Journal of Experimental Psychology:
Learning, Memory, and Cognition
1995, Vol. 21, No. 2, 314–326

0278-7393/95/$3.00

Attention, Automaticity, and Levels of Representation in Number Processing

Stanislas Dehaene
Institut National de la Santé et de la Recherche Médicale

Rokny Akhavein
University of Oregon

Participants performed same–different judgments for pairs of numerals in 2 conditions: numerical matching (responding "same" to pairs such as 2–TWO), or physical matching (responding "different" to pairs such as 2–TWO). In most cases, a distance effect was obtained, with the *different* responses being slower when the 2 numbers were numerically close together (e.g., 1–2) than when they were further apart (e.g., 1–8). This indicates that numbers were automatically converted mentally into quantities, even when the participants had been told to attend exclusively to their physical characteristics. As postulated by several models of number processing, (e.g., Dehaene, 1992; McCloskey, 1992) Arabic and verbal numerals thus appear to converge toward a common semantic representation of quantities. However, the present results suggest that an asemantic transcoding route might allow for a direct mapping of Arabic and verbal numbers, perhaps by means of a common phonological representation.

The system of numbers is exceptional within written language because two radically different notations can be used to convey exactly the same meaning.[1] The verbal or alphabetical notation, on the one hand, provides a graphemic transcription of the phonemic content of the words (e.g., seven, eighty-six). The Arabic notation, on the other hand, conveys the same meaning using a single or a few arbitrary symbols whose visual appearance is largely unrelated to either meaning or pronunciation (e.g., 7, 86). Both the verbal and Arabic notations are normally mastered and used interchangeably in our written language. In this respect, any literate person may be considered bilingual, and indeed the very same questions that have been raised for bilingualism also apply to the number domain. First, is there a common semantic representation of numbers that is independent of the input notation (McCloskey, 1992)? Second, is there a direct mapping between the Arabic and verbal lexicons that bypasses semantic access (Dehaene, 1992; Deloche & Seron, 1987)? Third, is there a preferred language for numbers, to which numerals in the nonpreferred notation would have to be mentally transcoded before comprehension can take place (Noel & Seron, 1993)?

Stanislas Dehaene, Institut National de la Santé et de la Recherche Médicale (INSERM), Paris, France; Rokny Akhavein, Institute of Cognitive and Decision Sciences, University of Oregon.

We gratefully acknowledge stimulating discussions with Michael McCloskey, Michael Posner, and Gérard Roszavolgyi. This research was supported by INSERM (France), the Centre Nationale de la Recherche Scientifique (France), the National Science Foundation (United States) and by the Office of Naval Research (ONR Grant N00014-89-J3013, United States).

Correspondence concerning this article should be addressed to Stanislas Dehaene, Laboratoire de Sciences Cognitives et Psycholinguistique, Centre Nationale de la Recherche Scientifique (CNRS) and Ecole des Haute Etudes en Sciences Sociales (EHESS), 54 Boulevard Raspail, 75270 Paris, Cedex 06, France. Electronic mail may be sent via the Internet to stan@lscp.msh-paris.fr.

Issues in Number Processing

The evidence to date strongly suggests that the comprehension of Arabic and verbal numerals involves at least partially separate processing pathways. For instance, Anderson, Damasio, and Damasio (1990) described a patient suffering from alexia and agraphia, who, despite having very severe difficulties in reading and writing any word or letter, could still read and write multidigit Arabic numerals and perform numerical calculations. Other similar cases have been described (for review, see Cohen, Dehaene, & Verstichel, 1994). The fact that some of these patients suffered from large lesions of the left hemisphere suggests that Arabic numerals, unlike words, may be processed by the right hemisphere (see Coltheart, 1980). Indeed, some recent brain-imaging studies, using positron emission tomography and high-density recordings of event-related potentials, have suggested that Arabic numerals are processed bilaterally, as opposed to the left-lateralized processing of verbal numerals (Dehaene, in press). Accordingly, all current models of number processing concur in depicting distinct input stages for comprehending numbers in Arabic or verbal format (e.g., Campbell & Clark, 1988, 1992; Dehaene, 1992; McCloskey, Caramazza, & Basili, 1985).

The models differ, however, on the convergence of these two input systems. At what point do we know that *2* and *two* represent the same number? According to McCloskey's model (McCloskey, 1992; McCloskey et al., 1985), the separate Arabic and verbal comprehension pathways converge onto a

[1] In written language, variations in case, font, or handwriting also introduce significant variations in the surface form of words while preserving meaning. Indeed, same–different tasks quite similar to the ones used here, in which participants are required, for instance, to match words across variations in case or font type, have also been used to study the architecture of word processing (see, e.g., Posner, 1978). In the case of numbers, however, the structure of the notation system itself varies (i.e., alphabetical vs. logographic) which is a situation somewhat similar to the Japanese Kana and Kanji writing systems (Cohen et al., 1994).

common representation of quantity, which serves as a basis for arithmetic fact retrieval and mental calculation. Similarly, Dehaene's (1992) triple-code model postulates a convergence toward a common analogical representation of quantity, or number line. However, in the triple-code model the Arabic and verbal processing modules are also allowed to communicate directly by means of an asemantic transcoding route. More specifically, Dehaene's model assumes that written numerals connect to a phonological representation of the words and that Arabic numerals may also activate this phonological representation directly, without semantic mediation. Another hypothesis, Noel and Seron's (1993) preferred-entry code, supposes that all numerals are first internally transcoded into a single, preferred concrete notation, either verbal or Arabic, before any further processing may take place. Finally, Campbell and Clark (1988, 1992; Clark & Campbell, 1991) outlined an encoding-complex model that denies the existence of a single, central number representation and suggests an interactive network of specialized format-specific codes, including visual and phonological codes.

The schematic in Figure 1 shows the various levels of representation that have been postulated in these four models. Early visual processes include initial object segmentation and feature extraction. Arabic and verbal input lexicons are used to categorize instances of a specific digit or number word regardless of variations in size, font, or color. A semantic magnitude representation specifies the quantities associated with the numerals. Finally, a phonological representation specifies the sequence of sounds in the pronunciation of the corresponding number words.

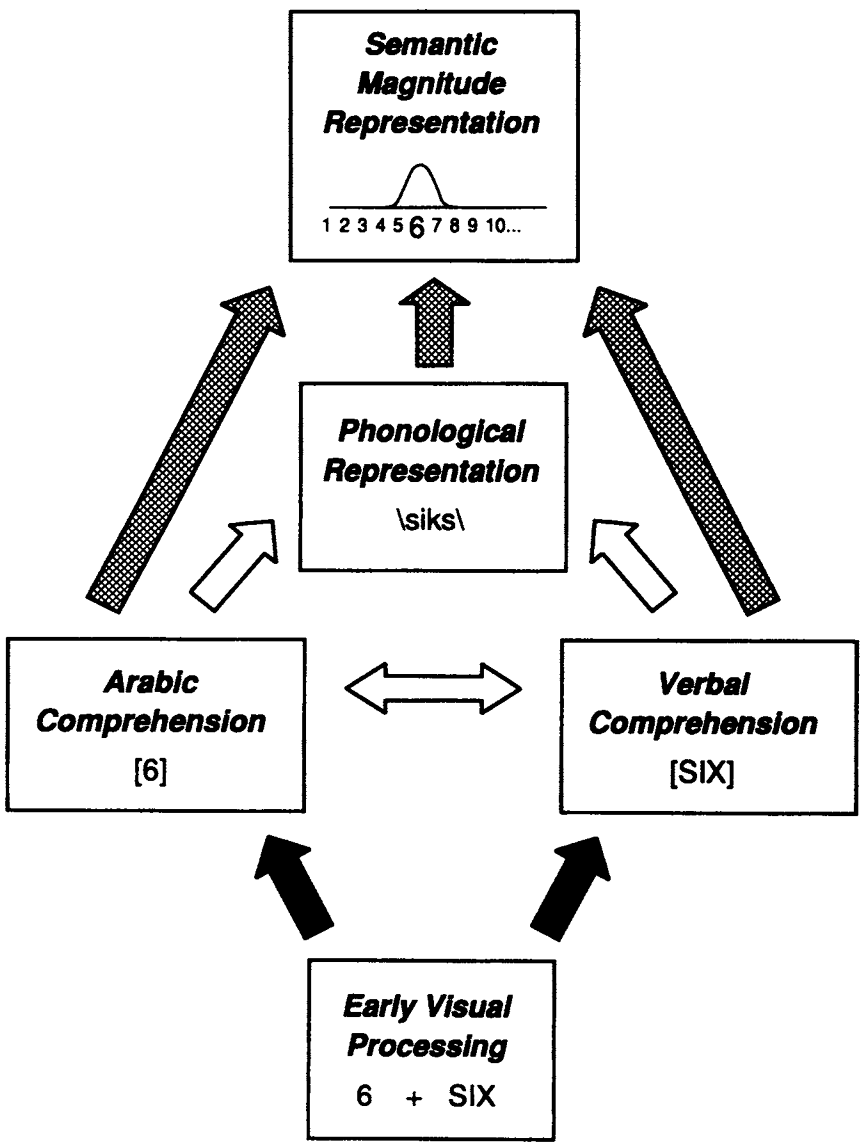

Figure 1. Schematic representation of the putative number processing pathways in a number matching task.

In the present article, we are not concerned about the internal structure of these various representations, so much as we are concerned about the architecture of the pathways by which they communicate with each other. In Figure 1, three levels of shading have been used to distinguish the pathways that are postulated by different models. Black arrows depict the connections from early visual to notation-specific lexicons; gray arrows depict the connections from Arabic, verbal, and phonological representations to the magnitude representation; and white arrows depict the direct connections between the Arabic, verbal, and phonological representations. Of critical importance is the existence of the pathways depicted by the white arrows, which can directly connect Arabic and verbal representations without going through the semantic representation (asemantic transcoding). McCloskey's (McCloskey et al., 1985; McCloskey, 1992) model takes the explicit stance that such asemantic transcoding pathways do not exist. Asemantic transcoding, on the other hand, is a central feature of Noel and Seron's (1993) preferred entry code model and of Dehaene's (1992) triple-code model.

Campbell and Clark's (1988, 1992) encoding-complex model is not as explicit about which pathways are available for number processing. Some of their descriptions imply a fully interactive model, in which all available representations are allowed to communicate directly. For instance, they stated that "number concepts and skills are based on modality- and format-specific mental codes that are interconnected in a complex and highly integrated associative structure" (Campbell & Clark, 1992, pp. 457–458). In this respect, the encoding-complex model is compatible with all the pathways depicted in Figure 1, including asemantic transcoding routes. In fact, a full depiction of the encoding-complex model would probably require many additional modules and pathways because many other number codes are postulated, including articulatory, motor, imaginal, and finger-counting representations. In that sense, Dehaene's (1992) triple-code model can be considered as a restricted implementation of the encoding-complex notion (Dehaene, Bossini, & Giraux, 1993).

At other times, however, Campbell and Clark (1988, 1992) have emphasized differences in the processing of Arabic and verbal numerals. The encoding-complex model assumes that "different number formats can differentially activate internal number representations and associations" (Campbell & Clark, 1992, p. 461). This corresponds to the assumption that the pathways in Figure 1 can have differential associative strengths or even that some of them might not exist. For instance, the phonological representation might be accessed only from numerals in verbal notation, and the magnitude representation might be preferentially accessed from the Arabic representation. Unfortunately, Campbell and Clark have not yet specified which pathways are assumed to be available in the encoding-complex model. Such flexibility renders the model largely untestable in its full generality (Dehaene et al., 1993; McCloskey, 1992; McCloskey, Macaruso, & Whetstone, 1992). Any finding that number processing varies as a function of input or output format could always be accounted for by variations in associative strength, but the absence of variations with format would also be compatible with the model.

Regardless of these assumptions, recent data indicate that format effects on number processing, if they exist, are relatively infrequent and minor. The bulk of the empirical evidence so far mostly supports the existence of a largely format-invariant processing architecture for numerals (for a critical evaluation of possible exceptions, see Campbell & Clark, 1992; McCloskey et al., 1992). Dehaene et al. (1993) found virtually parallel reaction times (RTs) for parity judgments of Arabic and verbal numerals. Likewise, Noel and Seron (1992) found few differences in the processing time of Arabic and Roman numerals, all of which could be attributed to an initial comprehension stage. Patient P.S., described by Sokol, McCloskey, Cohen, and Aliminosa (1991), made the same number and type of calculation errors regardless of the input notation. Finally Dehaene (in press), in a larger–smaller comparison task, found parallel RTs to Arabic and verbal stimuli, and the simultaneously recorded brain event-related potentials gave evidence for the convergence of activation toward a common right posterior brain area involved in the processing of quantities. In the present article, we assessed putative variations in number processing as a function of format in yet another task in which participants had to decide whether two numerals were the same or different. As we report, we again found very little variation in one index of semantic processing (the distance effect) when the numerals were presented in Arabic or in verbal notation.

Thus, at present most data suggest a considerable degree of convergence between the processing of Arabic and verbal numerals. It is clear that Arabic and verbal numerals are not processed by completely distinct notation-specific semantic and processing systems (Dehaene et al., 1993). The critical issue then becomes the following: At which level does the convergence between Arabic and verbal pathways occur? Does it necessarily occur only at the level of a unique semantic representation of magnitudes (McCloskey, 1992)? Is it based on more direct asemantic transcoding routes linking the Arabic and verbal modules, perhaps by means of an intermediate phonological representation (Cohen et al., 1994; Dehaene, 1992; Noel & Seron, 1993)? Or does it occur at several levels at once, with contributions from multiple articulatory, analogical, imaginal, or other numerical codes (Campbell & Clark, 1988, 1992)? As a first step toward answering these questions, we attempted to decide whether there exist any asemantic pathways linking the Arabic and verbal input lexicons.

Automaticity in Number Processing

One potential difficulty confounding the asemantic transcoding issue is that semantic access may proceed automatically and unintentionally, regardless of the task. Automatic semantic activation might then mask the function of the putative asemantic transcoding routes and might even give the impression that they do not exist. Therefore, a second goal of our work was to study the conditions under which an automatic semantic activation can be observed with Arabic and verbal numerals.

Several studies have suggested that semantic access is largely automatic for Arabic numerals. Number magnitude, even when it was irrelevant, has often been found to have an unexpected and sometimes deleterious effect on performance. For instance Henik and Tzelgov (1982) and Tzelgov, Meyer, and Henik (1992) found Stroop-like interference from number magnitude in a task in which the participants were asked to compare the physical sizes of digits. Likewise, Sudevan and Taylor (1987) found an interference of number magnitude on parity judgements. In their study, number magnitude was not totally irrelevant because on some trials the participants had to perform a number comparison task. However Dehaene et al. (1993) reported an effect of number magnitude in a pure task of parity judgements of verbal and Arabic numerals: Large numbers were responded to faster with the right-hand key than with the left-hand key, and the converse held for small numbers (the spatial–numerical association or SNARC effect). Number magnitude was irrelevant to the parity judgment task, so the effect was tentatively explained as an automatic activation of magnitude information encoded on a left-to-right oriented number line.

The numerical distance effect also appears as a salient marker of semantic number processing. The time to compare two numbers is an inverse function of the numerical distance between them (Moyer & Landauer, 1967). Numbers that are close in magnitude, such as 8 and 9, are harder to discriminate than numbers that are further apart, such as 2 and 9. It has been postulated that the internal representation of number magnitude is analogical and that the semantic representations of close numbers such as 8 and 9 overlap more than those of more distant numbers (Dehaene & Changeux, 1993; Gallistel & Gelman, 1992).

A numerical distance effect is found in several nonsemantic number tasks, suggesting that the mere presentation of an Arabic digit may suffice to activate its semantic representation. In Morin, Derosa, and Stultz's (1967) short-term memory study, after memorizing a set of consecutive digits such as 8, 5, 7, and 6, participants had to decide whether a probe digit was in the set or not. Responses were slower when the probe was outside but close to the set (e.g., 4) than when it was more distant (e.g., 1; see also Dehaene & Cohen, 1991). In an experiment involving a similar short-term memory task, LeFevre, Bisanz, and Mrkonjic (1988) showed that the mere presentation of a set of two digits such as 2 4 prompted an automatic activation of the addition $2 + 4 = 6$. Participants were slower to decide that a probe digit did not belong to the set when it was the addition result (6) than when it was not (e.g., 7). This finding was recently replicated by Lemaire, Barret, Fayol, & Abdi (1994).

In a digit naming task, Marcel and Forrin (1974) found that the amount of priming from one trial to the next was a continuous function of the numerical distance between the two digits. A similar variation of priming with numerical distance in a primed digit–letter classification task was reported by den Heyer and Briand (1986). Finally, Duncan and McFarland (1980, Experiment 2) found a distance effect in a same–different judgment task. Participants were presented with a pair of Arabic digits, and they had to decide whether the digits were identical or different. Even though they had to respond "different" to both close and distant pairs of digits (e.g., 1 2, 1 9), they were slower with the close pairs.

Outline of the Present Experiments

The present experiments may be viewed as extensions of Duncan and McFarland's (1980) same–different task. We asked participants to perform same–different judgments with pairs of Arabic numerals (e.g., 2 2), but also with pairs of verbal numerals (TWO TWO) and with mixed pairs (TWO 2 or 2 TWO). The same–different task is naturally ambiguous for mixed pairs. At the physical level, 2 and TWO are different stimuli, yet at the semantic level they represent the same number. Thus, two distinct same–different tasks were defined, which varied the level of representation to which the participant had to attend. In the numerical matching task, participants were told to attend to number meaning and therefore to respond *same* to pairs such as 2 TWO. In the physical matching task, they had to attend to the physical or lexical characteristics of the stimuli, and therefore they had to respond *different* to these pairs.

This design permitted us to simultaneously address questions of automaticity and asemantic transcoding in number processing. The critical issues for automaticity was whether a distance effect could be obtained regardless of the instructions and whether it could be obtained even in the physical matching condition in which the participants were trying to ignore semantic similarity. In this sense, we examined only the autonomy of semantic access, which is a restricted criterion for automaticity that concerns whether a mental process can proceed without intention (Zbrodoff & Logan, 1986). Experiment 1 introduced the numerical matching task. In Experiment 2, a new group of participants was tested with the same stimuli as in Experiment 1, but with the physical matching instructions. Thus, Experiments 1 and 2 permitted a direct test of the effect of instructions on the size of the distance effect. Finally, the entire design was replicated in Experiment 3, in which the numerical and physical matching tasks were compared in two new groups of participants and with a new set of number pairs.

The critical issue for asemantic transcoding was whether the participants could ever match an Arabic numeral to its corresponding verbal form without showing any evidence of semantic activation. In the present experiments, we considered the absence of any distance effect for mixed Arabic and verbal pairs, together with consistent Arabic–verbal matching (e.g., good performance in numerical matching of mixed pairs), to be evidence of asemantic transcoding.

In both cases, we used the distance effect as an index of semantic access. It is not inconceivable, however, that the distance would arise at least in part from the lexical level of processing rather than from the semantic level. For instance, each numeral could be associated with the next one within the lexicon, and spreading activation could then yield a distance effect. Experiment 3 was specifically designed to separate lexical and semantic contributions to the distance effect.

Experiment 1: Numerical Matching

In Experiment 1, participants were presented with pairs of numbers that could be written either in the same notation (*pure trials,* e.g., 8 9, ONE ONE) or in different notations (*mixed trials,* e.g., 8 EIGHT). Their task was to ignore superficial differences in notation and to judge whether the two stimuli represented equal or different numbers. The distance effect was used to assess the maximum level at which the numerals were processed in this task. We reasoned that if the numerals were processed all the way up to semantics, then semantic similarity would interfere with the same–different judgment. *Different* responses would be slower when the two numbers were numerically similar (e.g., 8 NINE) than when they were numerically far apart (e.g., 8 ONE). However, we reasoned that if there was a direct asemantic communication route between the Arabic and verbal lexicons, and if the participants were able to use it to recognize that 8 and EIGHT are identical before the meaning of these symbols was retrieved, then there would be no variation in the speed of responding *different* as a function of numerical distance.

Method

Participants. Ten University of Oregon students (7 of whom were women) participated and received payment. Their ages ranged from 18 to 23 years. All were right handed and were native speakers of English.

Apparatus. A Macintosh IIci (Apple Computers, Cupertino, California) was used to present the stimuli and record the RTs. Responses were recorded by means of two large Morse keys, 26 cm apart, and timed with millisecond accuracy using a NB-DMA-8 card (National Instruments, Austin, Texas). Stimulus presentation was synchronized with the 15-ms refresh rate of the computer monitor.

Procedure. Participants were seated about 50 cm from the computer screen and kept their eyes fixed on a small central cross throughout the experiment. On each trial, a pair of white numerals was flashed for 195 ms onto a black background. The numerals were centered on two points located 35 mm to the left and right of the central cross (4° of visual angle). Arabic digits subtended 6 mm × 8 mm (0.69° × 0.92°), and the largest verbal numeral (SEVEN) subtended 29 mm × 8 mm (3.32° × 0.92°). Arabic digits and capitalized verbal numerals were displayed in a Geneva bold 24-point font.

Participants were instructed to press one key if the two numerals represented the same quantity (e.g., 1 and ONE), and another key if they represented different quantities (e.g., 1 and 8). A few examples of pairs of numerals were presented, and it was stressed that participants sometimes had to respond *same* to numerals that were presented in different notations (e.g., 8 and EIGHT). Instructions emphasized both speed and accuracy.

Two blocks were used, each comprising 242 trials. In one block, the *same* response was assigned to the left key, and the *different* response to the right key. This assignment was reversed in the second block. The order of the two blocks was counterbalanced between participants. With an interstimulus interval of 2,700 ms, the entire experiment lasted less than 30 min.

In the experimental pairs, only the numbers 1, 2, 8, and 9 were used, and two variables were systematically varied: disparity between the two numerals and number notation. Three levels of disparity[2] were used with equal frequency: equal (distance = 0), close (distance = 1), and far (distance > 5). The actual pairs used were 1 1, 2 2, 8 8, and 9 9 for equal trials; 1 2, 2 1, 8 9, and 9 8 for close trials; and 1 8, 1 9, 2 8, 2 9, 8 1, 9 1, 8 2, and 9 2 for far trials. To equalize the overall frequency of equal, close, and far trials, we presented the four equal pairs and the four close pairs four times each in each condition of notation and presented the eight far pairs only twice each. This difference in the frequency of presentation of individual close and far pairs, however, could only go against the expected distance effect because the far pairs, which were presented the least often, were expected to yield the fastest responses. Finally, because the participants responded *different* to both close and far pairs, *same* responses and *different* responses constituted only one-third and two-thirds, respectively, of the responses in each block. This frequency difference is likely to affect the relative speed of *same* and *different* responses, but not the critical comparison of close versus far trials on which our discussion is focused.

Each number in each pair could be presented in either Arabic or verbal notation. Number notation, therefore, defined four categories of trials: pure Arabic (e.g., 1 1), pure verbal (e.g., ONE ONE), mixed Arabic–verbal (e.g., 1 ONE), and mixed verbal–Arabic (e.g., ONE 1). These four levels of number notation were crossed with the three levels of disparity, defining 12 trials types that were each presented 16 times, for a total of 192 experimental trials in each block. To prevent the participants from focusing exclusively on the numbers 1, 2, 8, and 9, we mixed these trials with 40 distracting trials using the digits 3–7. Each block also started with 10 random training trials. Both training and distracting trials were ignored in the analysis.

Results

Figure 2 shows RTs (left panel) and error rates (right panel) in Experiment 1 as a function of disparity and notation. Error rate averaged across the two blocks did not exceed 10.1% (average 4.1%). Median correct RTs were computed for each participant in each of the 12 cells of the design. These were then entered into an ANOVA with disparity (equal, close, or far) and notation (A-A, V-V, A-V, or V-A, with A representing Arabic notation and V representing verbal notation) as within-subject variables, and with order of the two blocks as a between-subjects variable. Order did not enter into any significant effect or interactions.

Notation had a significant effect, $F(3, 24) = 51.6, p < .0001$. The four-level variable of notation was decomposed using three contrasts: pure versus mixed notation, A-A versus V-V and A-V versus V-A. RTs on pure trials were 64 ms faster than on mixed trials, $F(1, 8) = 145.8, p < .0001$, and RTs on pure Arabic trials (A-A) were 58 ms faster than on pure verbal trials (V-V), $F(1, 8) = 45.1, p = 0002$. Within the mixed trials, it made no difference whether the left stimulus was in Arabic notation or in verbal notation, $F(1, 8) = 2.0$, *ns.* Follow-up comparisons revealed the following ordering: A-A < V-V < V-A = A-V (Newman–Keuls, $\alpha = .01$).

The main effect of disparity was also significant, $F(2, 16) = 31.7, p < .0001$, as was its interaction with notation, $F(6, 48) = 3.67, p = .004$. Two contrasts were used to analyze disparity effects, one contrasting equal and unequal pairs, and the other examining the effect of distance (close vs. far) within the unequal pairs.

[2] The term *disparity* is used here to avoid confusion with the distance effect. Disparity is entered into analyses of variance (ANOVAs) as a variable with three levels (equal, close, and far), and a significant main effect of disparity may be due to any significant difference between these three types of trials. The distance effect refers specifically to a significant difference between close and far trials, and it is therefore a subcomparison of the disparity effect.

First, the equal versus unequal contrast was significant, indicating that *same* responses were 16 ms faster overall than *different* responses, $F(1, 8) = 7.19, p = .028$. This contrast also interacted with notation $F(3, 24) = 4.42, p = .013$, an effect that could be traced to an Equal versus Unequal × Pure Versus Mixed interaction, $F(1, 8) = 8.58, p = .019$. The responses to equal pairs were 32 ms faster than the responses to unequal pairs in the pure trials, but only 1 ms faster in the mixed trials.

Second and more crucially, there was a highly significant difference between close and far trials, $F(1, 8) = 95.3, p < .0001$. *Different* responses were 42 ms faster when the two numbers were numerically far apart than when they were close in magnitude. This effect reached significance in all four conditions of notation (all $ps < .01$). It tended to be smaller in the pure Arabic condition, but the interaction of notation with the contrast for close versus far did not quite reach significance, $F(3, 24) = 2.39, p = .094$.

A similar ANOVA on error rates showed that most of these outcomes were not affected by a speed–accuracy trade-off. The sole exception was the equal–unequal effect: Participants made more errors on equal pairs than on unequal pairs, $F(1, 8) = 22.6, p = .001$, yet they responded faster to equal pairs than to unequal pairs. Otherwise, there was a significant effect of notation, $F(3, 24) = 14.9, p < .0001$. Decomposition into contrasts indicated that participants made fewer errors on pure trials than on mixed trials (1.9% vs. 5.5%), $F(1, 8) = 65.2$, $p < .0001$, but that there was no difference between A-A and V-V, nor between A-V and V-A. Finally, there was no significant difference between the error rates for close and far trials, $F(1, 8) < 1$.

Discussion

In analyses of RTs, a significant distance effect was observed in all conditions of notation in Experiment 1. This strongly suggests that number magnitude was accessed during the task.

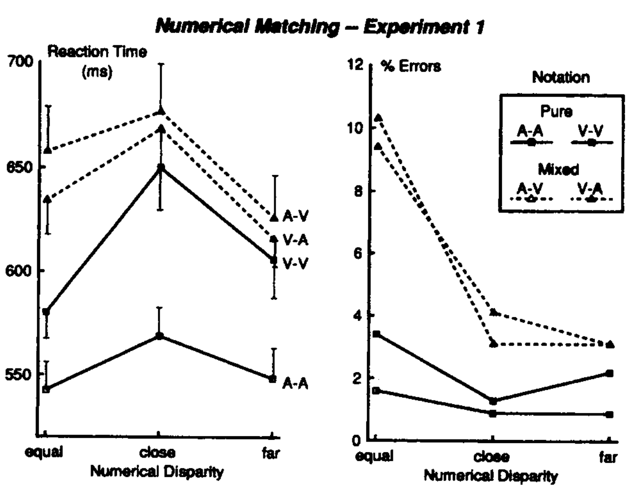

Figure 2. Reaction times and error rates in Experiment 1 as a function of notation (A = Arabic, V = verbal) and of numerical distance.

The distance effect remained highly significant even in the pure trials, in which the participants could conceivably have used physical or lexical identity as a cue for responding. Pure pairs were processed faster and more accurately than mixed pairs, and there was also a slight bias for responding *same* on pure trials and *different* on mixed trials, indicating some degree of influence of physical similarity on same–different judgments. However, the participants evidently were not able to rely exclusively on presemantic levels of processing.

Three minor additional effects confirmed that a semantic representation was accessed. First, the mental representation of number magnitudes is known from previous studies to obey Weber–Fechner's Law (Dehaene, 1992; Dehaene & Changeux, 1993; Krueger, 1989): For equal numerical distance, it is easier to discriminate small numbers (e.g., 1 vs. 2) than larger numbers (e.g., 8 vs. 9). Such a magnitude effect was found in the present experiment. When only equal and close trials were analyzed, RTs were significantly slower for trials using the numbers 8 and 9 (e.g., 8 8, 8 9) than for trials using the numbers 1 and 2 (e.g., 1 1, 1 2), $F(1, 9) = 16.3, p = .003$, and this effect held regardless of notation, interaction $F(3, 27) < 1$. Note, however, that number magnitude is highly correlated with digit or word frequency (Dehaene & Mehler, 1992). Thus, the magnitude effect should perhaps be attributed to frequency-sensitive lexical processes rather than to a semantic representation of numbers.

Second, previous studies have demonstrated that number magnitude maps naturally onto spatial coordinates. For participants who read from left to right, large numbers are automatically associated with the right-hand side of space and small numbers with the left-hand side (SNARC effect; Dehaene et al., 1993). In the present experiment, a somewhat similar influence of the left–right ordering of the numbers was found. On far trials, responses were significantly faster when the smaller number appeared on the left and the larger number appeared on the right (e.g., 1 NINE), as compared with the converse situation (e.g., 9 ONE), $F(1, 9) = 9.33, p = .014$. Again, this effect did not vary with notation, $F(3, 27) < 1$. The ordering effect was not found on close trials (e.g., 1 TWO vs. 2 ONE), presumably because larger–smaller relations were less obvious.

Third, our semantic knowledge of numbers is not limited to magnitude information. For instance, we may know that 8 is even, a power of 2, and so on. Previous studies suggest that semantic information about parity, for instance, is directly retrieved from an internal representation of numbers in Arabic format (Dehaene et al., 1993). In the present experiment, there was indeed an influence of number parity on the speed of same–different judgments. Half of the far trials had numbers of the same parity (1 9, 2 8), and half had numbers of different parities (1 8, 2 9). On pure trials, participants were slower to respond *different* when the parities matched than when they did not match, $F(1, 9) = 6.81, p = .028$, but this was not the case for mixed trials, interaction $F(1, 9) = 7.27$, $p = .025$. The parity matching effect was significant for A–A pairs, $F(1, 9) = 6.81, p = .028$, and did not reach significance for other notations (V-V, A-V, or V-A; all $Fs < 1$). However, the size of the effect did not differ significantly for A-A versus

V-V pairs (24 ms vs. 12 ms; $F < 1$). Thus, the data did not quite afford the conclusion that parity matching was a notation-specific effect, even if they tended to confirm the special role of Arabic notation in accessing parity knowledge (Dehaene et al., 1993).

The distance, magnitude, left–right ordering, and parity effects all suggested that numbers were processed all the way up to semantics in the same–different task. At first sight, this result might suggest that the direct asemantic transcoding routes depicted in Figure 1 did not exist and that the participants were obliged to access a semantic representation to recognize, for example, that 2 and TWO represent the same number. Alternatively, however, an asemantic route might have been available to the participants, but they would have been unable to rely exclusively on it if semantic access was automatic. To study the automaticity of semantic access in number processing, we examined in Experiment 2 whether a distance effect could still be obtained when participants were discouraged from attending to number meaning.

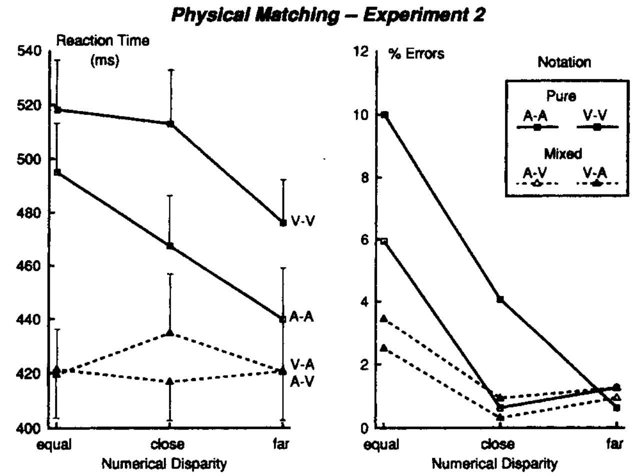

Figure 3. Reaction times and error rates in Experiment 2 as a function of notation (A = Arabic, V = verbal) and of numerical distance.

Experiment 2: Physical Matching

Experiment 2 used the same pairs of numerals that were used in Experiment 1, but the participants were asked to base their judgments solely on the physical similarity of the two targets. They were asked to respond *same* whenever the two stimuli were exactly the same (e.g., 1 1, ONE ONE) and to respond *different* otherwise. Because they now had to respond *different* to equal pairs, such as TWO 2, they were forced to focus attention to the physical or lexical levels of processing and to try to avoid interference from the semantic level.

Again, the distance effect was used to probe whether semantic access could thus be prevented or whether it proceeded automatically and unintentionally. If participants were still slower with numerically similar numbers (e.g., 8 9) than with more distant pairs (e.g., 8 1), this would indicate that they could not prevent accessing the meaning of these symbols.

Method

The stimuli and apparatus were identical to those used in Experiment 1. Only the instructions differed: Participants were instructed to press one key if the two targets were identical or another key if there was any physical difference between them. Instructions stated that participants would occasionally have to respond *different* to numerals representing the same quantity (e.g., FOUR and 4). Because the design was the same as in Experiment 1, only one-sixth of the trials now required a *same* response. This was different from Experiment 1, in which one-third of trials required a *same* response, and precluded any direct comparison of the same–different effect across experiments. Our analyses focused on the close versus far comparison, which used exactly the same stimuli and responses as in Experiment 1. Ten right-handed native speakers of English (6 of whom were women) participated in Experiment 2 and received payment. All were University of Oregon students ranging in age from 18 to 24 years; none had participated in the previous experiment.

Results

Figure 3 shows RTs (left panel) and error rates (right panel) in Experiment 2 as a function of disparity and notation. Error rate averaged across the two blocks did not exceed 4.4% (average 2.7%). Median correct response times for each participant within each cell were entered into an ANOVA similar to the one used in Experiment 1. Again, the order of the blocks did not have any effect, aside from an obscure Order × Notation interaction, $F(3, 24) = 3.7, p = .026$, indicating a slightly smaller difference between pure and mixed trials for participants who started in the same–right condition.

Notation had a significant effect, $F(3, 24) = 76.1, p < .0001$. Decomposition into contrasts showed that participants now responded 63 ms faster on mixed trials than on pure trials, $F(1, 8) = 147.2, p < .0001$, and that they responded 35 ms faster on pure Arabic trials than on pure verbal trials, $F(1, 8) = 41.0$, $p = .0002$. There was no difference between A-V and V-A trials, $F(1, 8) < 1$). Follow-up comparisons revealed the following ordering: A-V = V-A < A-A < V-V (Newman–Keuls, $\alpha = .01$).

The main effect of disparity was also significant, $F(2, 16) = 8.74, p = .0027$, as was its interaction with notation, $F(6, 48) = 4.94, p = .0005$. The latter interaction reduced to two interactions of simple contrasts. First, there was an interaction of Equal Versus Unequal × Pure Versus Mixed, $F(1, 11) = 16.7$, $p = .0035$. On mixed trials, equal and unequal pairs were responded to at the same speed (3 ms difference), whereas on pure trials, equal pairs were responded to 33 ms more slowly than unequal pairs, $F(1, 8) = 15.1, p = .005$. Note that participants responded *same* only on the one-sixth of trials with equal pairs in pure notation. Thus, it is perhaps not surprising that infrequent responses should be slower.

The second and most important effect was an interaction of the contrasts for close versus far and for pure versus mixed, $F(1, 11) = 10.5, p = .012$. There was no distance effect on mixed trials (5 ms difference). However, on pure trials close pairs were responded to 32 ms more slowly than far pairs, $F(1, 8) = 13.9, p = .006$. This distance effect did not differ for A-A versus V-V pairs, $F(1, 8) < 1$, and it reached significance in both cases (both ps $< .015$).

A similar ANOVA on error rates showed that none of these effects were affected by a speed–accuracy trade-off. Notation had a significant effect, $F(3, 24) = 4.34, p = .014$, replicating the ordering A-V = V-A < A-A < V-V. Disparity also had a significant effect, $F(2, 16) = 21.4, p < .0001$, and interacted with notation, $F(6, 48) = 2.60, p = .029$. This was due largely to a marginal Equal Versus Unequal × Pure Versus Mixed interaction, $F(1, 11) = 4.12, p = .077$. Participants made 6.4% more errors in response to equal pairs than to unequal pairs on pure trials, $F(1, 8) = 19.7, p = .002$, whereas the difference between equal and unequal pairs was only 2.1% on mixed trials, $F(1, 8) = 3.33, p = .105$. The contrast for close versus far did not reach significance in the error analysis, nor did it interact with any other contrasts.

Discussion

Automaticity of the distance effect. A distance effect was again observed in RTs on the pure trials of Experiment 2, but not on the mixed trials. On mixed trials, the stimuli were highly dissimilar: A single Arabic digit was compared with a three- or four-letter verbal number word. Thus, it is perhaps not surprising that the participants could respond *different* rapidly and before any semantic interference could take place. However, when the two numerals were written in the same notation (both Arabic or both verbal), participants were slower to respond *different* when the two numbers were numerically close than when they were numerically far apart. This distance effect suggests that there was interference from number semantics.

To directly test if instructions had any effect on the size of the distance effect, we entered the data from Experiments 1 and 2 into a single ANOVA with notation, disparity, and instruction (numerical vs. physical matching) as variables. We only report here interactions involving the distance contrast (close vs. far) because the responses to equal pairs were affected by the variable frequency of *same* responses depending on the instructions. There was a significant triple interaction between Close Versus Far, Pure Versus Mixed, and instruction, $F(1, 18) = 11.9, p = .003$. As expected, the distance effect on mixed pairs was significantly modulated by instructions, $F(1, 18) = 22.2, p = .0002$: It essentially disappeared in the physical matching task. The distance effect on pure pairs, however, was not affected by instruction, $F(1, 18) < 1$: It was of exactly the same size whether the instructions specified numerical or physical matching (32 ms in both cases). This suggests that attention had little influence on access to number meaning and provides evidence for the autonomy of semantic access in number processing.[3]

Confirming that a representation of number magnitude was accessed, the magnitude effect of Experiment 1 was replicated on pure trials: On equal and close trials, responses were marginally faster when the numbers were small (e.g., 1 1, 1 2) than when they were large (e.g., 8 8, 8 9), $F(1, 9) = 3.66$, one-tailed $p = .044$. This magnitude effect was found on pure trials, but not on mixed trials, interaction $F(1, 9) = 4.23$, one-tailed $p = .035$, which confirms that the *different* responses on mixed trials were initiated without semantic interference. The other minor effects of left–right ordering and of parity matching did not reach significance in Experiment 2.

Evidence for asemantic transcoding. The seemingly automatic access to number semantics creates difficulties for investigating the functioning, or even the existence, of direct asemantic transcoding routes linking the Arabic and verbal lexicons. It implies that even if such routes exist, they are often masked or overridden by the semantic route. In the present context, to obtain evidence in favor of asemantic transcoding, we would have to observe a significant difference between *same* and *different* responses in the absence of any difference between *close* and *far* responses. Such conditions were met only on the mixed trials of the physical matching task: No distance effect was observed, and the participants simply had to respond *different* to all these highly dissimilar pairs. Yet they tended to make more errors of responding *same* when equal numbers were presented (e.g., 2 TWO) than when unequal numbers were presented (e.g., 2 ONE or 2 NINE). This might suggest that 2 and TWO were recognized as representing the same number, thereby producing interference with the physical matching instructions even though number magnitude information was not accessed. Thus, asemantic transcoding might be dissociable from semantic access in this situation. However, the error trend did not quite reach significance in Experiment 2 (p = .105). Experiment 3 attempted to replicate this error trend with different pairs of numerals.

Origin of the distance effect. A crucial hypothesis underlying Experiments 1 and 2 was that the distance effect indicates an activation of number semantics. Because the digits 1 and 2 are not more visually similar than are 1 and 9, it is clear that the distance effect must reflect access to the sequential structure of numbers. However, with the particular stimuli used in those experiments, the effect could still be attributed solely to intralexical associations between consecutive numbers rather than to a semantic representation of quantities. The close pairs in Experiments 1 and 2 comprised only consecutive numbers (1–2 and 8–9), and it is possible that these were associatively connected within the Arabic and verbal input lexicons. Repeating the counting series "one, two, three. . . ," for instance, could impose associations between the lexical entries for one and for two, for two and for three, and so on, thus explaining at a purely lexical level why it would be more difficult to discriminate one from two than one from nine. The hypothesis of a lexical origin for the distance effect would readily explain the absence of any distance effect in the mixed trials of the physical matching task. Given distinct input lexicons for Arabic and verbal numerals, there would be no direct associative

[3] An alternative explanation of the distance effect found in the pure pairs in physical matching would be that when the participants perceived two quite similar stimuli, they actively processed them semantically before deciding that they were the same or different. Thus, the distance effect would not reflect an automatic interference from semantics but rather the active disambiguation of two physically similar stimuli by semantic access. This interpretation, however, can be refuted by comparing RTs to pure pairs in the physical and numerical matching tasks. Pure pairs were actually responded to faster in physical matching than in numerical matching, $F(1, 18) = 19.4, p = .0003$, which seems incompatible with the aforementioned notion of a sequential analysis of physical and semantic similarity in the physical matching task.

connections between, for example, the lexical entry for 1 in Arabic notation and the lexical entry for two in verbal notation.

The lexical hypothesis is not completely devoid of difficulties. First, it cannot explain the complete invariance of the distance effect with notation. If the distance effect was due to associations within the Arabic and verbal lexicons, it would be a remarkable coincidence that it should be of the same size regardless of notation. Attributing the distance effect to a semantic magnitude representation independent of input notation, however, readily accounts for the observed invariance. Second, some ad hoc hypotheses are required to explain the presence of a distance effect with mixed pairs in the numerical matching condition. One possibility is that, given that the participants had to respond *same* to mixed pairs such as 1 ONE, they would have actively transcoded the content of one lexicon to the other. Thus, they would have been slow to respond to the pair 1 TWO because they would have mentally transcoded the Arabic 1 to the verbal lexical token one and would have then suffered interference from the lexical association between one and two.

Experiment 3 was designed to separate these conflicting lexical and semantic accounts of the distance effect. Both the numerical matching and physical matching conditions were replicated with new and tightly controlled pairs of numerals. First, we used nonconsecutive numbers for the close pairs. This provided a control for direct lexical associations between consecutive numbers. Second, lexical distance was kept constant while numerical distance was varied. This was achieved by taking advantage of the decade structure of numerals. Several experimenters have suggested that the mental lexicons for Arabic and verbal numerals are organized in separate stacks for units, teens, and decades and that within each stack, words are accessed according to their ordinal position (Figure 4; Deloche & Seron, 1984; McCloskey, Sokol, & Goodman, 1986). The evidence for the stack notion comes mostly from neuropsychological data. The reading errors of some brain-lesioned patients are substitutions of number words across stacks that preserve the correct position within stack (e.g., the numeral *14* is read as *forty* or *four*, but not as *thirteen*). Other patients produce numbers from the correct stack but with the wrong position within the stack (e.g., the numeral *14* is read as *thirteen* or *sixteen*, but not as *forty*).

These data indicate that the lexical structure of number words is not linear but probably resembles more the two-dimensional matrix depicted in Figure 4. The neuropsychological data indicate that the lexical entry for the number *twenty* is more easily confused with *twelve* or with *two* than with *eighteen*. Thus, it is possible to select pairs of number words, such as *eighteen* and *twenty*, that are quite close in terms of number magnitude but not so closely related within the lexicon. Using such cross-decade pairs enabled us to contrast lexical and semantic distance in Experiment 3.

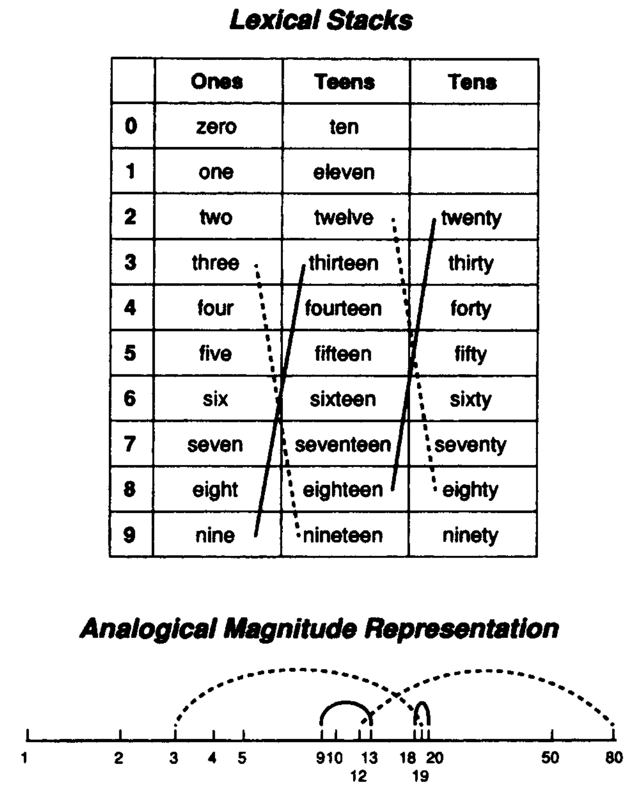

Figure 4. Putative lexical and semantic representations for numbers. The number pairs used in Experiment 3 are depicted on both representations as solid and dashed lines (close and far pairs). Numerical distance was varied, and the lexical distance was kept constant.

Experiment 3: Replication Across Decades

Experiment 3 used pairs of numerals, two close and two far, that were matched for digit or letter content and whose lexical characteristics were tightly controlled. First, the two numbers in each pair belonged to different lexical stacks (e.g., one "teen" and one "tens" word) and were always spaced at least 2 units apart (see Figure 4). If the distance effect in Experiments 1 and 2 was due only to spreading activation within the lexicon, we expected it to be reduced or altogether absent with these new pairs, because it seems less likely for activation to spread across lexical stacks and across nonconsecutive stack positions. In addition, the lexical distance, as measured by the difference in stack positions, was kept constant, and the numerical distance varied considerably between close and far pairs (see Figure 4).

Method

Twenty University of Oregon students (11 of whom were women) participated. All were right-handed native speakers of English between 17 and 28 years of age. Equal numbers of participants were assigned to the physical matching and the numerical matching conditions. None of the participants had been involved in the previous experiments.

The apparatus, procedure, and instructions were the same as in Experiments 1 and 2, with the following exceptions. The experimental pairs of numbers were 3 3, 9 9, 12 12, 13 13, 18 18, 19 19, 20 20, and 80 80 for equal trials; 9 13, 13 9, 18 20, and 20 18 for close trials; and 3 19, 19 3, 12 80, and 80 12 for far trials. Each equal pair was presented twice, and each close or far pair was presented four times in each condition of notation. Each number in each pair could appear in either Arabic or verbal notation. Arabic numerals were presented in the same font, size, and location as in Experiments 1 and 2. Verbal numerals were presented in a slightly smaller font (Geneva bold 18 point; 6 mm or 0.69° high). Their width ranged from 13 to 36 mm (1.5–4.1°). The 40 additional distractors were drawn from the numbers 1–19 and the decades 30–90, excluding the experimental targets.

Results

Figure 5 shows RTs (left panel) and error rates (right panel) in numerical matching as a function of disparity and notation. Figure 6 displays the same information for physical matching.

Numerical matching. The error rate averaged across the two blocks did not exceed 13.8% (average 6.1%). The median correct response times in each cell were entered into an ANOVA similar to that used in Experiment 1. The order of the two blocks did not enter into any significant effect or interactions. Notation had a significant effect, $F(3, 24) = 38.8, p < .0001$. Contrast decomposition showed that mixed pairs were responded to 60 ms slower than were pure pairs, $F(1, 8) = 113.1, p < .0001$, and that A-A pairs were responded to 83 ms faster than were V-V pairs, $F(1, 8) = 30.6, p = .0006$. Thus, the ordering of conditions was very similar to that in Experiment 1: A-A < V-V = V-A = A-V (Newman–Keuls, $\alpha = .01$).

The main effect of disparity fell short of significance, $F(2, 16) = 2.99, p = .079$, as did its interaction with notation, $F(6, 48) = 1.96, p = .090$. However, the contrast analyses essentially replicated the findings of experiment 1. First, there was an Equal Versus Unequal × Pure Versus Mixed interaction, $F(1, 8) = 5.15, p = .053$, indicating that the difference between *same* and *different* responses was greater in pure trials than in mixed trials. Second, with the *different* responses, RTs to close trials were significantly slower than RTs to far trials (16-ms difference), $F(1, 8) = 6.0, p = .040$. The distance effect was significant within the A-A condition (16-ms difference; $p = .023$), was marginal for the V-V and V-A conditions (18-ms and 28-ms difference, respectively; both $ps = .083$), and

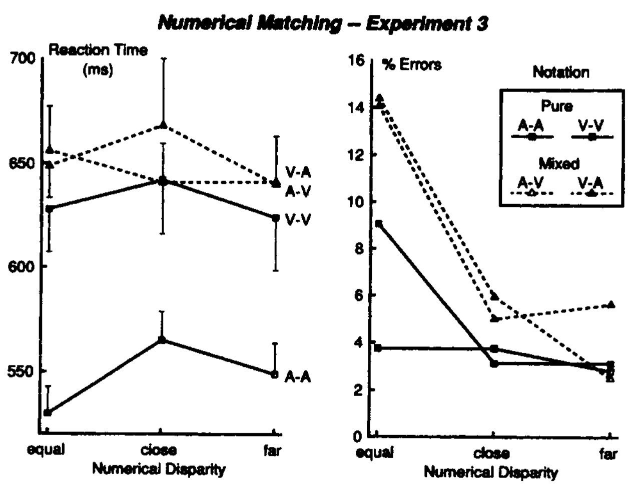

Figure 5. Reaction times and error rates in the numerical matching condition of Experiment 3, as a function of notation (A = Arabic, V = verbal) and of numerical distance.

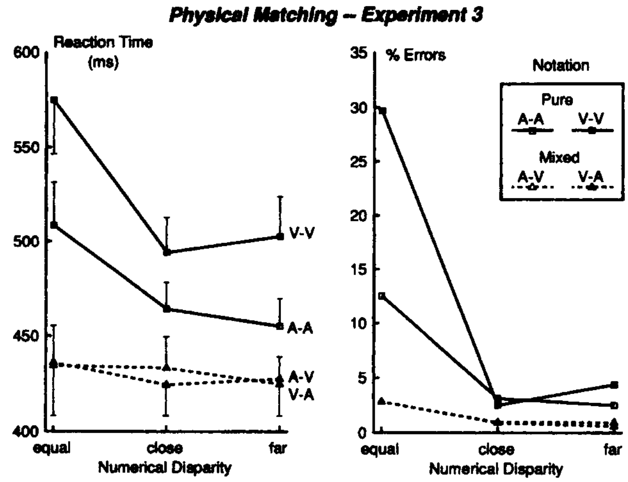

Figure 6. Reaction times and error rates in the physical matching condition of Experiment 3, as a function of notation (A = Arabic, V = verbal) and of numerical distance.

was nonsignificant for the A-V condition (0-ms difference). These differences between conditions did not seem reliable, however, because the distance effect did not interact with notation, $F(3, 24) = 1.21, p = 0.33$ or with any notation-related contrasts (pure vs. mixed, A-A vs. V-V, A-V vs. V-A).

The pattern of errors was also quite similar to that in Experiment 1. The effect of notation was significant, $F(3, 18) = 6.65, p = .003$. Contrast analyses indicated that pure trials yielded fewer errors than did mixed trials, $F(1, 8) = 26.8, p = .0008$, whereas the error rates for A-A versus V-V and for A-V versus V-A did not differ. There was also a main effect of disparity, $F(2, 12) = 12.2, p = .0006$, and a Disparity × Notation interaction, $F(6, 48) = 3.24, p = .009$. Contrast analyses showed that participants made more errors on same trials than on different trials, $F(1, 8) = 14.2, p = .006$, and that this effect was more pronounced on mixed trials than on pure trials, $F(1, 8) = 8.25, p = .021$. There were no significant main effects or interactions involving the close versus far contrast.

Physical matching. The error rate averaged across the two blocks did not exceed 10.9% (average 5.9%). In the ANOVA on median correct response times, the order of the two blocks did not enter into any significant effect or interactions. Notation had a significant effect, $F(3, 24) = 38.4, p < .0001$, and contrast decomposition showed that mixed pairs were responded to 70 ms faster than were pure pairs, $F(1, 8) = 59.4, p = .0001$, and A-A pairs were responded to 48 ms faster than were V-V pairs, $F(1, 8) = 24.5, p = .001$. There was no difference between A-V and V-A. Thus, the ordering of conditions was identical to that in Experiment 2: V-A = A-V < A-A < V-V (Newman–Keuls, $\alpha = .01$).

Disparity had a significant main effect, $F(2, 16) = 9.56, p = .002$, and interacted with notation, $F(6, 48) = 4.21, p = .002$. Contrast analyses revealed a significant interaction between equal versus unequal and pure versus mixed, $F(1, 8) = 7.12, p = .028$, indicating that equal pairs were responded to 63 ms slower than were unequal pairs on pure trials, but only 8 ms slower on mixed trials. This replicated the findings of Experiment 2 that *same* responses, which made up only one-sixth of responses, were slower. There were no main effects or interactions involving the contrast for close versus far (no distance effect).

The ANOVA on error rates showed a main effect of notation, $F(3, 24) = 62.2, p < .0001$. Again, pure trials yielded more errors than did mixed trials, $F(1, 8) = 75.5, p < .0001$, and V-V pairs yielded more errors than did A-A pairs, $F(1, 8) = 46.4, p = .0001$, but there was no difference between A–V and V–A. Disparity also had a significant main effect, $F(2, 16) = 36.1, p < .0001$, and interacted with notation, $F(6, 48) = 39.3, p < .0001$. In the contrast analyses, these effects reduced to interactions of the contrast for equal versus unequal pairs with the contrast for pure versus mixed notations, $F(1, 8) = 62.8, p < .0001$, and with the contrast for A-A versus V-V, $F(1, 8) = 89.1, p < .0001$. Thus, there was an especially large number of errors in the one-sixth of trials in which the stimuli were physically identical and the participant had to respond *same* (up to 30% errors for equal V-V pairs). As in Experiment 2, even within the mixed pairs, the participants tended to make more errors of saying *same* when the two numbers were equal (e.g., 12 TWELVE) than when they were unequal, $F(1, 8) = 4.70, p = .062$. No significant differences were found in the error rates for responses to close and far pairs.

Additional analyses. Because the target numbers did not fall neatly into distinct small and large categories, and because their parities always matched, magnitude and parity effects were not studied in Experiment 3. The effect of the left–right ordering of the numerals did not reach significance in either the physical or the numerical matching conditions, $F(1, 9) < 1$.

Discussion

In the numerical matching condition of Experiment 3, the distance effect was again replicated. Regardless of notation, participants were 16 ms slower to respond *different* when the two numbers were numerically close than were they were further apart. Because the number pairs were constructed to control for lexical similarity, this effect cannot be easily explained by spreading activation within the Arabic and verbal lexicons. Rather, it suggests that participants were accessing a representation of number magnitude, which interfered with their same–different judgments.

The conclusion that semantic distance effects cannot be explained by a simple process of lexical association concurs with previous experiments concerning two-digit number comparison (Dehaene, Dupoux, & Mehler, 1990; Hinrichs, Yurko, & Hu, 1981). When participants had to classify two-digit numerals, such as 49, as larger or smaller than a fixed standard, such as 65, the RTs were affected almost exclusively by a continuous distance effect rather than by the lexical characteristics of the numerals. In the generic case, the distance effect extended continuously both within and across decades, just like in the present experiment. Furthermore, experiments involving French participants showed no effect of the peculiar lexical structure of decade names 70, 80, and 90 in French (Dehaene et al., 1990). The findings from these experiments, together with those from the present experiment, confirm the semantic nature of the numerical distance effect (Moyer & Landauer, 1967).

There may be several plausible and nonexclusive explanations for the reduction of the distance effect from 42 ms in Experiment 1 to 16 ms in Experiment 3. In Experiment 1, the close pairs comprised only consecutive numbers, whereas Experiment 3 used a spacing of 2 (18 vs. 20) or 4 (9 vs. 13) for close trials. This may have facilitated a *different* response to close pairs in Experiment 3, thereby reducing the distance effect. It also seems possible that access to number magnitude was slower or less automatic for the teens and tens number words used in Experiment 3, either because they were physically longer and less frequent or because their larger magnitudes were more difficult to represent mentally (Dehaene & Mehler, 1992). This might also explain the absence of a left–right ordering effect.

This conclusion was strengthened by the observation that no distance effect was found in the physical matching condition of Experiment 3 (see Figure 6). Interpretation of this finding must remain tentative because in a global reanalysis of both the numerical and the physical matching conditions of Experiment 3, the interaction of distance with instruction did not reach significance, $F(1, 18) = 2.41, p = .14$. Nevertheless, it seems that when participants were instructed to attend to physical differences, interference from the numerical properties of the targets tended to disappear. A simple explanation is that the new number pairs were now physically quite different in almost all conditions. Even when both numbers appeared in the same notation, their length and digit or letter content were often very distinct (e.g., 3 vs. 19; EIGHTEEN vs. TWENTY). This dissimilarity could have been used to respond fast, before number magnitude could possibly interfere. Retrospectively, only one set of number pairs, 18–20 vs. 12–80, seemed sufficiently visually similar to preclude fast rejection based on physical dissimilarity. A specific analysis showed that the *different* responses were indeed 44 ms slower with the pair 18–20 than with the pair 12–80, $F(1, 9) = 4.71, p = .058$. This suggests that a distance effect can be obtained with two-digit number pairs, across decades, and in a physical matching condition, provided visual dissimilarity is not too great.

There were also indications of response bias in the physical matching task. The participants made a rather large number of errors of responding *different* when the stimuli were actually identical (up to 30% errors for equal V-V pairs). Note that five-sixths of the trials required a *different* response, and that most such trials were easily classified as different (one or two digits vs. a long string). Thus, participants might have become biased to respond too fast, sacrificing accuracy for the more difficult A-A and V-V equal pairs. This would have contributed to the disappearance of the distance effect. A similar, though less pronounced, response bias was observed in physical matching of single-digit numerals in Experiment 2. However, in Experiment 2, we did observe a distance effect. Thus, the disappearance of the distance effect in Experiment 3 must be caused, at least in part, by some properties of the stimuli, such as their larger magnitude or lower frequency in the language.

Interestingly, even though the absence of a distance effect in physical matching in Experiment 3 suggested little or no access to semantics, there was evidence for communication between the Arabic and verbal lexicons. The participants were very fast and accurate at responding *different* in mixed trials, in which one numeral was in Arabic notation and the other in verbal notation. Yet, as in Experiment 2, they tended to err and respond *same* more often when the two numbers were equal (e.g., 20 TWENTY) than when they were not (e.g., 18 TWENTY). When the results from the physical matching conditions of Experiments 2 and 3 were pooled, this trend was quite significant, $F(1, 19) = 8.27, p = .0097$. Thus, it seems that even when attending to physical differences between the two stimuli, participants could not help noticing the match between Arabic and verbal numerals such as 20 and TWENTY. This comparison was presumably performed by an asemantic mapping of the Arabic and verbal lexicons because no semantic distance effect was observed (e.g., 18 TWENTY did not yield more errors than 12 EIGHTY).

General Discussion

The goal of the experiments presented in this article was threefold: (a) to study whether the number comprehension modules specialized for Arabic and verbal notation eventually converge onto a common semantic representation of number magnitude; (b) to examine whether such semantic access can be modified by attentional instructions; and (c) to probe the existence of more direct asemantic transcoding routes between the Arabic and verbal lexicons.

Convergence Toward a Common Semantics

Experiments 1 and 3 gave evidence that numerals in either Arabic or verbal notation are automatically converted to a common magnitude representation. In the numerical matching task, the distance effect was similar whether the stimuli were two Arabic digits, two verbal number words or one Arabic digit and one number word. Thus, the present experiments add to a growing list of failures to find variations in number processing as a function of input format (Dehaene, in press; Dehaene et al., 1993; Noel & Seron, 1992; Sokol et al., 1991). Although the present experiments indicated convergence toward a common semantic representation, they could not determine whether both Arabic and verbal numerals enjoyed direct semantic access (McCloskey, 1992) or whether the numerals were converted to another format, perhaps phonological, before semantic access (see Dehaene et al., 1993; Noel & Seron, 1993).

Automaticity

The repeated finding of a distance effect in Experiments 1, 2, and 3 indicated that a representation of number magnitude plays a central role in number processing and is automatically accessed whether or not the task requires semantic access. In the numerical matching task, a distance effect was found in all conditions, even in pure pairs such as 2 3 or TWO THREE, in which the participant could have relied on physical cues to respond *different*. With pure pairs, a distance effect was even found in the physical matching task, in which participants were instructed to avoid semantic processing. It thus seems that instructions and attentional set had little or no influence on the activation of number semantics. The only conditions in which the distance effect was not found were when the stimulus pairs were visually quite dissimilar, which allowed for a fast *different* response. Even in this situation, however, we cannot exclude the possibility that semantic activation did occur, but without having time to interfere with the response. Our findings concur with previous reports of automatic semantic access for numbers (Dehaene et al., 1993; Henik & Tzelgov, 1982; Sudevan & Taylor, 1987; Tzelgov et al., 1992).

Asemantic Transcoding

Our results suggest that the participants relied mostly on a semantic magnitude representation to decide whether two numerals represented the same quantity. This raises the issue of whether it is ever possible to compare an Arabic number to a verbal numeral without using semantic cues. In other words, are there direct asemantic transcoding routes between the Arabic and verbal lexicons? Testing this hypothesis was made difficult by the automaticity of semantic activation, which tended to mask lower level processing stages. The only situation in which no distance effect was found was the mixed pairs condition in the physical matching task. In this condition, an analysis of error rates suggested some form of Arabic–verbal transcoding without semantic mediation. Significantly more errors of responding *same* were observed for mixed equal pairs (e.g., 2 TWO) than for mixed pairs that were either close or far (e.g., 2 ONE, 2 NINE). Obviously, the fact that 2 and TWO represent the same number interfered with the participant's decision that these were physically different stimuli. Such interference, however, apparently did not come from semantic mediation because no distance effect was observed. Hence, it seems that an asemantic route should be postulated.

The present evidence for asemantic transcoding is still weak, however, and alternative explanations of our error findings remain plausible. First, it could be that on a small percentage of the physical matching trials, the participants' attention slipped away from the physical level and to the semantic level. This would have yielded a significant increase of erroneous *same* responses, yet might have had no measurable effect on mean RTs. Second, even within the semantic system, same–different relations might be more salient and yield a stronger interference effect than close–far relations (M. McCloskey, personal communication, October 1993). This would explain, at a purely semantic-level, the simultaneous occurrence of a significant same–different effect and of an insignificant close–far effect on error rates in the physical matching task.

Thus, the existence of asemantic transcoding routes remains a vexing question requiring further research. In particular, the exact role of phonological representations in this process remains to be specified, perhaps by systematically varying the phonological similarity of the targets. We note that in a recent study of a brain-lesioned, patient with deep dyslexia, reading by directly mapping Arabic digits onto a phonological representation appeared disturbed, but reading by means of a semantic route was partially preserved (Cohen, Dehaene, & Verstichel, 1994). Such neuropsychological studies may provide more clear-cut evidence for multiple-route models.

In conclusion, how do people decide if two numerals are the same? It seems that instructions can only partially modulate the level of information processing used for same–different judgments. In the numerical matching task, participants gave priority to the semantic similarity of the targets. Yet it seems that they also noticed the occasional physical identity of the targets and used this information to respond faster in pure trials. Conversely, in the physical matching task, participants gave priority to the physical and lexical similarity of the targets. Yet there was occasional interference from a magnitude representation, and there was also evidence that participants recognized the identity of numerals such as *2* and *two*, perhaps by using an intermediate phonological representation. Thus, the participants' decisions seemed to be based, in each case, on a weighting of similarity ratings obtained at different levels of the processing hierarchy.

References

Anderson, S. W., Damasio, A. R., & Damasio, H. (1990). Troubled letters but not numbers: Domain specific cognitive impairments following focal damage in frontal cortex. *Brain, 113,* 749–766.

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

Campbell, J. I. D., & Clark, J. M. (1992). Cognitive number processing: An encoding-complex perspective. In J. I. D. Campbell (Ed.), *The nature and origins of mathematical skills,* (pp. 457–491). Amsterdam: North Holland.

Clark, J. M., & Campbell, J. I. D. (1991). Integrated versus modular theories of number skills and acalculia. *Brain and Cognition, 17,* 204–239.

Cohen, L., Dehaene, S., & Verstichel (1994). Number words and number non-words: A case of deep dyslexia extending to arabic numerals. *Brain, 117,* 267–279.

Coltheart, M. (1980). Deep dyslexia: A right-hemisphere hypothesis. In M. Coltheart, K. Patterson, & J. C. Marshall (Eds.), *Deep dyslexia* (pp. 326–380). London: Routledge.

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

Dehaene, S. (in press). The organization of brain activations in number comparison: Event-related potentials and the additive-factors method. *Journal of Cognitive Neuroscience.*

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

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

Dehaene, S., & Cohen, L. (1991). Two mental calculation systems: A case study of severe acalculia with preserved approximation. *Neuropsychologia, 29,* 1045–1074.

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

Dehaene, S., & Mehler, J. (1992). Cross-linguistic regularities in the frequency of number words. *Cognition, 43,* 1–29.

Deloche, G., & Seron, X. (1984). Semantic errors reconsidered in the procedural light of stack concepts. *Brain & Language, 21,* 59–71.

Deloche, G., & Seron, X. (1987). Numerical transcoding: A general production model. In G. Deloche & X. Seron (Eds.), *Mathematical disabilities: A cognitive neuropsychological perspective* (pp 137–170). Hillsdale, NJ: Erlbaum.

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

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

Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition, 44,* 43–74.

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.

Hinrichs, J. V., Yurko, D. S., & Hu, J. M. (1981). Two-digit number comparison: Use of place information. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 890–901.

Krueger, L. E. (1989). Reconciling Fechner and Stevens: Toward a unified psychophysical law. *Behavioral and Brain Sciences, 12,* 251–267.

LeFevre, J., Bisanz, J., & Mrkonjic, L. (1988). Cognitive arithmetic: Evidence for obligatory activation of arithmetic facts. *Memory & Cognition, 16,* 45–53.

Lemaire, P., Barrett, S. E., Fayol, M., & Abdi, H. (1994). Automatic activation of addition and multiplication facts in elementary school children. *Journal of Experimental Child Psychology, 57,* 224–258.

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

McCloskey, M. (1992). Cognitive mechanisms in numerical processing: Evidence from acquired dyscalculia. *Cognition, 44,* 107–157.

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

McCloskey, M., Macaruso, P., & Whetstone, T. (1992). The functional architecture of numerical processing mechanisms: Defending the modular model. In J. I. D. Campbell (Ed.), *The nature and origins of mathematical skills* (pp. 493–537). Amsterdam: North Holland.

McCloskey, M., Sokol, S. M., & Goodman, R. A. (1986). Cognitive processes in verbal-number production: Inferences from the performance of brain-damaged subjects. *Journal of Experimental Psychology: General, 115,* 307–330.

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

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

Noel, M. P., & Seron, X. (1992). Notational constraints and number processing: A reappraisal of the Gonzalez and Kolers (1982) study. *Quarterly Journal of Experimental Psychology, 45(A),* 451–478.

Noel, M. P., & Seron, X. (1993). Arabic number reading deficit: A single case study. *Cognitive Neuropsychology, 10,* 317–339.

Posner, M. I. (1978). *Chronometric explorations of the mind.* Hillsdale, NJ: Erlbaum.

Sokol, S. M., McCloskey, M., Cohen, N. J., & Aliminosa, D. (1991). Cognitive representations and processes in arithmetic: Inferences from the performance of brain-damaged subjects. *Journal of Experiment Psychology: Learning, Memory, and Cognition, 17,* 355–376.

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

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

Zbrodoff, N. J., & Logan, G. D. (1986). On the autonomy of mental processes: A case study of arithmetic. *Journal of Experimental Psychology: General, 115,* 118–130.

Received October 22, 1993
Revision received April 29, 1994
Accepted May 3, 1994 ■