Exploring the Developmental Changes in Automatic Two-digit Number Processing

Winnie Wai Lan Chan, Terry K. Au, & Joey Tang

The University of Hong Kong, HKSAR

Author note

Correspondence concerning this article should be sent to Joey Tang, Department of Psychology, The University of Hong Kong, HKSAR. E-mail: joeytang@hku.hk. Phone: (+852) 2859 2295. Fax: (+852) 2858 3518; or Winnie Chan, Department of Psychology, The University of Hong Kong, HKSAR. E-mail: cwinnie@hku.hk. Phone: (+852) 2241 5650. Fax: (+852) 2858 3518.

Exploring the Developmental Changes in Automatic Two-digit Number Processing

Word count: 5,067

Abstract

Even when two-digit numbers are irrelevant to the task at hand, adults process them. Do children process numbers automatically, and if so, what kind of information is activated? In a novel dot-number Stroop task, children (Grades 1 to 5) and adults were shown two different two-digit numbers made up of dots. Participants were asked to select the number that contained the larger dots. If numbers are processed automatically, reaction time for dot-size judgment should be affected by numerical characteristics. The results suggest that, like adults, children process two-digit numbers automatically. Based on the present findings, we propose a developmental trend for automatic two-digit number processing that goes from decomposed sequential (activation of decade digit followed by that of unit digit) to decomposed parallel processing (simultaneous activation of decade and unit digits).

Keywords: Two-digit numbers, automaticity, development, numerical processing, number magnitude, Stroop

Exploring the Developmental Changes in Automatic Two-digit Number Processing

Two-digit number processing is essential to early acquisition of mathematics. Automatically activating the numerical magnitude without conscious monitoring can facilitate calculations. The classical number Stroop task has demonstrated automatic two-digit number processing in adults (Ganor-Stern, Tzelgov, & Ellenbogen, 2007). In each Stroop trial, participants were presented with two numbers differing in physical size (e.g., 24 39) and were asked to select the physically larger number. Adult participants generally responded faster when the physically larger number was also numerically larger (i.e., congruent trial, 24 39) than when it was numerically smaller (i.e., incongruent trial, 24 39). Such congruity effect represents facilitation and interference from the automatic activation of numerical magnitude in the congruent and incongruent trials respectively.

The classical number Stroop task works well for adults but probably not for children. For one thing, a typical experiment contains many trials (e.g., over 700 trials; Ganor-Stern et al., 2007), which are likely to cause fatigue and lose children's attention. For another, complying with the request to "pick the *physically* larger number" can be difficult for children who do not understand the word "*physically*". Moreover, in a language like Chinese, there is no simple, child-friendly translation for this particular instruction.

The classical number Stroop task has been administered to children in Belgium (Mussolin & Noël, 2007, 2008). No congruity effect was found among third graders. The

authors postulated that perhaps numerical processing had taken place automatically but was too slow to interact with the much faster physical-size comparison. To ensure sufficient time for numerical processing, the task was modified so that children could see the numbers in identical size beforehand (e.g., 12 24 was shown first, followed by 12 24). Although the third graders showed the congruity effect in the modified task, the numerical processing had been purposely facilitated and hence might not have been truly automatic. Perhaps a better alternative is to slow down the physical-size processing instead (e.g., Schwarz & Ischbeck, 2003). This was our strategy in developing a novel number Stroop task – the dot-number Stroop task. In Study 1, we validated the novel task, and in Study 2, we used it to explore the development of automatic two-digit number processing among Chinese children.

For the dot-number Stroop task, we created a set of 48 trials, far fewer than the over 700 trials in the classical task (Ganor-Stern et al., 2007). In each trial, we presented two numbers with some dots inside and asked the participants to choose the number containing the larger dots. We thus avoided the expression "*physically larger*" — difficult for young children and without a simple translation in the native language of our Chinese participants — used in the classical number Stroop task. Also, in our dot-number Stroop task, the physical dimension (i.e., dot size) was less salient and should take longer to process than that in the classical task (i.e., the physical size of numbers), thereby giving automatic numerical processing a chance in the race against physical-size processing.

Like the classical task, the dot-number Stroop task included congruent pairs (with

the numerically larger number containing larger dots; e.g., 13 **26**; for ease of exposition, the

number containing the larger dots would be boldfaced hereafter) and incongruent pairs

(with the numerically larger number containing smaller dots; e.g., **13** 26). If the irrelevant

numerical magnitude was automatically activated, it should facilitate physical-size

processing in the congruent condition and interfere with it in the incongruent condition.

Hence, congruent pairs should elicit faster and more accurate responses than incongruent

pairs, yielding a congruity effect.

We also examined what kind of numerical information was automatically activated.

Consider the number "26". In holistic processing, the magnitude of the whole number "26"

is activated (Brysbaert, 1995; Dehaene, Dupous, & Mehler, 1990; Reynvoet & Brysbaert,

1999). In decomposed processing, both the decade digit "2" and unit digit "6" are activated

(Nuerk, Weger, & Willmes, 2001, 2004; Ratinckx, Brysbaert, & Fias, 2005). When the

decade digit is activated first followed by the unit digit (e.g., Nuerk, Kaufman, Zoppoth, &

Willmes, 2004), decomposed sequential processing occurs; whereas when the decade digit

and unit digit are activated simultaneously (e.g., Nuerk, Kaufman et al., 2004), decomposed

parallel processing occurs. Note that holistic processing and decomposed processing can

occur together (i.e., hybrid processing; see studies on intentional two-digit number

processing, e.g., Liu, Wang, Corbly, Zhang, & Joseph, 2006; Nuerk, Weger et al., 2001,

2004; Wood, Mahr, & Nuerk, 2005). To date, the only study that investigated different types

of numerical activation in automatic two-digit number processing documented decomposed

parallel processing in adults (Ganor-Stern et al., 2007). Using the novel task, we made the

first attempt to explore this with children.

In both Studies 1 and 2, we attempted to differentiate between decomposed

sequential processing and decomposed parallel processing[1]. Each stimulus number pair was

either unit-decade compatible, where both the decade digit and unit digit of a number were

larger than those of the other (e.g., 36 and 58, with $3 < 5$ and $6 < 8$), or unit-decade

incompatible, where the number with the smaller decade digit had the larger unit digit (e.g.,

36 and 54, with $3 < 5$ but $6 > 4$). In order to match the overall numerical distance between

the pairs of numbers in the stimulus set, unit-decade incompatible pairs consistently had a

larger decade distance (i.e., 1 vs. 3 in "17 34") than unit-decade compatible pairs (e.g., 1 vs.

2 in "13 26"; likewise in Nuerk, Kaufman et al., 2004; Nuerk, Weger et al., 2001).

Decomposed processing would be indicated by an interaction between unit-decade

compatibility and congruity, and its direction would determine the subtype of activation

(sequential vs. parallel). Figure 1 shows the predicted pattern of reaction times for

sequential processing and parallel processing. In decomposed sequential processing, a larger

congruity effect would be predicted for unit-decade incompatible pairs compared to

unit-decade compatible pairs; whereas in decomposed parallel processing, a larger congruity

effect would be predicted for the unit-decade compatible pairs. The rationale for these

predictions is as follows.

Consider decomposed sequential processing where the decade digit is activated

first. Since the decade distance in unit-decade incompatible pairs is larger than that in

unit-decade compatible pairs, as noted earlier, the decade digits in the former would be

activated faster due to the distance effect of the decade digits (Moyer & Laudauer, 1967)

and hence generate a greater impact (facilitation or interference) on the physical-size

comparison, leading to a larger congruity effect for the unit-decade incompatible pairs

compared to the unit-decade compatible pairs.

The opposite would be predicted for decomposed parallel processing, where both

decade digit and unit digit are activated simultaneously. In the congruent condition,

unit-decade compatible pairs (e.g., 13 **26**) should be responded to faster than unit-decade

incompatible pairs (e.g., 17 **34**), because numerical comparisons between the decade digits

(i.e., 1 vs. 2) and between the unit digits (i.e., 3 vs. 6) should speed up the target response

(i.e., 26) in the former, whereas numerical comparison between the unit digits (i.e., 7 vs. 4)

should slow down the target response (i.e., 34) in the latter. By contrast, in the incongruent

condition (i.e., compatible pair: **13** 26; incompatible pair: **17** 34), unit-decade compatible

pairs should be responded to more slowly than unit-decade incompatible pairs, because both

numerical comparison between the decade digits and between the unit digits should slow

down the target response for the compatible pairs, whereas numerical comparison between

the unit digits at least should speed up the target response for the incompatible pairs.

Overall, decomposed parallel processing should yield a larger congruity effect for

unit-decade compatible pairs compared to unit-decade incompatible pairs.

Study 1

The goal of Study 1 was to validate the novel dot-number Stroop task by

replicating previous findings on adults' automatic numerical activation (Ganor-Stern et al.,

2007). One group of adults did an abbreviated classical number Stroop task, and another

group did the dot-number Stroop task. Both contained the same set of 48 number pairs; they

differed only in the target physical dimension (font size vs. dot size). Based on Ganor-Stern

et al. (2007), we predicted that both tasks would yield decomposed parallel processing in

adults.

Method

Participants.   One hundred and seventy Chinese adults participated in this study:

77 (mean age = 21.5 years, SD = 2.8; 54 female, 23 male) performed an abbreviated

classical number Stroop task, and 93 (mean age = 21.5 years, SD = 3.2; 66 female, 27 male)

performed the dot-number Stroop task. In both Study 1 and 2, all participants had normal or

corrected-to-normal vision and gave written informed consent.

Stimuli.

Numerical dimension.   The number pairs are shown in Appendix. In both tasks, the same 24 pairs of two-digit numbers were presented twice to each participant: once in the congruent condition (i.e., physical size congruent with numerical magnitude) and once in the incongruent condition. Decade numbers (e.g., 20, 30) and tie numbers (e.g., 22, 33) were excluded. The decade digits and unit digits always differed within a pair.

Within each congruity condition, the 24 number pairs were evenly divided into unit-decade compatible (e.g., for 12 24, both $1 < 2$ and $2 < 4$) and incompatible (e.g., for 17 34, $1 < 3$ but $7 > 4$). The compatible and incompatible pairs did not differ significantly in average unit distance, overall distance, and overall magnitude (all ps $> .05$). However, due to matching the average unit distance and overall distance, unit-decade incompatible pairs necessarily had larger decade distance than compatible pairs, $t(23) = -4.80$, $p < .01$. The numbers in each of the compatibility conditions were evenly divided into large ratio (1:2, e.g., 12 24) and small ratio (2:3, e.g., 24 36).

In addition to the 48 test pairs, four practice pairs and four dummy pairs, different from the test pairs, were included.

Physical dimension.   Figure 2 shows some sample stimuli used in the two tasks. In the abbreviated classical number Stroop task, the 48 number pairs were presented in Arial font, evenly divided among the four contrasts in physical size: font sizes of points 49 and points 65, font sizes of points 52 and points 72, font sizes of points 61 and points 82,

and finally font sizes of points 70 and points 90. The numbers in each pair were shown in

black against a grey background, side-by-side along the horizontal axis, and separated by 50

mm.

In the dot-number Stroop task, the number pairs were shown in white against

a black background. The paired numbers were presented side by side along the horizontal

axis, separated by 45 mm. The number pairs were shown in Arial Black font, with font size

of points 80. The black dots were identical in size within a number and differed across

paired members. The dots were shown in a grid pattern, with the centres of the dots spaced

6 mm apart. The 48 test pairs were evenly divided among four dot-size contrasts in

diameters: (a) 1.00 mm vs. 1.80 mm, (b) 1.50 mm vs. 2.70 mm, (c) 1.70 mm vs. 3.06 mm,

and (d) 2.00 mm vs. 3.60 mm.

In both tasks, the size contrasts were counterbalanced across the congruity

and unit-decade compatibility conditions. The number with a larger physical size (in the

abbreviated classical task) or larger dots (in dot-number Stroop task) within a pair appeared

equally often on the left and the right sides.

Procedure.   The tasks were presented as online computer tasks. To ensure smooth

execution, the computer programmes were actually downloaded to and ran off the local

computer. For each task, the testing session began with a block of four practice trials which

were repeated if necessary. When a participant achieved at least three correct, the test block

would then start with four dummy trials as a warm up, followed by the 48 test pairs. Both

the practice and dummy trials were excluded from analysis. In both tasks, the 48 test trials

were presented in a fixed, pseudorandom order. For each pair, participants were told to

select the physically larger number in the abbreviated classical number Stroop task or the

number containing the larger dots in the dot-number Stroop task as quickly and as

accurately as possible. Each trial was preceded by a fixation cross for 500 ms. When the

number pair appeared, participants were required to indicate which number was physically

larger or contained the larger dots by pressing the F key with their left index finger if the

target number was on the left and pressing the J key with their right index finger if it was on

the right. Both the response and the reaction time were recorded by the computer. Each

number pair stayed on the screen for a maximum of 5,000 ms or until the participant

pressed a key. Between trials, there was a 2,000 ms pause. Each task took less than 5 min.

Results

The data were first trimmed using Nuerk et al.'s procedure (2001). Reaction times

shorter than 200 ms or longer than 2,000 ms were excluded. Reaction times more than three

standard deviations above or below a participant's average were likewise excluded. As a

result, 1.8% and 1.5% of trials were excluded from the abbreviated classical task and the

dot-number Stroop task respectively. Mean error rate for both tasks was 1%. Since both

speed and accuracy were important performance indicators in a number Stroop task, we

combined them in the measure of inverse efficiency by dividing the mean reaction time by

the proportion of correct responses. Only reaction times of correct trials were considered.

Given the same reaction time, the higher the proportion of correct responses, the smaller the

inverse efficiency score would be, indicating higher efficiency. In effect, this score is an

adjusted reaction time. Similar inverse efficiency measure has also been used in

computerized tasks on numerical processing (Butterworth, 2003; Iuculano, Tang, Hall, &

Butterworth, 2008).

A $2 \times 2 \times 2$ mixed-design ANOVA was conducted on inverse efficiency, with

congruity (congruent or incongruent) and unit-decade compatibility (compatible or

incompatible) as within-subject factors, and task (abbreviated classical task or dot-number

Stroop task) as a between-subject factor. Mean inverse efficiency scores in each condition

are shown in Figure 3.

A significant congruity effect was found, $F(1, 168) = 76.73$, $p < .01$, $\eta^2 = .13$,

showing that congruent pairs (476 ms) were responded to more efficiently than incongruent

pairs (495 ms). Moreover, the congruity $\times$ unit-decade compatibility interaction was

significant, $F(1, 168) = 32.71$, $p < .01$, $\eta^2 = .06$, where the congruity effect for unit-decade

compatible pairs was larger than that for incompatible pairs, $t(169) = 5.85$, $p < .01$. Together,

these results indicated that adults activated the numbers automatically in a decomposed

parallel manner. There was a significant task effect, $F(1, 168) = 47.55$, $p < .01$, $\eta^2 = .22$,

where the dot-number Stroop task (526 ms) was generally responded to less efficiently than the abbreviated classical number Stroop task (446 ms). This task effect reflects our success in making the physical dimension (dot size) of the dot-number Stroop task less salient to slow down the physical size comparison. This may give children's numerical processing – generally slower than adults—a chance to interfere with physical magnitude processing, which was critical for exploring children's automatic number processing in Study 2. More importantly, task did not interact significant with congruity and unit-decade compatibility (ps >.05), indicating that automatic numerical activation was similar in the two versions of the task.

Discussion

In this study, both the abbreviated classical number Stroop task and the dot-number Stroop task replicated the previous findings (Ganor-Stern et al., 2007), demonstrating decomposed parallel processing in adults' automatic activation of two-digit numbers. The dot-number Stroop task has the additional advantage of being more child-friendly. It should work better for children who may not understand the expression "physically larger" and in languages without a simple way to say "physically larger", as is the case for Chinese. In Study 2, we used it to examine children's automatic numerical processing.

Study 2

Method

Participants.   Three hundred and ninety Chinese children from a girl's primary

school in Hong Kong participated in the study. Among them, 90 were from Grade 1 (mean

age = 7.1 years, SD = 0.3), 81 were from Grade 2 (mean age = 8.1 years, SD = 0.3), 78 were

from Grade 3 (mean age = 9.1 years, SD = 0.3), 63 were from Grade 4 (mean age = 10.1

years, SD = 0.4), and 78 were from Grade 5 (mean age = 11.2 years, SD = 0.4). Parental

consent was obtained. The study took place at the end of the second semester, when Grade 1

students had already learned the numbers up to 100, according to the curriculum in Hong

Kong.

Stimuli and procedure.   The participants performed the dot-number Stroop task

described in Study 1.

Results

The data trimming procedure used in Nuerk, Weger et al.'s study (2004) of children

in Grades 2 to 5 was adopted here. Reaction times shorter than 200 ms or longer than 2,000

ms for Grade 5, 3,000 ms for Grade 4, 4,000 ms for Grade 3, 5,000 ms for Grade 2, and

6,000 ms for Grade 1 were excluded from analysis. Reaction times more than three standard

deviations above or below each child's average were also excluded. Altogether, 1.6% of

trials were excluded from further analysis. Mean error rates for Grade 1 to Grade 4 were 2%,

and for Grade 5 was 1%. No child had more than 20% errors or excluded responses. Inverse

efficiency was calculated for each child. Mean inverse efficiency scores in each condition

for the five grade levels are shown in Table 1.

A $2 \times 2 \times 5$ mixed-design ANOVA was conducted on inverse efficiency, with congruity and unit-decade compatibility as the within-subject factors, and grade as a between-subject factor. A significant main effect for grade was found, $F(4, 385) = 38.28$, $p < .01$, $\eta^2 = .28$, showing that inverse efficiency decreased with grade (1166 ms for Grade 1; 1093 ms for Grade 2; 943 ms for Grade 3; 832 ms for Grade 4; and 706 ms for Grade 5). Post-hoc t-tests revealed significant difference between every pair of grades (all ps $< .05$), except for Grade 1 and Grade 2 ($p > .05$). This suggests that children in the upper grade levels responded more efficiently.

There was a significant congruity effect, $F(1, 385) = 74.88$, $p < .01$, $\eta^2 = .03$, with congruent pairs (919 ms) responded to more efficiently than incongruent pairs (977 ms). The congruity effect was modulated significantly by grade, $F(4, 385) = 3.01$, $p = .018$, $\eta^2 = .005$. Further analysis revealed that the congruity effect for Grade 4 was significantly smaller than that for Grade 1, $t(134) = 2.41$, $p = .02$, and the congruity effect for Grade 5 was significantly smaller than that for Grade 3, $t(123) = 2.06$, $p = .04$, and Grade 1, $t(104) = 2.74$, $p = .01$. A smaller congruity effect for older children might be due to higher efficiency in physical-size processing, leaving little room for further facilitation by the congruent numerical information. Alternatively, it might be due to the older children's better attentional control, thereby minimizing interference from the incongruent numerical

information.

The congruity × unit-decade compatibility interaction was also significant, $F(1,$ 385) = 5.27, p = .02, η^2 = .002, which is shown in Figure 4. The congruity effect for unit-decade compatible pairs was larger than that for incompatible pairs, $t(389)$ = 2.20, p = .03, indicating decomposed parallel processing. Because grade only interacted significantly with congruity, the type of numerical information processing appeared to be the same across the five grades. However, this view may be too simplistic and neglects possible individual differences. The ANOVAs reported thus far have not taken into account individual response patterns. Here, we used a two-step cluster analysis to identify subgroups of individuals across grades based on their inverse efficiency across four conditions derived from congruity (congruent vs. incongruent) and unit-decade compatibility (compatible vs. incompatible). Log-likelihood distance and Bayesian Information Criterion were used to define the distance between two clusters and choose the number of clusters, respectively. Two clusters (N = 70 for Cluster 1; N = 320 for Cluster 2) were identified. The composition of each cluster is shown in Figure 5. In general, the majority of children in Cluster 1 were in the lower grade levels, whereas those in Cluster 2 were in the upper grade levels. Hence, there appeared to be a trend of moving towards Cluster 2 as children developed.

A 2 × 2 × 2 mixed-design ANOVA was performed on the inverse efficiency with the two within-subject factors as before (i.e., congruity and unit-decade compatibility) and

cluster (Clusters 1 and 2) as a between-subject factor. The congruity × unit-decade

compatibility interaction was modulated by cluster, $F(1, 388) = 3.71$, $p = .055$, $\eta^2 = .003$.

Figure 6 depicts the interaction for each of the two clusters. Like adults, children in Cluster

2 showed a larger congruity effect for unit-decade compatible pairs than for incompatible

pairs, $t(319) = 4.55$, $p < .01$. By contrast, children in Cluster 1 showed similar congruity

effects for both compatible and incompatible pairs, $t(69) = -0.41$, $p = .68$. There is no clear

indication these younger children were engaging in parallel processing. At the same time,

their pattern of performance does not suggest sequential processing. If automatic two-digit

number processing follows the developmental trend observed with intentional processing

(Nuerk, Kaufman et al., 2004), the speculation from our results is that children in Cluster 1

might be in transitional stage from sequential to parallel processing.

Discussion

The present study demonstrated that Chinese children as young as first graders

could automatically process two-digit numbers, indicating that Chinese children developed

such automaticity earlier than their counterparts (third graders) in Belgium (Mussolin &

Noël, 2007, 2008). Nearly all children (93%) in the upper grade levels (Grades 3 to 5)

adopted an adult-like pattern of processing, i.e., decomposed parallel processing. The

pattern of numerical processing was, however, mixed for the lower graders (Grades 1 and 2).

While the majority (68%) had developed decomposed parallel processing, some of them

(32%) had not. We speculate that these children might be in the transition from sequential to parallel processing, with sequential processing on the way out and parallel processing yet to emerge. Such a developmental shift from sequential to parallel processing has also been observed in children's intentional processing of two-digit numbers (Nuerk, Kaufman et al., 2004).

General Discussion and Conclusions

Much research has focused on children's intentional processing of two-digit numbers, but until now, little was known about automatic processing of such numbers during childhood. In the present studies, we first developed a more child-friendly number Stroop task for examining automatic processing of two-digit numbers. We then administered this novel task to primary school children to see whether they could process two-digit numbers automatically, and if so, what magnitude information was activated.

In Study 1, we replicated previous findings of decomposed parallel processing among adults (Ganor-Stern et al., 2007) using the dot-number Stroop task. The dot-number Stroop task is more child-friendly because it substantially reduces the number of trials from over 700 (Ganor-Stern et al., 2007) to 48, and it does not require the linguistic differentiation between "*physically larger*" and "*numerically larger*" in the verbal instruction, rendering it more feasible to use with young children and in languages that do not have an easy way to express "physically larger". This novel task offers a better tool for

future investigation into automatic two-digit number processing in different populations.

Study 2 revealed that Chinese children as young as age 7 years could process two-digit numbers automatically. At that age, some Western children have been documented to begin processing single-digit numbers automatically (Mussolin & Noël, 2007; Rubinsten, Henik, Berger, & Shahar-Shalev, 2002), while others do so at age 8 (Girelli, Lucangeli, & Butterworth, 2000). Chinese children, by contrast, begin to process single-digit numbers automatically at the age of 5 (Zhou, Chen, Chen, & Jiang et al, 2007). It appears that Chinese children develop automaticity in numerical processing ahead of their Western counterparts, perhaps due to early formal and informal number-related experiences such as counting and simple calculations in Chinese families and preschools (Huntsinger, Jose, Liaw, & Ching, 1997; Miller, Kelly, & Zhou, 2005). Because two-digit numbers are relatively more complex and learned at a later stage compared to single-digit numbers, Western children may develop automaticity in two-digit number processing after age 7 to 8 years (Mussolin & Noël, 2007, 2008).

Study 2 is the first to explore the developmental changes of automatic two-digit number processing. Using cluster analysis, two groups of children were identified showing different patterns of numerical processing. Because children seemed to progress from one cluster to another with age, these clusters could prove to characterise the developmental stages of automatic two-digit number processing.

Like adults, children in Cluster 2 (covering nearly all of the third, fourth and fifth graders) processed two-digit numbers in a decomposed parallel pattern. Although children in Cluster 1 who were mainly first and second graders had not yet developed this adult-like pattern, they appeared to be moving away from the probably less mature way of processing (decomposed sequential) towards the adult-like one (decomposed parallel). Together, these results suggest a developmental story that goes from sequential processing in preschoolers, transits towards parallel processing in first and second graders, and finally arrives at parallel processing in third and fourth graders. Further studies are needed to verify this trend by extending the investigation to preschoolers.

In summary, the dot-number Stroop task has proven to be a child-friendly tool for investigating automatic processing of two-digit numbers. With this novel tool, we have demonstrated that Chinese children as young as first graders can process two-digit numbers automatically. Our studies constitute a first attempt to discover the developmental story of automatic two-digit number processing, which seems to go from decomposed sequential to parallel processing.

References

Brysbaert, M. (1995). Arabic number reading: on the nature of the numerical scale and the

origin of phonological reading. *Journal of Experimental Psychology: General, 124*,

434–452.

Buckley, P. B., & Gillman, C. B. (1974). Comparisons of digits and dot patterns. *Journal of

Experimental Psychology, 103*, 1131–1136.

Butterworth, B. (2003). Dyscalculia Screener. London: nfer Nelson Publishing Company

Ltd.

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.

Ganor-Stern, D., Tzelgov, J., & Ellenbogen, R. (2007). Automaticity and two-digit numbers.

Journal of Experimental Psychology: Human Perception and Performance, *33*,

483–496.

Girelli, L., Lucangeli, D., & Butterworth, B. (2000). The Development of Automaticity in

Accessing Number Magnitude. *Journal of Experimental Child Psychology, 76*,

104–122.

Huntsinger, C. S., Jose, P. E., Liaw, F.-R., & Ching, W.-D. (1997). Cultural differences in

early mathematics learning: A comparison of Euro-American, Chinese-American,

and Taiwan-Chinese families. *International Journal of Behavioral Development, 21*, 371–388.

Iuculano, T., Tang, J., Hall, C. W., & Butterworth, B. (2008). Core information processing deficits in developmental dyscalculia and low numeracy. *Developmental science*, *11*, 669–80.

Liu, X., Wang, H., Corbly, C. R., Zhang, J., & Joseph, J. E. (2006). The Involvement of the Inferior Parietal Cortex in the Numerical Stroop Effect and the Distance Effect in a Two-digit Number Comparison Task. *Journal of Cognitive Neuroscience*, *18*, 1518–1530.

Loftus, G. R., & Masson, M. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review, 1*, 476–490.

Miller, K. F., Kelly, M., & Zhou, X. (2005). Learning mathematics in China and the United States: Cross-cultural insights into the nature and course of preschool mathematical development. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 163–177). New York: Psychology Press.

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

Mussolin, C., & Noël, M.-P. (2007). The non-intentional processing of Arabic numerals in children. *Journal of Clinical and Experimental Neuropsychology*, *29*, 225–234.

Mussolin, C., & Noël, M.-P. (2008). Automaticity for numerical magnitude of two-digit

Arabic numbers in children. *Acta Psychologica*, *129*, 264–272.

Nuerk, H.-C., Kaufmann, L., Zoppoth, S., & Willmes, K. (2004). On the development of the

mental number line: More, less, or never holistic with increasing age?

Developmental Psychology, *40*, 1199–1211.

Nuerk, H.-C., Weger, U., & Willmes, K. (2001). Decade breaks in the mental number line?

Putting the tens and units back in different bins. *Cognition*, *82*, B25–B33.

Nuerk, H.-C., Weger, U., & Willmes, K. (2004). On the perceptual generality of the

unit-decade compatibility effect. *Experimental Psychology*, *51*, 72–79.

Ratinckx, E., Brysbaert, M., & Fias, W. (2005). Naming two-digit Arabic numerals:

Evidence from masked priming studies. *Journal of Experimental Psychology:*

Human Perception and Performance, *31*, 1150–1163.

Reynvoet, B., & Brysbaert, M. (1999). Single-digit and two-digit numerals address the same

semantic number line. *Cognition, 72*, 191–201.

Rubinsten, O., Henik, A., Berger, A., & Shahar-Shalev, S. (2002). The Development of

Internal Representations of Magnitude and Their Association with Arabic Numerals.

Journal of Experimental Child Psychology, 81, 74–92.

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 and Performance, 29, 507–522.

Wood, G., Mahr, M., & Nuerk, H.-C. (2005). Deconstructing and reconstructing the

10-Base structure of Arabic numbers. *Psychology Science, 47*, 84–95.

Zhou, X., Chen, Y., Chen, C., Jiang, T., Zhang, H., & Dong, Q. (2007). Chinese

kindergartners' automatic processing of numerical magnitude in Stroop-like tasks.

Memory & Cognition, 35, 464–470.

Footnote

[1] Holistic processing (i.e., activation of the whole number magnitude) could be indicated by an interaction between congruity and ratio. In both the abbreviated classical number Stroop task and the dot-number Stroop task, the large-ratio pairs (e.g., 12 24) were necessarily numerically smaller than the small-ratio pairs (e.g., 24 36) due to the matching of the overall numerical distance within pairs. Predicted from the magnitude effect (Buckley & Gillman, 1974) and the psychological distance effect (Mussolin & Noël, 2008), the numbers in the large-ratio pairs would be activated faster than those in the small-ratio pairs. Hence, if the whole-number magnitude was activated automatically during the physical processing, the large-ratio pairs should yield a larger congruity effect than the small-ratio pairs. Such interaction was consistently observed in Studies 1 and 2, supporting holistic processing. Yet, in an additional study (not presented in this paper), there was no such interaction when the numbers in the pairs were not multiples of one another, implying that the activation of whole-number magnitude in Studies 1 and 2 was probably facilitated by the manipulation of the ratio relationship between the paired numbers. Additional experiments are needed to clarify the "holistic processing" observed in the present studies.

- We developed a novel task for examining automatic two-digit number processing.
- Children develop from sequential to parallel processing of two-digit numbers.
- Chinese children develop automatic number processing earlier than Western children.

Figure 1

Figure 1. Predicted Reaction Times (RT) for Decomposed Sequential Processing and
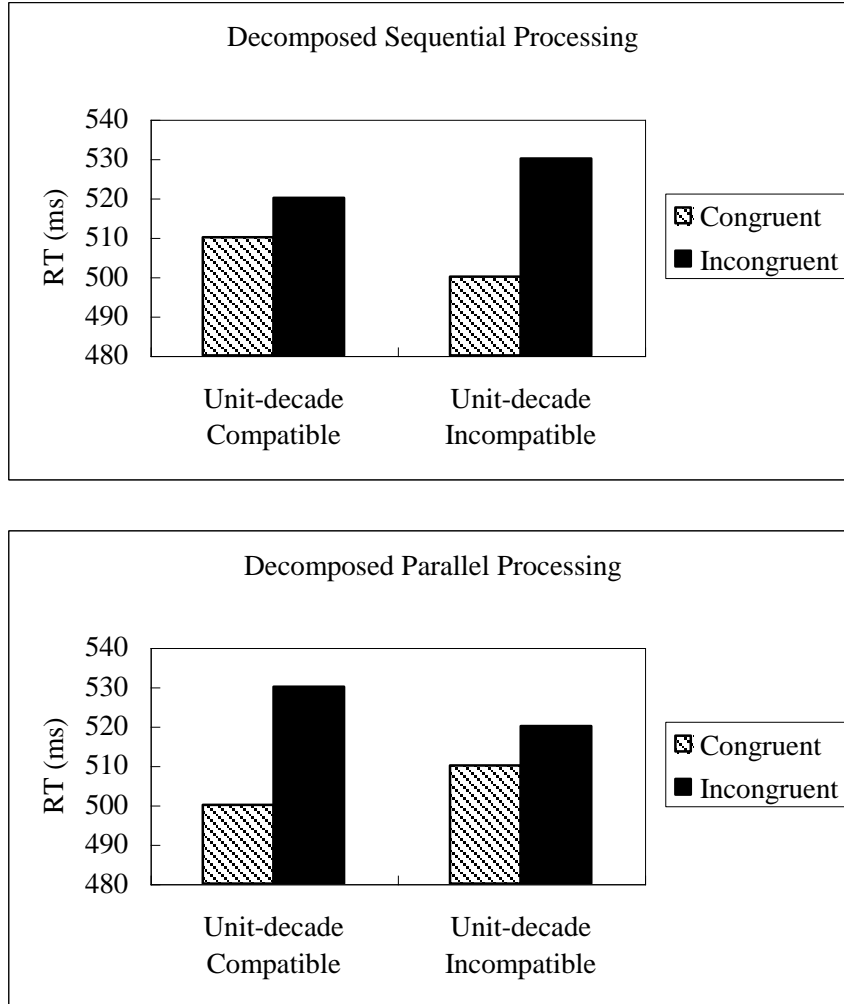
Decomposed Parallel Processing.

Figure 2

Figure 2. Examples of Stimuli.

Abbreviated Classical Number Stroop Task

	Compatible	Incompatible
Congruent	13 26	27 54
Incongruent	26 13	54 27

Dot-number Stroop Task

	Compatible	Incompatible
Congruent	13 26	27 54
Incongruent	26 13	54 27
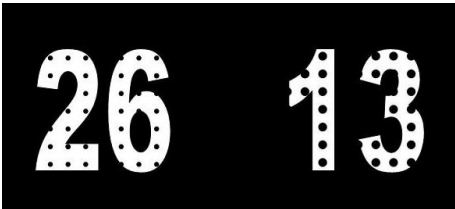
Figure 3

Figure 3. Inverse Efficiency (IE) as a Function of Congruity and Unit-decade

Compatibility among Adults. Bars present 95% confidence intervals computed

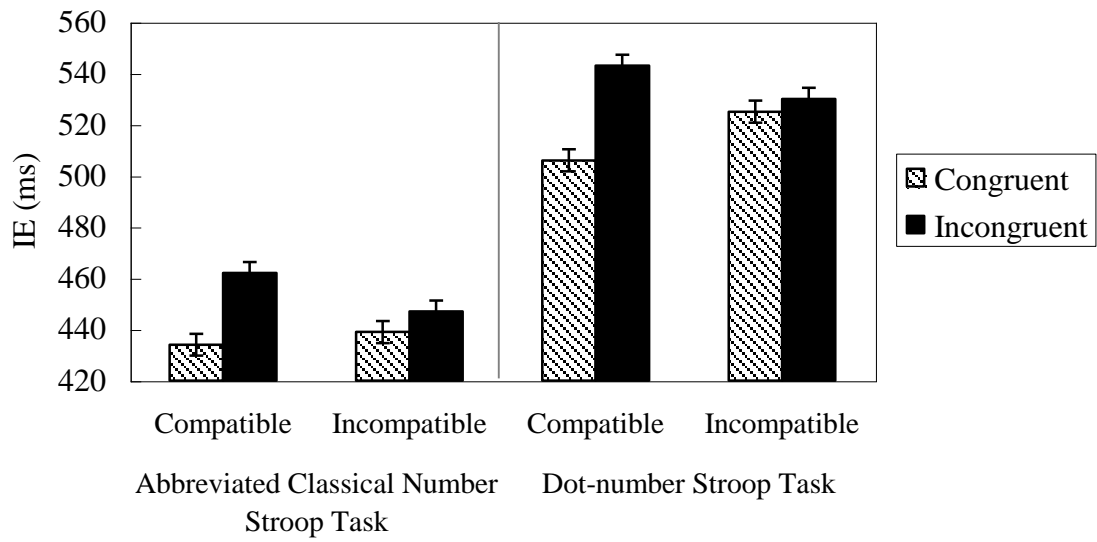following Loftus and Masson (1994).

Figure 4

Figure 4. Inverse Efficiency (IE) as a Function of Congruity and Unit-decade

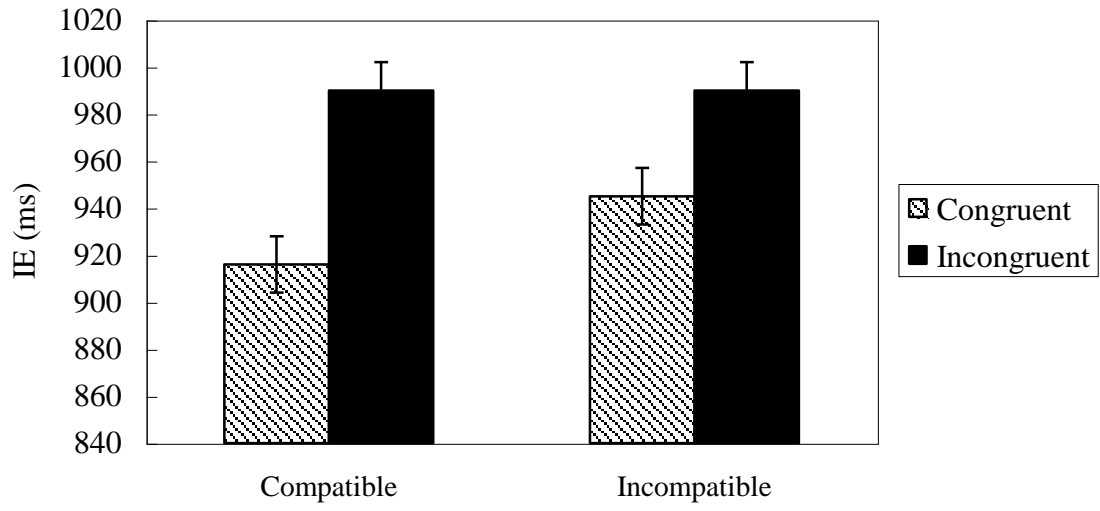Compatibility among Children. Bars present 95% confidence intervals.

Figure 5

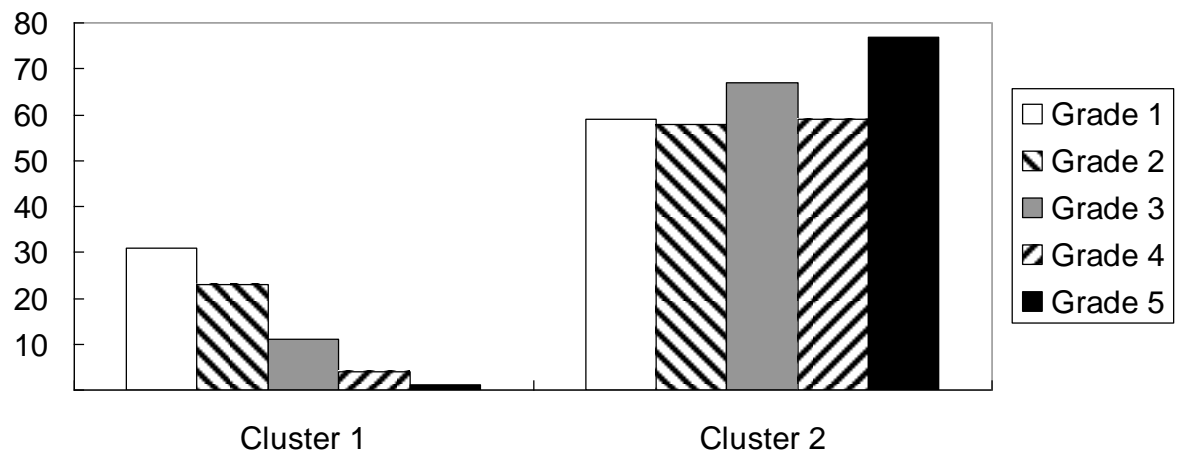*Figure 5.* Numbers of Children in the Two Clusters.

Figure 6

Figure 6. Inverse Efficiency (IE) as a Function of Congruity and Unit-decade

Compatibility among Two Clusters of Children. Bars present 95% confidence
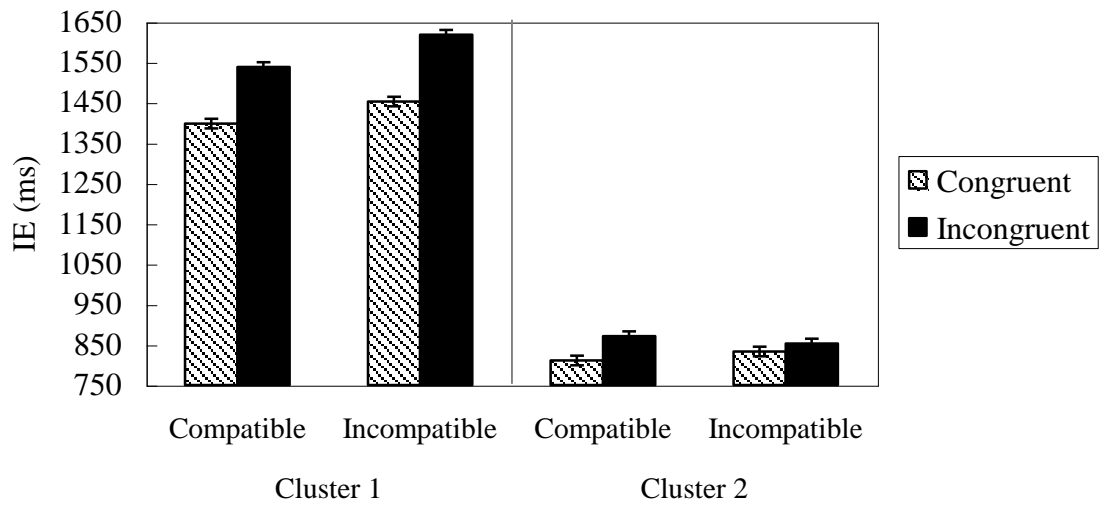
intervals.

Table 1

Table 1

Inverse Efficiency (in ms) as a Function of Congruity and Unit-decade Compatibility among Children from Grade 1 to Grade 5

	Unit-decade Compatible	Unit-decade Incompatible
Grade 1		
Congruent	1099 (334)	1136 (325)
Incongruent	1203 (366)	1225 (480)
Grade 2		
Congruent	1054 (293)	1072 (291)
Incongruent	1122 (320)	1126 (358)
Grade 3		
Congruent	903 (252)	923 (257)
Incongruent	979 (295)	967 (321)
Grade 4		
Congruent	794 (251)	833 (257)
Incongruent	859 (245)	841 (293)
Grade 5		
Congruent	675 (133)	702 (136)
Incongruent	726 (153)	720 (173)

Note. Standard deviations are in parentheses

Appendix

Number Pairs Repeated in Congruent and Incongruent Conditions

	Unit-Decade Compatibility	Number Pair		Overall Distance	Decade Distance	Unit Distance
Large Ratio	Compatible	12	24	12	1	2
		13	26	13	1	3
		21	42	21	2	1
		23	46	23	2	3
		31	62	31	3	1
		32	64	32	3	2
	Incompatible	17	34	17	2	3
		18	36	18	2	2
		19	38	19	2	1
		27	54	27	3	3
		28	56	28	3	2
		29	58	29	3	1
Small Ratio	Compatible	24	36	12	1	2
		26	39	13	1	3
		42	63	21	2	1
		46	69	23	2	3
		62	93	31	3	1
		64	96	32	3	2
	Incompatible	34	51	17	2	3
		36	54	18	2	2
		38	57	19	2	1
		54	81	27	3	3
		56	84	28	3	2
		58	87	29	3	1