

Empirical Research

The Association Between Number Magnitude and Space Is Dependent on Notation: Evidence From an Adaptive Perceptual Orientation Task

Tianwei Gong^a, Baichen Li^a, Limei Teng^a, Zijun Zhou^b, Xuefei Gao^c, Ting Jiang^{*a}

[a] Faculty of Psychology, Beijing Normal University, Beijing, China. [b] Department of Energy and Power Engineering, Tsinghua University, Beijing, China. [c] Queensland University of Technology, Brisbane, Australia.

Abstract

Research on adults' numerical abilities suggests that number representations are spatially oriented. This association of numbers with spatial response is referred to as the SNARC (i.e., spatial–numerical association of response codes) effect. The notation-independence hypothesis of numeric processing predicts that the SNARC effect will not vary with notation (e.g., Arabic vs. number word). To test such assumption, the current study introduced an adaptive experimental procedure based on a simple perceptual orientation task that can automatically smooth out the mean reaction time difference between Arabic digits and traditional Chinese number. We found that the SNARC effect interacted with notation, showing a SNARC effect for Arabic digits, but not for verbal number words. The results of this study challenged the commonly held view that notation does not affect numerical processes associated with spatial representations. We introduced a parallel model to explain the notation-dependent SNARC effect in the perceptual orientation judgment task.

Keywords: automatic processing, SNARC effect, perceptual orientation task, mathematical cognition, notation effect

Journal of Numerical Cognition, 2019, Vol. 5(1), ##–##, <https://doi.org/10.5964/jnc.v5i1.169>

Received: 2018-01-22. Accepted: 2018-06-22. Published (VoR): 2019-03-22.

*Corresponding author at: Faculty of Psychology, Beijing Normal University, No. 19 Xijiekouwai Street, Haidian District, 100875, Beijing, P. R. China.
E-mail: psytingjiang@gmail.com



This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International License, CC BY 4.0 (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Research into cognitive number processing has made considerable progress over the last two decades (Butterworth, 1999; Campbell, 2005; Dehaene, 2011). In numerical cognition studies, the interaction between numbers and space (Sekuler, Rubin, & Armstrong, 1971) has been a major area of inquiry. Dehaene and his colleagues (Dehaene, Bossini, & Giraux, 1993; Viarouge, Hubbard, & Dehaene, 2014) found the Spatial-Numerical Association of Response Codes effect (the SNARC effect) by using parity (odd-even) judgment tasks, in which they asked participants to indicate the parity of the digits 0–9 presented to them. They found that the larger the target number was, the faster the response on the right-hand side relative to the response on the left-hand side. From then on, interpretation of the SNARC effect has disputed for some time. The debate over whether the SNARC effect is dependent on notation or not is one of those issues.

In general, it has been suggested that the SNARC effect is insensitive to notation type (Iversen, Nuerk, & Willmes, 2004; Nuerk, Iversen, & Willmes, 2004; Nuerk, Wood, & Willmes, 2005). Dehaene et al. (1993) detected the SNARC effects for Arabic digits and, albeit somewhat weaker, for number words. Iversen et al. (2004) observed the SNARC effect for Arabic digits, number words, and German number signs in German signers.

Moreover, [Nuerk et al. \(2005\)](#) used four different formats (Arabic digits, number words, auditory numbers, and dice patterns) and they found that there was no difference among different input notations and modalities of SNARC effects. All these studies were conducted by using the parity judgment task. Based on such data patterns, [Nuerk et al. \(2005\)](#) proposed the notation/modality-independence hypothesis. Assuming that notation type constitutes a manipulation that might modulate the visual complexity (e.g., simplicity vs. complexity) and frequency of use (e.g., high vs. low), one would expect that the SNARC effect did not appear to be sensitive to these characteristics ([Dehaene et al., 1993](#), p. 380).

A different prediction might be derived from the work of Fias and his colleagues ([Fias, 2001](#); [Fias, Brysbaert, Geypens, & d'Ydewalle, 1996](#)). They replicated the parity judgment task in Arabic digits and Dutch number words, and as expected, the SNARC effects were observed during the processes of different notations. However, they also developed a new phoneme-monitoring task that could be completed without direct access to the semantic information. They conducted two experiments ([Fias et al., 1996](#); [Fias, 2001](#)) with one using Arabic numbers and the other using Dutch number words. The participants were asked to judge whether the number name contained an /e/ sound. In this perceptual task, they provided evidence against the notation-independence hypothesis: The SNARC effect emerged in the Arabic digits' situation ([Fias et al., 1996](#)), but not in verbal numbers ([Fias, 2001](#)). Thus, whether the SNARC effect is stable and replicable for different notations may depend greatly on the particular task requirements (e.g., parity judgment vs. phoneme-monitoring task).

Researchers have not yet reached an agreement on whether the mechanism underlying the processes in parity judgments ([Dehaene et al., 1993](#)) is different from that in perceptual judgments (e.g., the phoneme-monitoring task, [Fias et al., 1996](#)). To be precise, most attention is focused on whether the numerical-spatial information is merely automatically activated without being essential for solving the parity judgment (i.e., magnitude-irrelevant) or the parity task requires the manipulation of numerical magnitude features that is used to make the parity judgment (i.e., magnitude-relevant). Owing to the fact that the processing of magnitude information was not part of the task requirements, the parity judgment task has generally been considered as the magnitude-irrelevant task (e.g., [Gevers, Verguts, Reynvoet, Caessens, & Fias, 2006](#), p. 32). However, this idea has been disputed by some. [Tzelgov and Ganor-Stern \(2005\)](#), p. 56 and p. 59 argued that the parity judgment task is a numerical task. Wood and his colleagues carried out a meta-analysis of 46 studies with a total of 106 experiments, and they also emphasized that the SNARC effect might be stronger in parity judgment tasks which require deeper semantic number processing than that in perceptual tasks, where the required analysis of numerical stimuli features, such as color or orientation discrimination, are superficial ([Wood, Willmes, Nuerk, & Fischer, 2008](#), p. 492). If the parity task was considered as a magnitude-irrelevant task, we would expect similar data patterns (e.g., the interaction between notation and SNARC) in the parity and perceptual tasks. However, the results in previous SNARC studies did not support the conventional viewpoint that these two tasks shared similar mechanisms (see the details, [Wood et al., 2008](#), p. 506).

Albeit indirectly, [Cohen Kadosh et al. \(2008\)](#) also tested the notation-independence hypothesis using the number Stroop paradigm. There are two comparison tasks (numerical vs. physical) in the number Stroop paradigm. When participants were instructed to refer to the numerical dimension and to decide which of two stimuli were larger while ignoring the physical dimension, a Numerical-SCE (i.e., the size congruity effect, or SCE, [Paivio, 1975](#)) appeared. When participants were instructed to judge which of two digits is physically larger and to ignore the numerical dimension, a Physical-SCE ([Cohen Kadosh et al., 2008](#), p. 1378) happened. They examined whether the Physical-SCE would appear with verbal numbers as well as with Arabic numbers and found

that only Arabic notation affected visual number processing. Therefore, they suggested that the task-irrelevant magnitude information should be processed to a lesser degree for verbal numbers than for Arabic digits when these two notations were presented in blocked (see their Figure 1, [Cohen Kadosh et al., 2008](#), p. 1381) or mixed order (see their Figure 4, p. 1385). However, as mentioned by [Cohen Kadosh et al. \(2008\)](#), the elimination of the effect with verbal numbers might be attributed to the fact that the physical comparison task was carried out too fast for the verbal numbers (see their Figure 1 and Figure 4), and the fast processing may not allow enough time for numerical information to interfere. Given the SNARC effect might have derived from a similar mechanism as the number Stroop effect ([Cohen Kadosh et al., 2008](#); [Tzelgov & Ganor-Stern, 2005](#)), we assume a common explanation that three potential factors, namely, task requirements (e.g., parity judgment vs. perceptual judgment), materials (e.g., Arabic digits vs. verbal numbers), and processing speed (e.g., fast vs. slow) will be responsible for the inconsistent results in previous SNARC studies.

The main purpose of this research was to test the notation-independence hypothesis with the current perceptual orientation task. In this task, numbers were surrounded by a square frame, and the whole stimulus slightly tilted to the left or right. Participants were asked to judge whether the frame tilted to the left side or the right side (see methods for details). There were three significant improvements compared with previous perceptual tasks (e.g., [Fias, 2001](#); [Fias et al., 1996](#); [Fias, Lauwereyns, & Lammertyn, 2001](#); [Lammertyn, Fias, & Lauwereyns, 2002](#)).

First and foremost, this study applied a similar but easier procedure relative to [Lammertyn et al.'s \(2002\)](#) study. In their study, Lammertyn et al. used upright vs. tilted (10° to the right) digits for the judgment of the target digits' orientation (upright – left key & tilted – right key in one block, and counterbalanced key assignment in another block). The difference between [Lammertyn et al.'s \(2002\)](#) task and the current task was that the current task did not contain upright digits. Thus, there was no need for participants to memorize the key assignments for target digits or to switch the key assignment across blocks which might make our task easier to complete than [Lammertyn et al.'s \(2002\)](#).

Second, the perceptual characteristics might lead to different processing difficulties of different numbers (without frames) in the perceptual judgment task and conceal the effect of numerical-spatial associations. For instance, the rotated digit “1” could be easier to define its orientation than “3” in the same tilt degree because the tilt status of straight lines might be easier to clarify than that of curved lines ([Lammertyn et al., 2002](#)). To reduce the perceptual confounds across different numbers, our paradigm made participants' direct attention away from the numbers by asking them to judge the orientation of the frames.

Third, since researchers have discovered that the response latencies would affect the emergence and the strength of the SNARC effect ([Gevers et al., 2006](#); also see [Wood et al., 2008](#)), it is necessary to equate processing time in both notation conditions when SNARC effects of different notations are compared. [Gevers et al. \(2006\)](#) found that the SNARC effect became stronger with increasing reaction time in a parity judgment task by ranking reaction times and dividing them into different bins. They suggested that the longer the time needed to reach a motor response, the stronger the impact of number magnitude and spatial-numerical associations on the response. Therefore, to smooth out the mean RT difference between Arabic digits and verbal numbers, we developed an adaptive algorithm to micro-modulate the response latencies by changing the degree of stimuli tilt, which was proved to be effective in the pre-experiment. This operation of reaction time is easier to imple-

ment for the present frame-orientation task than other perceptual tasks such as using triangles pointing upwards or downwards or lines superimposed on the number oriented horizontally or vertically (Fias et al., 2001).

Based on the summary above, with the improvements in experimental design and methods, this study aimed to find out whether the pattern of SNARC with traditional Chinese verbal numbers would match that pattern produced by Arabic digits with equal cross-notation response latencies. It is of huge significance as such result would challenge the commonly-held view that the SNARC was notation-independent, and further resolve the inconsistency in the literature of SNARC.

Pre-Experiment

The study has been approved by the ethical committee of Faculty of Psychology, Beijing Normal University. This pre-experiment had two goals: one was to demonstrate that reaction time decreases with the increase of the degree of tilt, namely, the larger the degree of tilt, the easier to judge. The second goal was to estimate a mathematical function used to reduce the response latency difference between the Arabic digits and Chinese verbal numbers in formal experiments.

Method

Participants

Eighteen students (9 female and 9 male, mean age = 20.56 years, $SD = 2.11$ years) participated in the pre-experiment study. All of them had normal or corrected-to-normal vision. Participants were right-handed, and they were not informed about the purpose of the study. Chinese was their mother tongue. They were paid 35 RMB (approximately 5.1 USD) for their participation in the pre-experiment.

Stimuli

Numbers in two different notations: Arabic digits (1, 2, 3, 4, 5, 6, 7, 8, and 9) and traditional Chinese number words (壹, 貳, 叁, 肆, 伍, 陆, 柒, 捌, and 玖) were put in a square-shape frame (visual angle: 2.87°) individually and filled the frame completely. Each number as well as its frame tilted in different degrees, ranging from 2° to 12° .

The motivation behind comparing Arabic digits with Chinese verbal numbers is threefold. In the first place, although many previous studies used number words in western or European languages, which are made up of several horizontally organized letters, Chinese number words are uniquely special because they resemble much more closely to Arabic digits in their basic iconographic formats (Jiang et al., 2010). Additionally, the dominant reading and writing experience of Chinese characters (including traditional Chinese number words) and Arabic numbers are both horizontal with a left-to-right direction in Mainland China, which are different in Taiwan where Chinese number words are aligned vertically with a top-to-bottom direction (Hung et al., 2008). Thirdly, the horizontal SNARC effect has been observed with traditional Chinese number words in the classic parity task (see Appendix).

Procedure

The SNARC task was programmed and run using Matlab R2012b Psychtoolbox on a Lenovo ThinkPad E431 laptop. Participants were seated approximately 50 cm from the screen. They made manual responses using the

“F” key with the left index finger for the left-tilted stimuli and the “J” key with the right index finger for the right-tilted stimuli.

During each trial, a white circle (1.5 cm diameter, about 1.72°) appeared firstly for 500 ms as a fixation, followed by a black screen for 400-600 ms randomly to reduce premature responses. After that, the white target stimulus appeared for 150 ms. Reaction time was defined as the interval between the stimulus onset and the key press. Each stimulus was presented centrally on black background, and the next trial would start after a response was made (there was not a time out interval in which participants had to answer before the next trial started). There was an interval of 1000 ms between two successive trials (Figure 1). Each participant subsequently performed only one block with 792 trials, i.e., 9(numbers: 1-9) \times 2(notations: Arabic vs. Chinese) \times 11(degrees: 2-12) \times 2(tilt sides: left vs. right) \times 2(times), and all stimuli were selected in a random order.

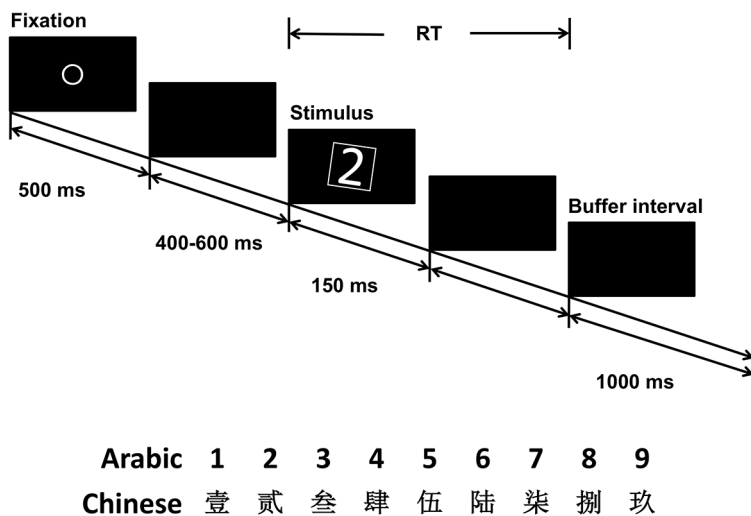


Figure 1. The procedure of one trial in pre-experiment and Experiment 1 (where the stimulus corresponded to the “J” key). The Arabic digits and Chinese verbal numbers used (mixed within a block) are shown at the bottom.

Results

Data analysis was first conducted with SPSS 20.0. Trials with reaction times which were regarded as outliers (i.e., less than 150 ms or more than 1500 ms or outside ± 3 standard deviations from the individual's means) were excluded (2.37%). Four participants with the error rate beyond 15% were excluded from the further analysis, resulting in a final sample of 14 individuals with a mean error rate of 6.60% ($SD = 3.18\%$).

Subsequently, we used MatlabR2012b CFTool to estimate functions between the mean RTs and the degrees (D) of stimuli rotation: $RT(\text{Arabic}) = 403.0 + 215.2 \times D^{-0.6155}$ ($SSE = 152.3$, $r^2 = .9824$); $RT(\text{Chinese}) = 418.7 + 201.0 \times D^{-0.7506}$ ($SSE = 110.8$, $r^2 = .9853$). Both two conditions fit well, indicating that reaction time was associated with tilt angles to some extent (Figure 2). Furthermore, we chose to implement the better Chinese RT function in the formal experiments, to make the RTs for Chinese verbal numbers as close as possible to those for Arabic digits.

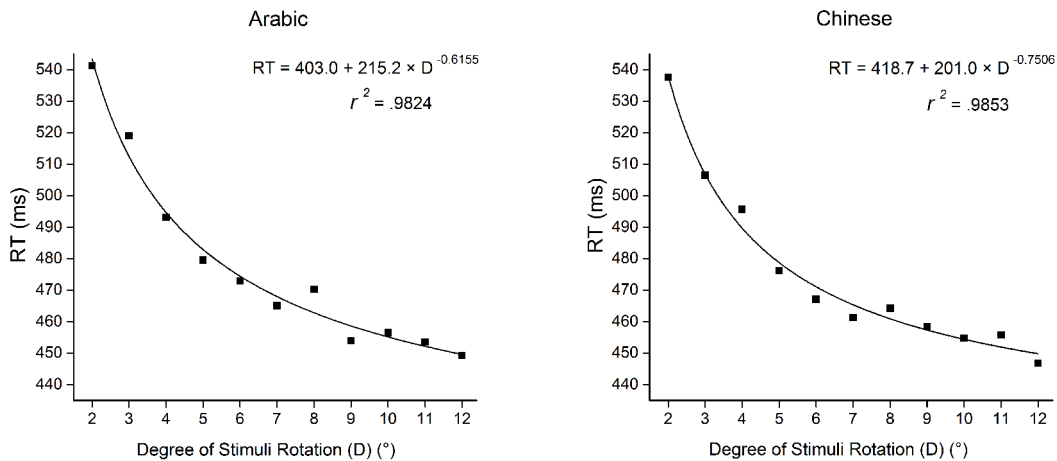


Figure 2. The power function of mean RTs and the rotation degree in pre-experiment.

Experiment 1

Method

Participants

Twenty-five students (21 female and 4 male, mean age = 20.48 years, $SD = 1.86$ years) participated in the study. All of them had normal or corrected-to-normal vision and were right-handed. Chinese was their mother tongue. They were paid 15 RMB (approximately 2.5 USD) for their participation in the experiment. None of these participants contributed to the pre-experiment.

Stimuli

The stimuli were similar to those in pre-experiment. Taking into consideration the pre-experiment result, we decided to initialize the tilt degree of stimuli as 5° and the slight readjustment were made under the control of the adaptive program after each block.

Procedure

The procedure for the participants was almost the same as that in the pre-experiment, except that the program in Psychtoolbox could adapt the tilt degree of Chinese verbal numbers based on the RT difference between these two notations during the procedure.

There were 36 different items: 2(notations: Arabic vs. Chinese) \times 2(orientation: left vs. right) \times 9(numbers: 1-9), which were regarded as a block. Each participant finished 8 blocks (i.e., 288 trials) without breaks. The sequence of stimuli in each block was arranged at random (i.e., Arabic and Chinese stimuli were mixed within blocks).

The program took 5° as the initial value for both the Arabic and Chinese stimuli. After the first block, the program calculated the mean RTs for the Arabic condition (RT_a) and Chinese condition (RT_c) of the 36 trials (18 trials for Arabic and 18 trials for Chinese), except for the outliers (see pre-experiment). Subsequently, if the difference between RT_a and RT_c was 5 ms or larger, the program translated the function which was created in pre-experiment along the Y axis, according to the difference between RT_c and the standard RT_c of the pre-ex-

periment. Then, taking RTa as the ordinate, it found a corresponding rotation degree, which was applied for the Chinese notation in the next block.

After the second block, the program also calculated the RTa and RTc of the new 36 trials. However, it adapted the degree in a new way: if the RTc was 5 ms or larger than RTa, the program would add 0.2° to make the Chinese condition easier; if the RTc was 5 ms or smaller, the program would subtract 0.2° ; or, the parameter would not change. Such parameters would be applied to the next block.

The adaption of the rest of the blocks was very similar to Block 2, except for the third block and onward, the calculations of RTa and RTc would include all trials from the beginning from the third block (i.e., the seventy-third trial). The reason why we adopted such a complicated adapting mechanism was that we wanted to anchor the parameter fast at first, and then fine-tune it (Figure 3). As demonstrated in Figure 3, RTs for Arabic and Chinese notations converged across the eight designated blocks for a single participant.

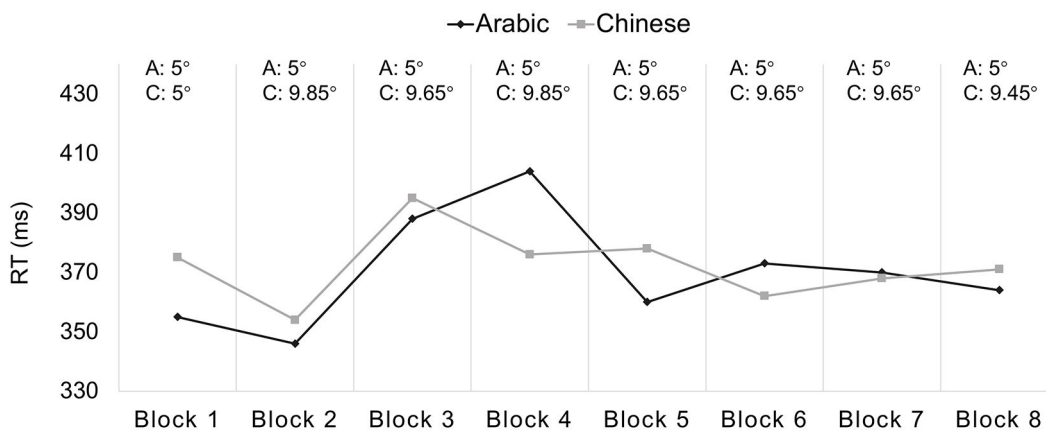


Figure 3. An example for the adaptive program based on data from one participant.

Results

Data analysis was conducted with SPSS 20.0. Trials with reaction times which were regarded as outliers (see the detailed information mentioned in the pre-experiment) were excluded (1.78%). Two participants with an error rate beyond 15% were excluded from further analysis, and one participant was excluded because of a high error rate (67%) on a single condition, resulting in a final sample of 22 individuals with a mean error rate of 5.90% ($SD = 2.83\%$). A correlation analysis was conducted between RTs and accuracy crossed 36 levels (2 notations \times 2 hands \times 9 numbers) and showed that there was no speed-accuracy tradeoff ($r = -.552$). The range of corrected angles of the Chinese condition in the final block was 2.89 – 10.35° ($M = 5.52^\circ$, $SD = 1.86^\circ$).

A three-way repeated ANOVA with Notation (Arabic or Chinese numbers), Hand (left or right), Number (1–9) was conducted. The result revealed that the main effect of hand was significant, $F(1,21) = 9.59$, $p < .01$, $\eta^2 = .31$, the RT for the right hand ($M = 420$ ms, $SD = 9$ ms) was smaller than that for the left hand ($M = 433$ ms, $SD = 10$ ms). The main effect of number was not significant, $F(8,168) = 1.41$, $p = .195$, $\eta^2 = .06$. The main effect of notation was not significant, $F(1,21) = 1.60$, $p = .220$, $\eta^2 = .07$, which meant there was no significant difference between RTs for Arabic digits ($M = 427$ ms, $SD = 9$ ms), and that for Chinese verbal numbers ($M = 425$ ms, $SD = 9$ ms).

None of the two-way interactions were statistically significant (all $p > .1$). However, the three-way interaction was significant, $F(8,168) = 2.65$, $p < .01$, $\eta^2 = .11$. Breaking down the three-way interaction based on notation into two two-way interactions showed that the interaction between the hand and the number was significant for Arabic digits, $F(8,168) = 3.32$, $p < .01$, $\eta^2 = .14$, but not significant for Chinese verbal numbers, $F(8,168) = 0.56$, $p = .808$, $\eta^2 = .03$.

Separate regressions of mean dRT (right-hand response minus the left one) for each participant on different notations were conducted to evaluate the SNARC effect (Fias et al., 1996). The result of regression analysis was significant for Arabic notation situation ($dRT = -2.25 \times \text{Magnitude} - 1.68$, $t(21) = -2.41$, $p < .05$, $d = 0.51$), but not significant in Chinese notation ($dRT = 0.75 \times \text{Magnitude} - 16.68$, $t(21) = 0.84$, $p = .408$, $d = 0.18$). Paired-sample t-test result showed that the slope of the Arabic condition was significantly more negative than that of the Chinese condition, $t(21) = -2.36$, $p < .05$, $d = 0.50$, (Figure 4).

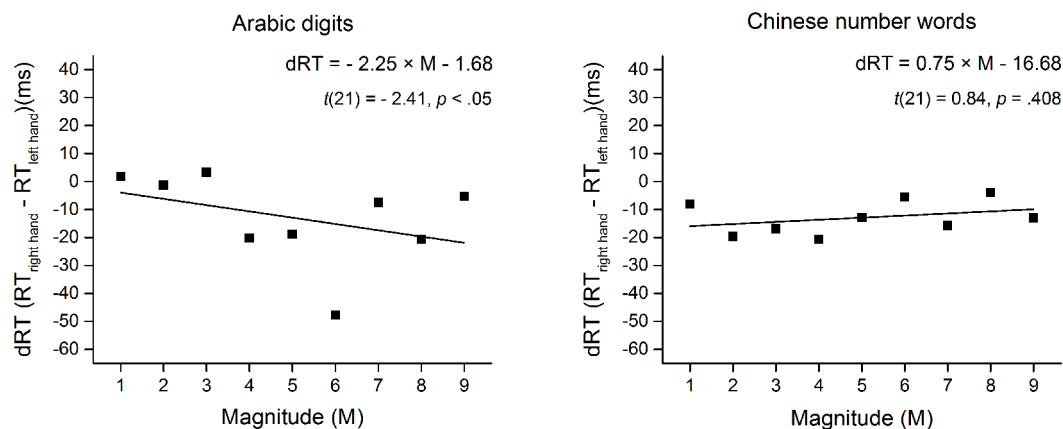


Figure 4. Observed data in Experiment 1 and fitted line representing RT differences between two hands as a function of magnitude for the Arabic digit condition and the verbal number condition.

Discussion

In this perceptual orientation task with an adaptive procedure, we found that the SNARC effect was only observed for Arabic digits, but not for Chinese verbal numbers. In the current experiment, the brief presentation time of the target (150 ms) might make complex Chinese number words difficult to perceptually process or pay attention to it. Thus, we decided to conduct Experiment 2 to replicate the result of Experiment 1 with unlimited presentation time.

Experiment 2

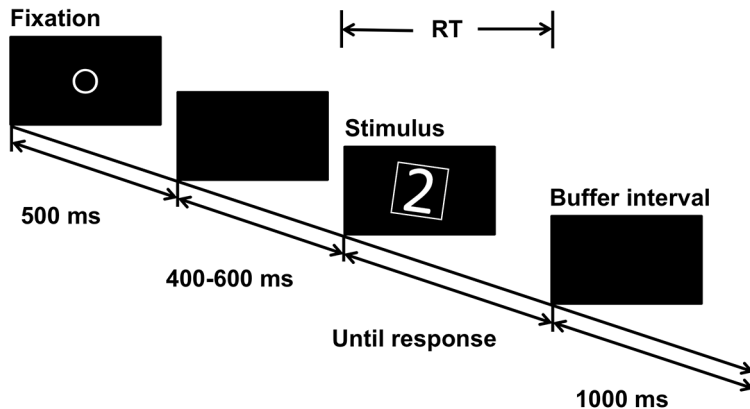
Method

Participants

Twenty-five students (17 female and 8 male, mean age = 21.28 years, $SD = 2.54$ years) participated in the study. The participants in Experiment 2 were met the same conditions as those in Experiment 1 and received the same reward (15 RMB). None of these participants contributed to the pre-experiment or Experiment 1.

Stimuli and Procedure

The stimuli and procedure were similar to Experiment 1. The only modification was the presentation time was changed from 150 ms to indefinitely until a response was made (Figure 5).



Arabic	1	2	3	4	5	6	7	8	9
Chinese	壹	貳	叁	肆	伍	陆	柒	捌	玖

Figure 5. The procedure of one trial in Experiment 2.

Results

Data analysis was conducted with SPSS 20.0. Two participants who ended with extremely large adapted angles in the Chinese condition (28.06° and 24.18°) were excluded from further analyses, because huge angles indicated large performance variations during the adaptation procedure, and general maladaptation and insensitivity to our critical experimental manipulation. One participant was excluded due to the failure of data recording. The final sample contained 22 participants. Trials with reaction times which were regarded as outliers were excluded (2.11%). The mean error rate was 4.85% ($SD = 2.02\%$). The correlation analysis between RTs and accuracy showed that there was no speed-accuracy tradeoff ($r = -.277$). The range of corrected angle of the Chinese condition in the final block was 2.72 - 9.45° ($M = 4.87^\circ$, $SD = 1.61^\circ$).

A three-way repeated ANOVA with Notation (Arabic or Chinese numbers), Hand (left or right), and Number (1-9) was conducted. The result revealed that the main effect of hand was significant, $F(1, 21) = 8.50$, $p < .01$, $\eta^2 = .29$, the RT for the right hand ($M = 459$ ms, $SD = 15$ ms) was smaller than that for the left hand ($M = 476$ ms, $SD = 16$ ms). The main effect of number was not significant, $F(8, 168) = 0.25$, $p = .981$, $\eta^2 = .01$. The main effect of notation was not significant, $F(1, 21) = 2.97$, $p = .100$, $\eta^2 = .12$, which meant there was no significant difference between RTs for Arabic digits ($M = 465$ ms, $SD = 15$ ms), and that for Chinese verbal numbers ($M = 471$ ms, $SD = 16$ ms).

None of the two-way interactions were statistically significant (all $p > .1$). However, the three-way interaction was significant, $F(8, 168) = 2.90$, $p < .01$, $\eta^2 = .12$. Breaking down the three-way interaction based on notation into two two-way interactions showed that the interaction between the hand and the number was significant for

Arabic digits, $F(8, 168) = 2.65$, $p < .01$, $\eta^2 = .11$, but not significant for Chinese verbal numbers, $F(8, 168) = 1.54$, $p = .145$, $\eta^2 = .07$.

The result of regression analysis was significant for the Arabic notation situation, $dRT = -4.95 \times \text{Magnitude} + 9.56$, $t(21) = -4.25$, $p < .001$, $d = 0.91$, but not significant in Chinese notation, $dRT = 1.31 \times \text{Magnitude} - 26.00$, $t(21) = .99$, $p = .334$, $d = 0.21$. Paired-sample t -test result showed that the slope of the Arabic condition was significantly more negative than that of the Chinese condition, $t(21) = -3.56$, $p < .01$, $d = 0.76$, (Figure 6).

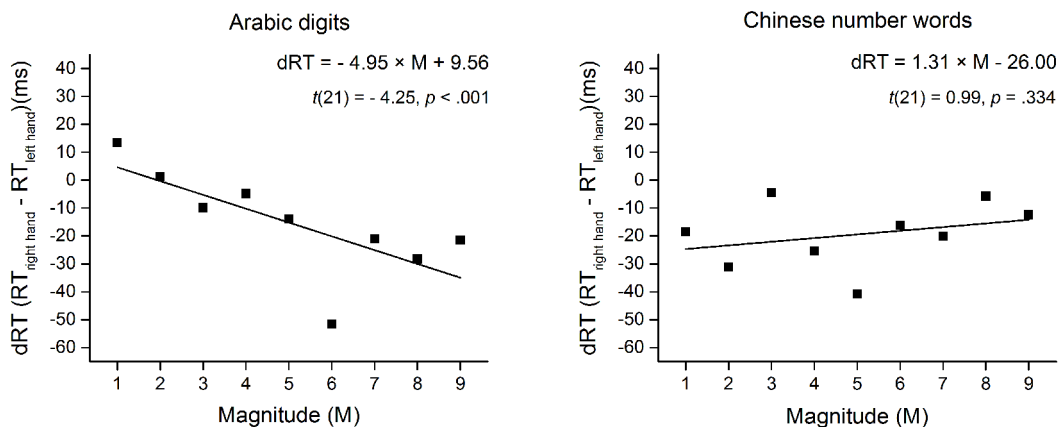


Figure 6. Observed data in Experiment 2 and fitted line representing RT differences between two hands as a function of magnitude for the Arabic digit condition and the verbal number condition.

Discussion

In Experiment 2, with unconstrained stimulus presentation time, we obtained a similar data pattern as in Experiment 1, where the SNARC effect was observed for Arabic digits, but absent for Chinese verbal numbers, indicating that this pattern was reliable and replicableⁱ.

General Discussion

This paper investigated the notation-independence hypothesis with a perceptual orientation task. Results showed that with better maneuver and calibration of task difficulty for two notation conditions, SNARC effect was observed for Arabic digits, but not for Chinese verbal numbers.

The current results, which were obtained from a within-subject design, are in accordance with the previous results reported in a “phoneme monitoring” version of perceptual judgment task (Fias, 2001; Fias et al., 1996). The results are also in accordance with previous results reported in a simple version of a perceptual orientation (upright or right-tilted) study for Arabic stimuli only (Lammertyn et al., 2002). Nonetheless, the current study introduced a novel paradigm based on a simple perceptual orientation task. Compared to the perceptual judgment tasks mentioned above, there are three major advantages, which were mentioned in the introduction of this task. From Table 1, we could see that the perceptual confounds across different numbers (inter-number confounds) and the cross-notation variation had already been reduced compared to Fias et al.’s studies (Fias, 2001; Fias et al., 1996). Consequently, our procedure guarantees that the interaction between Notation and SNARC was uncontaminated by subtle processing time.

Table 1

Mean RT (ms) Comparison Between the Current Study and Fias et al.'s Studies

Study	Numbers									Notations	
	1	2	3	4	5	6	7	8	9	Arabic	Verbal
Experiment 1	430	423	425	423	426	426	431	427	423	427	425
Experiment 2	465	467	469	466	465	469	468	469	470	464	471
Fias et al., 's ^a	569	558	650	624	591	613	568	605	540	591	517

^aFor number comparison, data were obtained from Fias et al., 1996; for notation comparison, data were obtained from Fias et al., 1996 (Arabic) and Fias, 2001 (Verbal).

An Explanation for Different Results From Parity and Perceptual Tasks

We took two factors into consideration when looking at the two opposing results obtained from the notation-independent hypothesis from parity and perceptual tasks. One is automaticity. Automaticity has been defined as the ability to process information without monitoring, where monitoring means an intentional setting of the goal of behavior, and a continuous intentional evaluation of the output of the process (Tzelgov, 1997). Logan (1992) emphasized that the definition of automaticity should specify the learning mechanisms that lead to automaticity, namely, automaticity is usually the result of learning, repetition and practice. Since numbers can be represented in several symbolic formats (e.g., Arabic, English, Roman, and Chinese), or in non-symbolic stimuli (e.g., sets of dots), the question raised that how numbers in the different formats are processed. Cohen Kadosh et al. (2008) addressed this question more specifically: are Arabic digits processed more automatically than verbal numbers?

Arabic digits are frequently encountered in calculation and other numerical comparison contexts in our daily life and conversely, verbal numbers are rarely used. Adopting a learning perspective of automaticity, it is not unreasonable to expect that the strength of the numerical-spatial associations should be influenced by the frequency of use (i.e., experience). Robust findings in the field of numerical cognition confirmed such assumption that Arabic digits can be processed more rapidly than verbal numbers in semantic tasks (Brysbaert, 2005; Cohen Kadosh et al., 2008; Dehaene & Akhavein, 1995; Ito & Hatta, 2003). Jiang et al. (2010), using the classification task, further found that the electrophysiological results for notation effect were in agreement with behavioral results. The shorter reaction times for Arabic digits were accompanied by shorter peak latencies of the VPP (vertex-positive potential) and other later ERP components.

The other is reaction time. As mentioned in the introduction, the slower responses were found to be relevant with the stronger SNARC effect (Gevers et al., 2006; Wood et al., 2008). Gevers et al. (2006) provided a dual-route model to explain it that more intense number magnitude processing should be observed in slower responses because the unconditional route carrying up information on the association between number magnitude and response codes (task-irrelevant) has more time to interfere with the selection of a response button (task-relevant).

Automaticity and reaction time difference have functioned differently in different studies, leading to drastically disparate conclusions regarding automaticity of the SNARC effect. In Nuerk et al.'s (2005) study, slower RTs for verbal numbers (574 ms on average) than those for Arabic digits (487 ms on average) might have amplified the

SNARC effect for verbal numbers, thus overshadowing the processing priority (automaticity) for digits as reflected in SNARC, resulting in a non-interactive effect between SNARC and notation.

In contrast, the slower response for Arabic digits (591 ms, [Fias et al., 1996](#)) than that for verbal numbers (517 ms, [Fias, 2001](#)) in Fias et al.'s SNARC studies, had made it hard to isolate the observed notation by SNARC interaction to a single factor (automaticity), as slower RTs for digits could have simultaneously inflated the SNARC in that condition. In the present study, however, we employed an adaptive program to reduce the processing time difference between Arabic digits and Chinese verbal numbers, where the mean RTs for Arabic digits and verbal numbers were 427 ms and 425 ms (Experiment 1). To our knowledge, this has been the best-controlled study demonstrating strikingly that SNARC effect varies with notation, supporting the claim that the transformation from Arabic input notation to the numerical-spatial associations seems to be more automatic than verbal numbers, an argument that is consistent with many other studies ([Brysbaert, 2005](#); [Cohen Kadosh et al., 2008](#); [Dehaene & Akhavein, 1995](#); [Ito & Hatta, 2003](#); [Jiang et al., 2010](#)).

In summary, the perceptual orientation task might be another alternative way to test the hypothesis regarding the advantage of Arabic digits over verbal numbers, as it can examine the effect of notation on the task-irrelevant dimension when the difference of processing time on the task-relevant dimension is eliminated. In this study, we confirmed [Cohen Kadosh et al.'s hypothesis \(2008\)](#). More importantly, when we confirmed that the SNARC depended on automaticity, we raised a new possible explanation for the notation-independent findings in previous SNARC studies.

The Processing of Notation and Numerical-Spatial Associations

Despite recent advances in the field of numerical cognition, theories regarding the basic architecture of number representations have remained controversial. The abstract-code model ([McCloskey, 1992](#); [McCloskey & Macaruso, 1994, 1995](#)) implies that regardless of input notation, numerical and parity information should be calculated from an abstract, notation-independent representation, which was shown by the classification task ([Dehaene, 1996](#)). The encoding complex hypothesis ([Campbell & Epp, 2004](#)), however, proposed that visual number processing is mediated by modality-specific processes, rather than an abstract representation.

These studies all tried to explain the processing of notation and numerical information by exploring the interaction between notational factor and numerical distance factor in the additive-factors method (i.e., AFM, [Sternberg, 1969](#)). [Nuerk et al.'s \(2005\)](#) study followed the additive-factors method and concluded that there were overlapping stages in processing toward numerical-spatial association as their results showed a non-significant interaction between notation and SNARC effect. However, unlike the numerical distance factor, the SNARC factor might not fit into the additive-factorial design, because the prerequisite of using additive-factors method is that the mental processes should proceed in a series of successive stages whereas most of SNARC studies have suggested that the SNARC effect was based on a dual-route model (e.g., [Gevers et al., 2006](#); [Wood et al., 2008](#)).

Therefore, the traditional paradigms (comparison and parity tasks) used to investigate SNARC effect may not be appropriate to test whether there is a general and abstract magnitude representation of numbers. Nevertheless, in a magnitude-irrelevant task, the SNARC effect can be used to give indications about the automatic activations of the spatially oriented representations ([Tzelgov & Ganor-Stern, 2005](#)). Particularly, for the perceptual orientation task, we provide a dual-route explanation for the SNARC effect (see also [Gevers et al., 2006](#)). In

this parallel model, we assume that: (a) the perceptual orientation information and the numerical/spatial information are first processed along parallel, functionally independent routes can separately activate the left/right response selection stage, and (b) the SNARC effect reflects at least two stages: the activation of semantic magnitude information, and its response-related numerical-spatial associations (Fias & Fischer, 2005; Gevers, Caessens, & Fias, 2005; Gevers et al., 2006). Our framework also coincided with recent neural models of numeric processing (Chen & Verguts, 2010), in which a third layer/stage (decision) was inserted between semantic activation (representation) and response execution. In addition, it should be indicated that this model was established based on a combination of previous conceptions that were proposed to account for both the SNARC effects (Dehaene et al., 1993; Fias et al., 1996; Fias, 2001; Gevers et al., 2006; Wood et al., 2008) and the number Stroop effects (Cohen Kadosh et al., 2008; Henik & Tzelgov, 1982; Schwarz & Ischebeck, 2003). This framework could be useful for future number Stroop experiments, considering the SNARC effect and the number Stroop effect shared similar mechanisms (e.g., the relative speed account, Schwarz & Ischebeck, 2003), and the effect size of the SNARC should be modulated by the processing time of the task-relevant dimension.

Conclusion

To sum up, many researchers have concluded that the perceptual task would be a preferable approach to studying the SNARC effect, but without being able to propose a confound-free paradigm that provides strong evidence to support their hypotheses (e.g., Fias, 2001; Fias et al., 1996; Lammertyn et al., 2002). Using the adaptive perceptual orientation task, we tried to investigate the SNARC effect by disentangling these possible confounds (e.g., controlling for the inter-number variations and reducing the cross-notation processing time difference). We observed that the SNARC effect was sensitive to notation type. One fundamental conclusion derived from the present study is that Arabic digits should be processed more automatically than verbal numbers. We argued that the commonly accepted notation-independence hypothesis (e.g., Nuerk et al., 2005) concerning the SNARC effect should be rejected. Future developmental or cross-culture research may further our understanding of the developmental trajectory and the cultural determination of how the spatial information becomes an integral part of numeric representation.

Notes

- i) The SNARC effect for Arabic digits was larger (Figure 6 vs. Figure 4) with unconstrained presentation time (Experiment 2) than with constrained presentation time (Experiment 1, marginally significant in cross-experiment slope comparison, $t(42) = 1.81$, $p = .077$, $d = 0.55$).
- ii) Stimuli were generated by Words2Picture tool (<https://github.com/psyJT/Words2Pictures>), which was developed for transforming words stimuli to pictures instantly.

Funding

This work was supported by the National Fund for Fostering Talents of Basic Science (Code: J1103601, J1210048) and the National Social Science Fund for A National Study of Chinese Children's Cognitive Development (Code: 14ZDB160).

Competing Interests

The authors have declared that no competing interests exist.

Acknowledgments

We would like to thank the editors and two anonymous reviewers for their helpful comments and suggestions.

Author Contributions

The first two authors contributed equally to this work. In the current article, TJ contributed the idea and experimental design, TG, BL, ZZ carried out the experiments and processed data analysis, and BL, TG, LT, ZZ, XG and TJ completed the writing.

Data Availability

For this study datasets are freely available (see the [Supplementary Materials](#) section).

Supplementary Materials

The datasets of all experiments in this article.

Index of Supplementary Materials

Gong, T., Li, B., Teng, L., Zhou, Z., Gao, X., & Jiang, T. (2019). *Supplementary materials to "The association between number magnitude and space is dependent on notation: Evidence from an adaptive perceptual orientation task"*. PsychOpen. <https://doi.org/10.23668/psycharchives.2376>

References

- Brysbaert, M. (2005). Number recognition in different formats. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 23-42). New York, NY, USA: Psychology Press.
- Butterworth, B. (1999). *The mathematical brain*. London, United Kingdom: Macmillan.
- Campbell, J. I. D. (Ed.). (2005). *Handbook of mathematical cognition*. New York, NY, USA: Psychology Press.
- Campbell, J. I. D., & Epp, L. J. (2004). An encoding-complex approach to numerical cognition in Chinese-English bilinguals. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 58(4), 229-244. <https://doi.org/10.1037/h0087447>
- Chen, Q., & Verguts, T. (2010). Beyond the mental number line: A neural network model of number-space interactions. *Cognitive Psychology*, 60(3), 218-240. <https://doi.org/10.1016/j.cogpsych.2010.01.001>
- Cohen Kadosh, R. C., Henik, A., & Rubinsten, O. (2008). Are Arabic and verbal numbers processed in different ways? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(6), 1377-1391. <https://doi.org/10.1037/a0013413>
- Dehaene, S. (1996). The organization of brain activations in number comparison: Event-related potentials and the additive-factors method. *Journal of Cognitive Neuroscience*, 8(1), 47-68. <https://doi.org/10.1162/jocn.1996.8.1.47>
- Dehaene, S. (2011). *The number sense: How the mind creates mathematics*. New York, NY, USA: Oxford University Press.

- Dehaene, S., & Akhavein, R. (1995). Attention, automaticity, and levels of representation in number processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(2), 314-326. <https://doi.org/10.1037/0278-7393.21.2.314>
- Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General*, 122(3), 371-396. <https://doi.org/10.1037/0096-3445.122.3.371>
- Fias, W. (2001). Two routes for the processing of verbal numbers: Evidence from the SNARC effect. *Psychological Research*, 65(4), 250-259. <https://doi.org/10.1007/s004260100065>
- Fias, W., Brysbaert, M., Geypens, F., & d'Ydewalle, G. (1996). The importance of magnitude information in numerical processing: Evidence from the SNARC effect. *Mathematical Cognition*, 2(1), 95-110. <https://doi.org/10.1080/135467996387552>
- Fias, W., & Fischer, M. H. (2005). Spatial representation of numbers. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 43-54). New York, NY, USA: Psychology Press.
- Fias, W., Lauwereyns, J., & Lammertyn, J. (2001). Irrelevant digits affect feature-based attention depending on the overlap of neural circuits. *Brain Research: Cognitive Brain Research*, 12(3), 415-423. [https://doi.org/10.1016/S0926-6410\(01\)00078-7](https://doi.org/10.1016/S0926-6410(01)00078-7)
- Gevers, W., Caessens, B., & Fias, W. (2005). Towards a common processing architecture underlying Simon and SNARC effects. *European Journal of Cognitive Psychology*, 17(5), 659-673. <https://doi.org/10.1080/09541440540000112>
- Gevers, W., Verguts, T., Reynvoet, B., Caessens, B., & Fias, W. (2006). Numbers and space: A computational model of the SNARC effect. *Journal of Experimental Psychology: Human Perception and Performance*, 32(1), 32-44. <https://doi.org/10.1037/0096-1523.32.1.32>
- Henik, A., & Tzelgov, J. (1982). Is three greater than five: The relation between physical and semantic size in comparison tasks. *Memory & Cognition*, 10(4), 389-395. <https://doi.org/10.3758/BF03202431>
- Hung, Y. H., Hung, D. L., Tzeng, O. J. L., & Wu, D. H. (2008). Flexible spatial mapping of different notations of numbers in Chinese readers. *Cognition*, 106(3), 1441-1450. <https://doi.org/10.1016/j.cognition.2007.04.017>
- Ito, Y., & Hatta, T. (2003). Semantic processing of Arabic, Kanji, and Kana numbers: Evidence from interference in physical and numerical size judgments. *Memory & Cognition*, 31(3), 360-368. <https://doi.org/10.3758/BF03194394>
- Iversen, W., Nuerk, H.-C., & Willmes, K. (2004). Do signers think differently? The processing of number parity in deaf participants. *Cortex*, 40(1), 176-178. [https://doi.org/10.1016/S0010-9452\(08\)70940-7](https://doi.org/10.1016/S0010-9452(08)70940-7)
- Jiang, T., Qiao, S., Li, J., Cao, Z., Gao, X., Song, Y., . . . Chen, C. (2010). Effects of symbol type and numerical distance on the human event-related potential. *Neuropsychologia*, 48(1), 201-210. <https://doi.org/10.1016/j.neuropsychologia.2009.09.005>
- Lammertyn, J., Fias, W., & Lauwereyns, J. (2002). Semantic influences on feature-based attention due to overlap of neural circuits. *Cortex*, 38(5), 878-882. [https://doi.org/10.1016/S0010-9452\(08\)70061-3](https://doi.org/10.1016/S0010-9452(08)70061-3)
- Logan, G. D. (1992). Attention and preattention in theories of automaticity. *The American Journal of Psychology*, 105, 317-339. <https://doi.org/10.2307/1423031>

- McCloskey, M. (1992). Cognitive mechanisms in numerical processing: Evidence from acquired dyscalculia. *Cognition*, 44(1-2), 107-157. [https://doi.org/10.1016/0010-0277\(92\)90052-J](https://doi.org/10.1016/0010-0277(92)90052-J)
- McCloskey, M., & Macaruso, P. (1994). Architecture of cognitive numerical processing mechanisms: Contrasting perspectives on theory development and evaluation. *Cahiers de Psychologie Cognitive / Current Psychology of Cognition*, 13, 275-295.
- McCloskey, M., & Macaruso, P. (1995). Representing and using numerical information. *The American Psychologist*, 50(5), 351-363. <https://doi.org/10.1037/0003-066X.50.5.351>
- Nuerk, H. C., Iversen, W., & Willmes, K. (2004). Notational modulation of the SNARC and the MARC (linguistic markedness of response codes) effect. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, 57(5), 835-863.
- Nuerk, H. C., Wood, G., & Willmes, K. (2005). The universal SNARC effect – The association between number magnitude and space is amodal. *Experimental Psychology*, 52(3), 187-194. <https://doi.org/10.1027/1618-3169.52.3.187>
- Paivio, A. (1975). Perceptual comparisons through the mind's eye. *Memory & Cognition*, 3(6), 635-647. <https://doi.org/10.3758/BF03198229>
- 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(3), 507-522. <https://doi.org/10.1037/0096-1523.29.3.507>
- Sekuler, R., Rubin, E., & Armstrong, R. (1971). Processing numerical information: A choice time analysis. *Journal of Experimental Psychology*, 90(1), 75-80. <https://doi.org/10.1037/h0031366>
- Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica*, 30, 276-315. [https://doi.org/10.1016/0001-6918\(69\)90055-9](https://doi.org/10.1016/0001-6918(69)90055-9)
- Tzelgov, J. (1997). Note. *Consciousness and Cognition*, 6(2-3), 441-451. <https://doi.org/10.1006/ccog.1997.0303>
- Tzelgov, J., & Ganor-Stern, D. (2005). Automaticity in processing ordinal information. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 55-67). New York, NY, USA: Psychology Press.
- Viarouge, A., Hubbard, E. M., & Dehaene, S. (2014). The organization of spatial reference frames involved in the SNARC effect. *Quarterly Journal of Experimental Psychology*, 67(8), 1484-1499. <https://doi.org/10.1080/17470218.2014.897358>
- Wood, G., Willmes, K., Nuerk, H.-C., & Fischer, M. H. (2008). On the cognitive link between space and number: A meta-analysis of the SNARC effect. *Psychology Science Quarterly*, 50(4), 489-525.

Appendix

To make sure that the traditional Chinese number words are able to trigger a SNARC effect, we conducted a pilot experiment with 18 Chinese students (14 female, 4 male, mean age = 19.47 years, $SD = 1.46$ years). The paradigm we used was the classical parity task with traditional Chinese number wordsⁱⁱ. Chinese number word stimuli were the same as those we described in our formal experiment. We included “0” as a stimulus, in order that the numbers of stimuli were equal for odd and even conditions.

We analyzed the data with repeated measured ANOVA and regression analysis to examine a SNARC effect. There was a significant main effect of the number magnitude, $F(9, 153) = 5.90$, $p < .001$, $\eta^2 = .26$, the RT for number 9 was significantly larger than that for number 1, 5, 6, and 7 (all $p < .05$), respectively. The interaction between magnitude and response hand was also significant, $F(9, 153) = 2.28$, $p < .05$, $\eta^2 = .12$. The regression analysis revealed a significant slope on dRT ($dRT = -7.71 \times \text{Magnitude} + 25.38$, $t(17) = -4.30$, $p < .01$, $d = 1.01$), indicating the presence of a SNARC effect for Chinese verbal numbers in the parity judgment task.

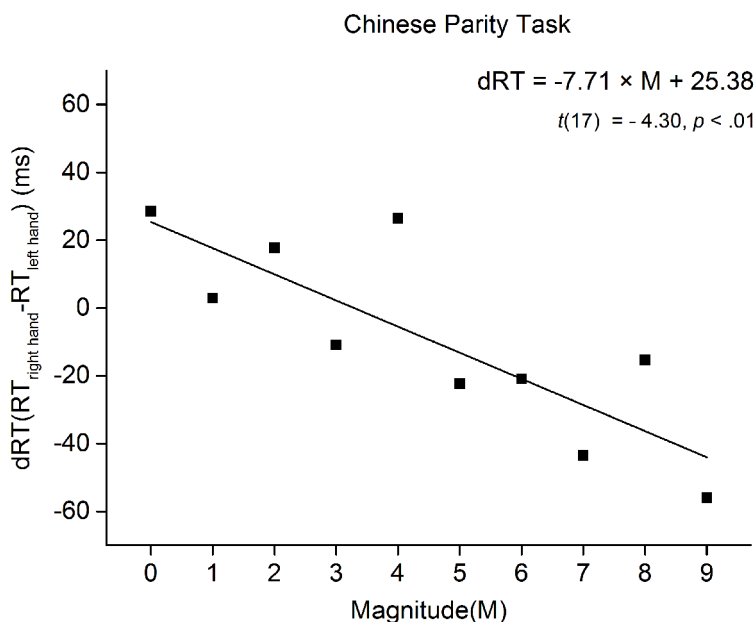


Figure A1. Observed data and fitted line representing RT differences between two hands as a function of magnitude for the traditional Chinese number words in the classical parity task.