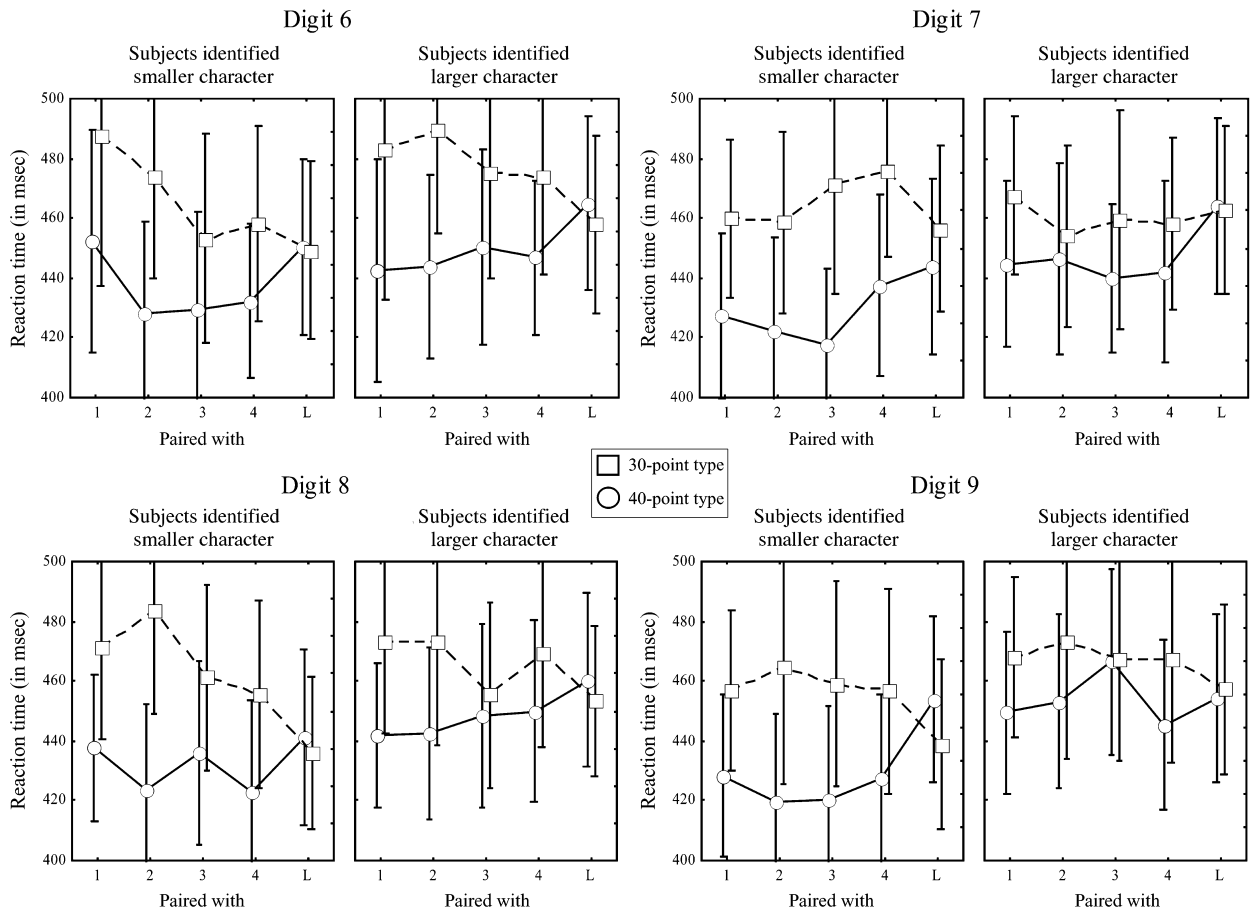


Figure 6. Results for each of the large digits 6, 7, 8, and 9 paired with the small digits 1, 2, 3, and 4 and letters. Interactions between font size and the digits with which each digit was paired were generally reliable. Vertical bars represent one standard error of the mean.

model, the lack of distance effects in this analysis is problematic for the graded associations model.

The results for the large digits 6, 7, 8, and 9 when they were paired with the small digits 1, 2, 3, and 4 and letters are presented in Figure 6. Note that these results are a rearrangement of the data that were presented in Figure 5. For example, the data points for digit 1 paired with digit 6 in Figure 5 are identical to the data points for digit 6 paired with digit 1 in Figure 6. As in the previous analysis, the results for the large digits 6, 7, 8, and 9 were analyzed using 2 (font sizes) \times 5 (character types with which each digit was paired) \times 2 (types of judgment) mixed-factor ANOVAs. Consistent with both the algorithmic model and the memory-based models, these ANOVAs revealed main effects of font size such that reaction times were longer when these large digits were small and paired with large characters than when they were large and paired with small characters. Reaction times identifying the relative sizes of the digit 6 paired with characters, for example, were longer when the digit 6 was presented in 30-point type and paired with characters presented in 40-point type ($M = 470$ msec) than when it was presented

in 40-point type and paired with characters presented in 30-point type [$M = 444$ msec; $F(1,46) = 60.14$, $MS_e = 1,369.3$, $p < .01$]. The same effect was found for the digit 7 [$F(1,46) = 41.62$, $MS_e = 1,644.9$, $p < .01$], the digit 8 [$F(1,46) = 31.03$, $MS_e = 2,007.7$, $p < .01$], and the digit 9 [$F(1,46) = 38.53$, $MS_e = 1,135.8$, $p < .01$]. Error rates were also greater when these large digits were small and paired with large characters than when they were large and paired with small characters. For example, subjects made more errors when the digit 6 was presented in 30-point type and paired with characters presented in 40-point type ($M = 2.87\%$ errors) than when it was presented in 40-point type and paired with characters presented in 30-point type [$M = 1.64\%$ errors; $F(1,46) = 15.91$, $MS_e = 0.1$, $p < .01$]. The same effect was found for the digit 7 [$F(1,46) = 32.11$, $MS_e = 0.2$, $p < .01$], the digit 8 [$F(1,46) = 22.73$, $MS_e = 0.2$, $p < .01$], and the digit 9 [$F(1,46) = 12.51$, $MS_e = 0.2$, $p < .01$].

Unlike the ANOVAs on the small digits, these ANOVAs did reveal interactions between font size (i.e., congruity) and the characters with which each digit was paired for both reaction times [$F(4,184) = 6.74$, $MS_e = 1,294.2$ for

the digit 6; $F(4,184) = 2.52$, $MS_e = 1,217.9$ for the digit 7; $F(4,184) = 7.39$, $MS_e = 1,213.6$ for the digit 8; and $F(4,184) = 3.97$, $MS_e = 1,307.7$ for the digit 9, respectively (all $ps < .05$) and error rates [$F(4,184) = 2.93$, $MS_e = 0.08$ for the digit 6; $F(4,184) = 4.98$, $MS_e = 0.2$ for the digit 8; and $F(4,184) = 5.38$, $MS_e = 0.09$ for the digit 9, respectively (all $ps < .05$); this interaction was not reliable for the digit 7; $F(4,184) = 2.17$, $MS_e = 0.1$, $p = .07$, although it might be described as marginal] such that reaction times were faster and error rates lower when these large digits (i.e., 6–9) were paired with smaller small digits (e.g., 1) than when they were paired with larger small digits (e.g., 4). That is, these analyses found distance effects for the large digits (i.e., 6–9). The analyses on these same results for the small digits (Figure 5) suggests that these interactions are an effect of associations between the small digit (i.e., the presence of 1, 2, 3, or 4) and the attribute “small” rather than an effect of distance per se. The overall pattern of these results is consistent with the retrieved instance model and contrary to the laterality model (which predicted that congruity effects ought not to differ among bilateral pairs) and the graded associations model (which predicted that congruity effects always ought to interact with distance). Consistent with the retrieved instance model and contrary to the laterality and graded associations models, post hoc least significant difference tests using the error terms from the interactions of font size \times character with which each item was paired interactions did not find effects of the sizes of the digits 6, 7, 8, or 9 paired with letters for either reaction times or error rates regardless of whether subjects identified the smaller or the larger character. This analysis confirms the findings of the analysis of the small+letter and large+letter conditions reported above and presented in Figures 3 and 4.

CONCLUSIONS

We investigated whether relational judgments of the numerosities represented by digits (e.g., *larger than* [8,2]) or associations between digits and the attributes “small” and “large” (e.g., the digit 2 is strongly associated with the attribute “small” and weakly associated with the attribute “large”) are responsible for size congruity effects. We reasoned that if—as assumed by the algorithmic model (Dehaene & Akhavein, 1995; Schwarz & Ischebeck, 2003; Tzelgov et al., 2000)—then comparisons between the represented values of digits were responsible for size congruity effects, size congruity effects ought to require two inputs and ought not to occur for individually presented digits. By contrast, if—as assumed by memory-based models (Algom et al., 1996; Pansky & Algom, 1999; Tzelgov et al., 1992)—then associations between digits and the attributes “small” and “large” were responsible for size congruity effects rather than relational judgments, analogous size congruity effects ought to be observable for individually presented digits. We found that congruity effects were just as large

for small digits (1–4) paired with letters as they were for small digits paired with other digits, but were not reliably greater than zero for large digits (6–9) paired with letters. This finding is problematic for the algorithmic model and consistent with memory-based models.

Of the three memory-based models, the results were most consistent with the retrieved instance model, in which the smallest digits are most strongly associated with the attribute “small” and most weakly associated with the attribute “large” (because they are almost always small and almost never large) and larger digits are approximately equally associated with the attributes “small” and “large” (because they are small in some contexts and large in other contexts; see Figure 2). Consistent with this model, congruity effects for the large digits differed as a function of the small digits with which they were paired (see Figure 6), while congruity effects for small digits did not differ as a function of the large digits with which they were paired (see Figure 5). This pattern of results was inconsistent with the laterality model in which the small digits (i.e., 1, 2, 3, and 4) were assumed to be equally associated with the attribute “small” and the large digits (i.e., 6, 7, 8, and 9) were assumed to be equally associated with the attribute “large.” As shown in Figure 6, congruity effects for the large digits differed as a function of the small digit with which they were paired. This pattern of results was also inconsistent with the graded associations model in which the strengths of the associations between digits and the attributes “small” and “large” were assumed to be graded according to each digit’s position among the small digits. As is shown in Figure 5, congruity effects for small digits did not significantly differ as a function of the large digit with which they were paired.

Tzelgov et al. (1992) proposed the laterality model as a supplement to the algorithmic model. In their view, both processes operate to produce size congruity effects. The results of our experiment do not support this dual-process model. If both processes were operating, both would have an effect on response times and error rates. However, size congruity effects were just as large for small digits (i.e., 1, 2, 3, and 4) paired with letters as they were for pairs of digits. Associations between digits and attributes seem sufficient to explain size congruity effects. Postulating effects of algorithmic comparison processes appears to be unnecessary to explain the results of our experiment. Future work will be needed to investigate this question in greater depth.

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.
- DEHAENE, S., & AKHAVEIN, R. (1995). Attention, automaticity, and level of representation in number processing. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *21*, 314-326.
- 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.

- LASSALINE, M. E., & LOGAN, G. D. (1993). Memory-based automaticity in the discrimination of visual numerosity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 561-581.
- LETH-STEENSEN, C., & MARLEY, A. A. J. (2000). A model of response time effects in symbolic comparison. *Psychological Review*, **107**, 62-100.
- LOGAN, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, **95**, 492-527.
- LOGAN, G. D., & ETHELTON, J. L. (1994). What is learned during automatization? The role of attention in constructing an instance. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 1022-1050.
- LOGAN, G. D., & ZBRODOFF, N. J. (2003). Subitizing and similarity: Toward a pattern-matching theory of enumeration. *Psychonomic Bulletin & Review*, **10**, 676-682.
- PALMERI, T. J. (1997). Exemplar similarity and the development of automaticity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **23**, 324-354.
- 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.
- RICKARD, T. C. (1997). Bending the power law: A CMPL theory of strategy shifts and the automatization of cognitive skills. *Journal of Experimental Psychology: General*, **126**, 288-311.
- RICKARD, T. C. (1999). A CMPL alternative account of practice effects in numerosity judgment tasks. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **25**, 532-542.
- 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.
- STRAYER, D., & KRAMER, A. F. (1990). An analysis of memory-based theories of automaticity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 291-304.
- 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.

NOTES

1. Algom and his colleagues' position differs from the memory-based automaticity position in that he assumes that attention is attracted to the irrelevant dimension, whereas memory-based views generally assume that attention to an object is sufficient to allow all of its dimensions to instigate memory retrieval. Both positions share the view that retrieval of information from memory is necessary to produce congruity effects.

2. We thank Maggie Xiong and Jeff Franks for suggesting that we investigate frequencies of comparisons by using Google.com.

3. Schwarz and Ischebeck (2003) presented single digits and had their subjects judge whether the numerosities represented by the presented digits were smaller or larger than 5. Their paradigm therefore involved a comparison. Algom et al. (1996) also presented single digits. Their procedure used both physical size classification tasks and numerical magnitude classification tasks. Our procedure, by contrast, used only physical size classification tasks.

4. The effects of the type of judgment subjects made are orthogonal to the immediate concerns of this article. We include this factor in our analyses and present the results for each type of judgment separately in Figures 3-6 because this factor might be important in its own right, but we leave full explication of the effects of this factor to future work.

(Manuscript received April 29, 2003;
revision accepted for publication May 3, 2004.)