The effect of working memory load on the SNARC effect: Maybe tasks have a word to say

Zhijun Deng[1] · Yinghe Chen[1] · Xiaoshuang Zhu[1] · Yanjun Li[1]

Abstract We investigated the effect of working memory load on the SNARC (spatial–numerical association of response codes) effect under different number judgment tasks (parity judgment and magnitude comparison), using a novel dual task. Instead of exerting load over the whole block of number judgment trials, in this dual task, number judgment trials were inserted into each interstimulus interval of an n-back task, which served as the working memory load. We varied both load type (verbal and spatial) and amount (1-load, 2-load, and 3-load). The results indicated that the SNARC effect disappeared even under the 1-load condition for a parity judgment, regardless of the type of load. However, during the magnitude comparison task, the SNARC effect increased with increasing load amounts under spatial load conditions; under verbal load conditions, the SNARC effect decreased with increasing amounts of load, and disappeared during the 3-load task. The difference between the parity and magnitude tasks was not attributable to the interval stimuli or task switching. These findings confirm that different spatial–numerical associations for comparing numerical magnitudes and judgments of parity have different needs with respect to working memory resources.

Keywords SNARC effect · Working memory load · Dual task · Parity judgment task · Magnitude comparison task

Processing numerical information is a key component of many human activities. Cognitive scientists therefore have shown keen interest in how such processing is carried out. How numbers are represented in the mind is one of the many important aspects of such processing. Several studies have indicated that the representation of numerical magnitude is closely related to the mental representation of space. Among the effects explored, the SNARC (spatial–numerical association of response codes) effect is a classical behavioral marker of the spatial coding of numbers (for reviews, see Fias, van Dijck, & Gevers, 2011; Fischer & Fias, 2005; Wood, Willmes, Nuerk, & Fischer, 2008). The SNARC effect, first reported by Dehaene, Bossini, and Giraux in 1993, means that individuals, when completing basic number-processing tasks, typically react faster with their left hands to relatively smaller numbers, and faster with their right hands to relatively larger numbers. The SNARC effect has been shown to be stable and robust, but its mechanism is still the subject of debate over 20 years later.

A primary issue in this debate is whether this spatial association is triggered automatically by the long-term representation underlying number processing, or whether it is temporally constructed in working memory during task execution. The SNARC effect has been observed in hemineglect patients (Halligan, Fink, Marshall, & Vallar, 2003; Hubbard, Piazza, Pinel, & Dehaene, 2009; Zorzi, Priftis, & Umiltà, 2002) and in animals such as 3-day-old domestic chicks (Rugani, Vallortigara, Priftis, and Regolin 2015) and chimpanzees (Adachi, 2014). Studies in these types of subjects have demonstrated that the SNARC effect takes the form of an inherently spatial orientation of a mental number line (Dehaene, 1992; Gevers, Reynvoet, & Fias, 2003; Restle, 1970). Studies of the attentional SNARC effect have also demonstrated that perceiving numbers could cause spatial shifts of attention even if the number is irrelevant (Dodd, Van der Stigchel, Leghari, Fung, & Kingstone, 2008; Fischer, Castel, Dodd, & Pratt, 2003; Hubbard, Piazza, Pinel, & Dehaene, 2005).

✉ Yinghe Chen
chenyinghe@bnu.edu.cn

[1] Institute of Developmental Psychology, School of Psychology, Beijing Normal University, Beijing, China

However, accumulating evidence has indicated a role for working memory in the SNARC effect. For example, the SNARC effect is range-dependent: The number 5 receives faster right responses when the overall range is 1 to 5, but faster left responses when the range is 4 to 9 (Ben Nathan, Shaki, Salti, & Algom, 2009; Dehaene et al., 1993; Fias, Brysbaert, Geypens, & d'Ydewalle, 1996); imagining a clock leads to a reversed SNARC effect (Bächtold, Baumüller, & Brugger, 1998; Ristic, Wright, & Kingstone, 2006); and the SNARC effect becomes less pronounced after the individual experiences an incongruent response–number mapping event (Fischer, Mills, & Shaki, 2010; Pfister, Schroeder, & Kunde, 2013). These studies indicate that the SNARC effect is flexible, context-dependent, and instant, so working memory may play a role in these temporary representations.

Several observations have confirmed that working memory is important to the SNARC effect, but debate on its specific role is ongoing. Using the dual-task paradigm, some researchers have found that the creation of number–space associations during different tasks required different modalities of working memory resources (Herrera, Macizo, & Semenza, 2008; van Dijck, Gevers, & Fias, 2009). Verbal working memory resources were found to be important for the SNARC effect with respect to parity judgment tasks (in which participants judged whether digits are odd or even), whereas visuospatial working memory resources were necessary for comparisons of magnitude (in which participants judged whether digits were smaller or larger than a reference number). However, researchers have also asked participants to react to only a single number in a newly memorized number sequence, and the results showed that the serial positions of the items stored temporally in working memory determined the direction of the SNARC effect (Fias, van Dijck, & Gevers, 2011; Ginsburg, van Dijck, Previtali, Fias, & Gevers, 2014; van Dijck & Fias, 2011). These studies demonstrated that different behavioral measures address different aspects of the relationship between working memory and the SNARC effect.

Building on the work of van Dijck et al. (2009), who found that the SNARC effects under different tasks required different resources, the type of task was taken into serious consideration here in assessing the mechanism underlying the SNARC effect. Answers to the question of whether working memory is needed, or is needed in what form, may vary by task. Both comparison-of-magnitude and judgment-of-parity tasks are common in research into the SNARC effect, but whether they involve the same processes for spatial–numerical associations is still unknown. Clearly, the two methods differ: Magnitude is by definition relevant to magnitude comparison tasks, but this is not the case for judgments of parity. Furthermore, in addition to van Dijck et al. (2009), other researchers have also found the cognitive processes involved in the SNARC effect to be dependent on the type of task or the

task instructions (Georges, Hoffmann, & Schiltz, 2014b; Georges, Schiltz, & Hoffmann, 2015). Georges et al. (2015) found that whether verbal–spatial or visuospatial mechanisms were activated depended on the type of instructions given (i.e., spatial or verbal instructions). For this reason, it was necessary to explore the potential mechanistic differences in the SNARC effect between the magnitude comparison and parity judgment tasks, with a main focus on the role of working memory.

Working memory is usually defined as a cognitive system that consists of a central executive and visual and phonological memory storage slave systems (Baddeley & Hitch, 1974). The central executive system controls and regulates cognitive processes and coordinates the two subsidiary (slave) storage systems (Ardila, 2014; Baddeley, 2012). It is a limited-capacity system responsible for the active maintenance of information and the control of attention (Chow & Conway, 2015; McCabe, Roediger, McDaniel, Balota, & Hambrick, 2010). In the present study, not only the type of load (verbal or spatial), but also the amount of load (1-load, 2-load, 3-load), was taken into consideration.

In the present experiment, a new dual task that combined an n-back (spatial or verbal) task with a number judgment task (either parity judgment or magnitude comparison) was adopted. As a classic measure of working memory, continuous performance on the n-back task can represent an individual's working memory capacity (Conway et al., 2005). Furthermore, the spatial n-back is functionally independent from the verbal n-back task (Fried, Rushmore, Moss, Valero-Cabré, & Pascual-Leone, 2014). To ensure that any effect would be truly attributable to the maintenance of working memory representations (Postle, D'Esposito, & Corkin, 2005), number judgment trials were added to every interstimulus interval of the n-back task. That meant that the offset of a number judgment occurred prior to the next n-back stimulus onset. This task was designed to meet the following two requirements. First, the gradually changed level of load could facilitate exploration of more complex relationships between working memory and the SNARC effect. Second, instead of exerting load across the whole block of number judgment trials, as had been the case in previous works (Herrera et al., 2008; van Dijck et al., 2009), the procedure used here exerted a load on every number judgment trial. Such an adjustment not only roughly equalizes the load of every number, but also increases the system's sensitivity to the load effect for every number. If working memory resources are generally needed for the SNARC effect, then, even at low loads, the magnitude of the SNARC effect should diminish, or it might even disappear entirely. If working memory resources are not needed for the SNARC effect, then the effect would not be affected even at high loads. If working memory resources are needed across some range, then the magnitude of the SNARC effect should change as the amount of load changed.

Experiment 1

Method

Experimental design

For the present experiment, we used a 2 (Type of Task: parity judgment, magnitude comparison) × 2 (Type of Load: spatial, verbal) × 4 (Amount of Load: 0, 1, 2, 3) repeated measures design. Among these variables, the type of task and the type of load were between-subjects variables, and the amount of load was a within-subjects variable. To simplify the description and clarify the design, the condition that combined the magnitude comparison task with the spatial load will here be named the *S-magnitude condition*, the magnitude comparison task combined with the verbal load will be the *V-magnitude condition*, the parity task combined with the spatial load will be the *S-parity condition*, and the parity task combined with the verbal load will be named the *V-parity condition*. Table 1 summarizes the tasks that participants completed under each of these four conditions. The basic dependent variable for number judgment tasks was the reaction time (RT). When the SNARC effect was analyzed, the dependent variable was dRT (i.e., differences in RTs, computed by subtracting the mean RT for left responses from the mean RT for right responses; Fias et al., 1996). When the differences in the SNARC effects among these conditions and tasks were analyzed, the dependent variables were the regression weights of the participants. The dependent variable for the load task was the accuracy of the response, and d' was used for further analysis.

Participants

We recruited 30 participants for the *S-magnitude* condition (18–30 years old, with a mean age of 23 years; 16 female,

14 male), 31 for the *V-magnitude* condition (18–29 years old, with a mean age of 22 years; 18 female, 13 male), 32 for the *S-parity* condition (18–30 years old, with a mean age of 23 years; 20 female, 12 male), and 29 for the *V-parity* condition (18–30 years old, with a mean age of 21 years; 14 female, 15 male). All participants had normal or corrected-to-normal vision and were native Chinese speakers with English as their second language. Participants each received a small monetary reward for participating. The purpose and procedures of the experiment were explained to all of the participants, and they provided informed consent. The experiment was approved by the local institutional ethics committee.

Materials and apparatus

The experiment was performed using the E-Prime 2 Professional software on a 17-in. LCD computer screen (1,280 × 1,024 pixels). Participants provided their responses by pressing specified keys on a standard qwerty computer keyboard, as we describe below.

All stimuli were given in black on a white background. The fixation point was an asterisk (*), 48 points in size. The primary and also a basic task was either parity judgment or magnitude comparison. The numbers 1 to 9 (except 5) were used. The numbers (Arial font, 48-point) appeared at the center of the screen, with each number (1–4, 6–9) repeated 16 times for the basic task, and 20 times for the dual task.

The secondary (load) task was either the verbal or spatial *n*-back task. For the verbal *n*-back, the stimuli were the eight letters "b," "f," " h," "k," "p," "q," "r," and "t" (Times New Roman, 48-point)—the same letters used by Jaeggi, Buschkuehl, Perrig, and Meier (2010) and by Nystrom and colleagues (2000). The letters appeared one at a time at the center of the screen. For the spatial *n*-back, the stimuli were eight possible locations (see Fig. 1), indicated by a black square (25 × 25 mm). In each trial block,

Table 1 Descriptions of tasks under each set of experimental conditions

	S-magnitude (30)	V-magnitude (31)	S-parity (32)	V-parity (29)
Basic task	0-load task (magnitude comparison task)	0-load task (magnitude comparison task)	0-load task (parity judgment task)	0-load task (parity judgment task)
Dual task	1-load task (magnitude comparison + spatial 1-back)	1-load task (magnitude comparison + verbal 1-back)	1-load task (parity judgment + spatial 1-back)	1-load task (parity judgment + verbal 1-back)
	2-load task (magnitude comparison + spatial 2-back)	2-load task (magnitude comparison + verbal 2-back)	2-load task (parity judgment + spatial 2-back)	2-load task (parity judgment + verbal 2-back)
	3-load task (magnitude comparison + spatial 3-back)	3-load task (magnitude comparison + verbal 3-back)	3-load task (parity judgment + spatial 3-back)	3-load task (parity judgment + verbal 3-back)

The experimental conditions were as follows: *S-magnitude*, spatial task with magnitude comparison; *V-magnitude*, verbal task with magnitude comparison; *S-parity*, spatial task with parity judgment; *V-parity*, verbal task with parity judgment. The basic task was a magnitude comparison or parity judgment task with no *n*-back manipulation. The dual task was a magnitude comparison or parity judgment task with varying spatial or verbal *n*-back manipulations (1-, 2-, or 3-back).
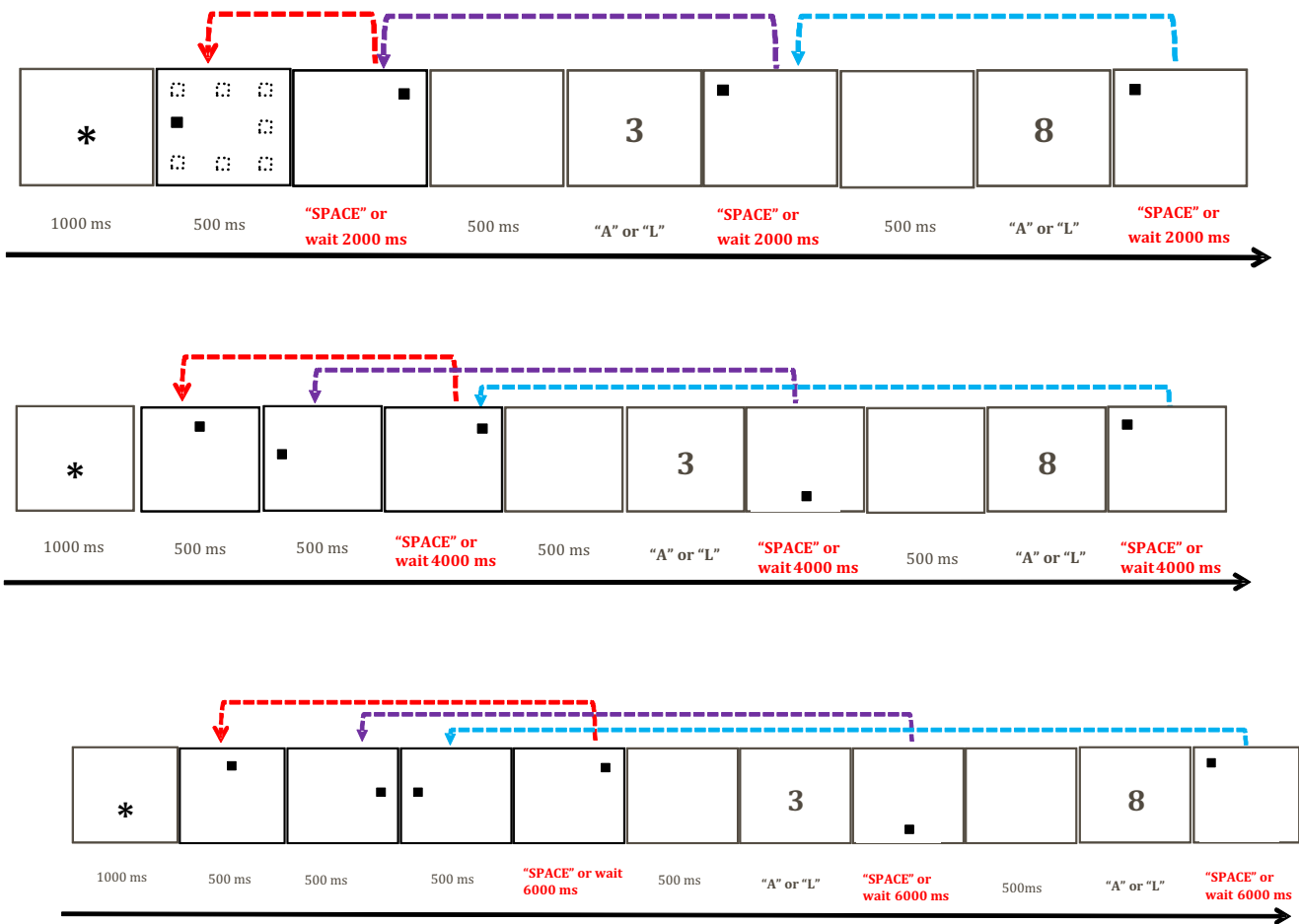
Fig. 1 Schematic illustrations of the experimental procedures for the S-magnitude and S-parity conditions; these three graphs show procedures for the 1-load task, 2-load task, and 3-load task, in that order. For each task, each block began with a fixation (*, 1,000 ms). The preceding *n* stimuli for the first *n*-back trial each lasted for 500 ms. The last stimulus for the first *n*-back trial and the stimuli for the following *n*-back trials required participants to press the space bar with their thumbs if the current stimulus was a repetition of the stimulus that had appeared *n* stimuli previously. Otherwise, the stimulus remained visible for 2×*n* s without any key-pressing response. The interstimulus interval separating the *n*-back stimuli was first a delay of one 500-ms blank screen, followed by one digital judgment task. During the digital judgment task, participants were required to press "A" (with the left forefinger) or "L" (with the right forefinger) to indicate the number's magnitude (S-magnitude condition) or parity (S-parity condition). The *dotted squares in the second picture in the first graph* represent the eight possible locations used in the present experiment

every letter or location occurred five times, on average. Each block included 40 *n*-back judgments, ten of which were "yes" responses. The "yes" trials and the sequence of stimuli appeared in random order.

Procedure

Each participant in each of the experimental groups (S-magnitude, V-magnitude, S-parity, or V-parity) completed four tasks. For the S-magnitude and V-magnitude conditions, the basic task was the magnitude comparison task. For the S-parity and V-parity conditions, the basic task was the parity judgment task. All participants first completed the basic task. Immediately following completion of the basic (0-load) task, participants made parity and magnitude judgments under the load conditions. The orders of the three load tasks (1-load, 2-load, and 3-load) were counterbalanced across participants. After completing the two tasks, the participant rested for 10 min. They finished all four tasks in about one and a half hours.

Basic task For both the magnitude comparison and parity judgment tasks, trials started with a 300-ms fixation (*), prior to a target number appearing at the center of the computer screen. The keys that the participants would press to respond to specific numbers were counterbalanced across two blocks. For example, in the parity judgment task, participants pressed the left key (A) with their left forefinger for odd numbers, and the right key (L) with their right forefinger for even numbers, and the positions were reversed for the next block. In the magnitude comparison task, participants pressed the left key (A) or the right key (L) for numbers numerically less than or greater than 5. The mappings between numbers and keys were

also counterbalanced across the two blocks. After either the participant had responded by pressing a key or 5,000 ms had elapsed with no action, the screen went blank for 1,000 ms, before the next trial started. All 128 trials (8 numbers × 16 presentations) were completed in two blocks for both kinds of basic tasks. The stimulus numbers appeared in random order. Preceding each trial block, participants completed six practice trials to familiarize themselves with the procedure.

Dual task Each participant in the four experimental groups (S-magnitude, V-magnitude, S-parity, or V-parity condition) completed three dual tasks (1-load, 2-load, and 3-load). The procedure was the same for all of these tasks, except for the changes in n ($n = 1$, 2, and 3). There were four blocks per task. Prior to each task, participants completed ten practice trials to familiarize themselves with the procedure.

Each block began with a fixation (*) that lasted for 1,000 ms, and then the dual task started. The load task was the n-back task (spatial n-back task for the S-magnitude and S-parity conditions; verbal n-back task for the V-magnitude and V-parity conditions), which consisted of serial presentation of the stimuli. For the first n-back trial, the preceding n stimuli each lasted 500 ms (for each block, the preceding n stimulus for the first n-back did not require a response). The last stimulus for the first n-back trial and the stimuli for the following n-back trials required participants to press the space bar with their thumb if the current stimulus was a repetition of the stimulus that had appeared n stimuli previously. Otherwise, the stimulus was presented for $2 \times n$ (s) if there was no key-pressing response. (According to the feedback from participants, the time was sufficient to finish comparing and to keep new stimuli in mind for each condition.) To reduce competition for number response resources, participants responded only to the "yes" trials, by pressing the space bar with their thumbs. The interstimulus interval (ISI) separating the n-back stimuli was the sum of the first 500-ms blank screen for delay and of one digital judgment task. The digital judgment task was the primary task, which required participants to press "A" (with the left forefinger) or "L" (with the right forefinger) to respond to its magnitude (for the S-magnitude and V-magnitude conditions) or parity (for the S-parity and V-parity conditions), just as in the basic tasks. The maps between magnitude (or parity) and response keys were counterbalanced across these four blocks. During one block, each of the 40 n-back judgments was followed by each of the 40 number judgments, which meant that participants had to perform the number judgment while retaining a working memory load. Figures 1 and 2 summarize the procedures used in the experiment.

Results and discussion

d′ for the load task The load task was an n-back task, so the d' for each condition was as shown in Table 2. d' was calculated using the formula $d' = Z_{Hit} - Z_{FA}$ (Macmillan & Creelman, 1990), where Hit represents the proportion of hits on "yes" trials [hits/(hits + misses)], and FA represented the proportion of false alarms, when the "yes" response was not correct [false alarms/(false alarms + correct negative)]. A high d' indicated that the target was easily detected, so the value of d' represented the difficulty of the load in part.

A repeated measures analysis of variance (ANOVA) was conducted on d', with the amount of load as a within-subjects variable (1-load, 2-load, and 3-load) and type of task and type of load as between-subjects variables. Only the main effect of load amount reached significance, $F(2, 236) = 92.173$, $p < .0001$, $\eta^2 = .439$. Post-hoc analyses indicated that the difference between the 1-load and 2-load conditions was significant ($p < .001$). Similarly, the 2-load and 3-load conditions differed significantly ($p < .001$). Those results showed that, in these tasks, the two types of load were considered roughly equal in difficulty, but the difficulty increased with increasing amounts of load.

Trade-offs between tasks For all of these dual tasks, none of the correlation coefficients between digital judgment RT and load task accuracy was significant (ranging in magnitude from 0 to .33; $n = 12$, all ps > .05, where n represents the number of trade-offs), and the correlation coefficients between digital judgment accuracy and load task accuracy were all nonsignificant (magnitudes from 0 to .31; $n = 12$, all ps > .05). All these results suggested no trade-offs between digital judgment task and load task. For the digital judgment tasks (the basic and primary tasks of dual tasks), there was no speed–accuracy trade-off (correlations ranged from .02 to .25; $n = 16$, all ps > .05).

RTs for the number-judging task The error rates for the number-judging task under each condition were low (all less than 5%), so only the correct trials were considered for further analysis. The RT data were trimmed to three standard deviations. The proportions of the remaining number-judging data under these conditions were all above 90%. The mean RTs and standard errors of the means (SEMs) were calculated for the remaining number-judging data; see Fig. 3.

A 2 (Type of Task) × 2 (Type of Load) × 4 (Amount of Load) repeated measures ANOVA was performed on the mean RTs. The results revealed a main effect of amount of load, $F(3, 354) = 587.63$, $p < .0001$, $\eta^2 = .833$. Post-hoc analyses indicated that under each condition, the mean RTs all showed the progression 0-load < 1-load < 2-load < 3-load (all pairwise comparisons were significant, $p < .001$).
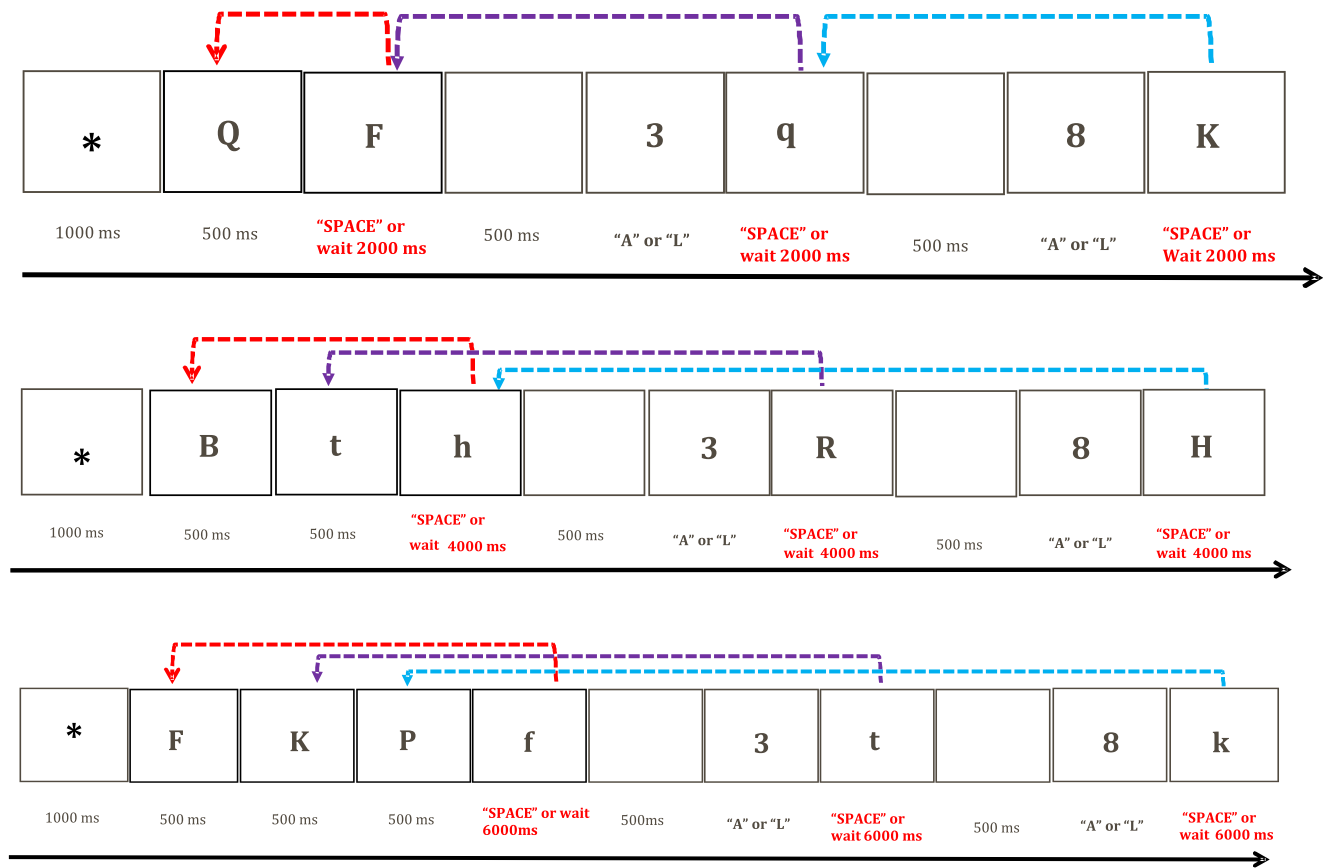
Fig. 2 Schematic illustrations of the experimental procedures for the V-magnitude and V-parity conditions; these three graphs show procedures for the 1-load task, 2-load task, and 3-load task, in that order. For each task, each block began with a fixation (*, 1,000 ms). The preceding *n* stimuli for the first *n*-back trial each lasted 500 ms. The last stimulus for the first *n*-back trial and the stimuli for the following *n*-back trials required participants to press the space bar with their thumbs if the current stimulus was a repetition of the stimulus that had appeared *n* stimuli previously. Otherwise, the stimulus remained visible for 2×*n* s without any key-pressing response. The interstimulus interval separating the *n*-back stimuli was first a delay of one 500-ms blank screen, followed by one digital judgment task. During the digital judgment task, participants needed to press "A" (with left forefinger) or "L" (with the right forefinger) to indicate the number's magnitude (V-magnitude condition) or parity (V-parity condition)

We observed a significant interaction between the amount and type of load, $F(3, 354) = 5.902$, $p < .001$, $\eta^2 = .048$. Simple-effects analyses suggested that, under the 1-load, 2-load, and 3-load conditions, the RTs of spatial load were shorter than those of verbal load: respectively, $F(1, 118) = 14.649$, $p < .0001$, $\eta^2 = .110$; $F(1, 118) = 9.292$, $p < .01$, $\eta^2 = .073$; $F(1, 118) = 5.767$, $p < .05$, $\eta^2 = .047$.

Table 2 Values of d' for the load task under each set of experimental conditions

Load	S-Magnitude	V-Magnitude	S-Parity	V-Parity
1-back	4.61	4.31	4.28	4.02
2-back	3.77	3.65	3.71	3.50
3-back	3.16	3.30	2.86	2.95

$d' = Z_{Hits} - Z_{False\ Alarm}$. The experimental conditions were as follows: *S-magnitude*, spatial task with magnitude comparison; *V-magnitude*, verbal task with magnitude comparison; *S-parity*, spatial task with parity judgment; *V-parity*, verbal task with parity judgment

The interaction of the amount of load and type of task was also significant, $F(3, 354) = 4.661$, $p < .01$, $\eta^2 = .038$. Simple-effects analyses suggested that only in the 2-load condition was the mean RT of the parity task longer than that of the magnitude comparison task, $F(1, 118) = 9.808$, $p < .01$, $\eta^2 = .077$.

In summary, RTs in the number judgment tasks increased with the amount of load. Furthermore, the RTs in the number judgment tasks under verbal load were longer than those under spatial load.

SNARC effect Linear regression was used to analyze the SNARC effect (Fias et al., 1996). For every participant and every number presented, we calculated the dRTs (mean RT for right responses – mean RT for left responses). Then, the dRT values were regressed against the number stimuli (1–4, 6–9), and we used the regression weights of every participant for further analysis.
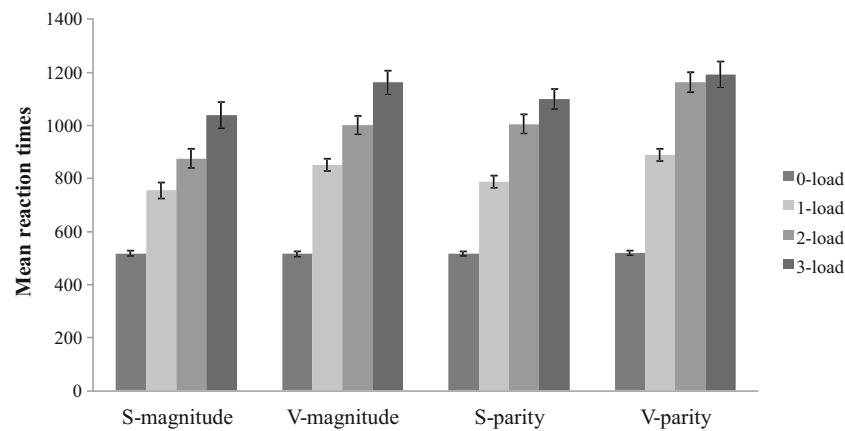
Fig. 3 Mean reaction times (in milliseconds) for the number judgment task under the four conditions. Error bars represent the *SEM*s

In the present work, we evaluated whether the regression weights of the group deviated significantly from zero using *t* tests. Figure 4 shows the mean unstandardized coefficients and *SD*s for each experimental condition.

As is shown in Fig. 4, all the SNARC effects in the basic task (0-load) were significant [S-magnitude condition: $M = -6.08$, $SD = 7.64$, $t(29) = -4.36$, $p < .01$; V-magnitude condition: $M = -7.09$, $SD = 10.04$, $t(30) = -3.93$, $p < .01$; S-parity condition: $M = -7.31$, $SD = 8.67$, $t(31) = -4.77$, $p < .01$; V-parity condition: $M = -6.10$, $SD = 7.91$, $t(28) = -4.16$, $p < .01$]. When the primary task was parity judgment, no matter the type of load (verbal or spatial) and no matter the amount of load (1-load, 2-load, or 3-load), the sizes of the SNARC effect all became nonsignificant. For the S-parity condition, the means and *SD*s for each load condition were $M = -0.92$, $SD = 17.43$, $t(31) = -0.30$, $p > .05$ (1-load); $M = -3.90$, $SD = 25.27$, $t(31) = -0.87$, $p > .05$ (2-load); $M = -5.88$, $SD = 38.25$, $t(31) = -0.87$, $p > .05$ (3-load). For the V-parity condition, the means and *SD*s for each load condition were $M = -3.09$, $SD = 23.04$, $t(28) = -0.72$, $p > .05$ (1-load); $M = -4.58$, $SD = 23.36$, $t(28) = -1.05$, $p > .05$ (2-load); $M = -6.15$, $SD = 35.85$, $t(28) = -0.92$, $p > .05$ (3-load).

However, when the primary task was magnitude comparison and the type of load was spatial *n*-back, the sizes of the SNARC effect increased with increasing load [S-magnitude condition: $M = -12.94$, $SD = 31.20$, $t(29) = -2.27$, $p < .05$ (1-load); $M = -15.56$, $SD = 41.42$, $t(29) = -2.06$, $p < .05$ (2-load); $M = -30.16$, $SD = 40.75$, $t(29) = -4.05$, $p < .01$ (3-load)]. When the primary task was magnitude comparison and the type of load was verbal *n*-back, the sizes of the SNARC effect were greater than that in the 0-load task, but they decreased with increasing load, and even became nonsignificant under the 3-load task [V-magnitude condition: $M = -32.31$, $SD = 40.71$, $t(30) = -4.42$, $p < .01$ (1-load); $M = -24.36$, $SD = 48.87$, $t(30) = -2.78$, $p < .01$ (2-load); $M = -15.36$, $SD = 61.63$, $t(30) = -1.39$, $p > .05$ (3-load)].

To further substantiate the differential influence of type and amount of load on the SNARC effect under parity and magnitude tasks, a 2 (Type of Task) × 2 (Type of Load) × 4 (Amount of Load) repeated measures ANOVA was performed on the regression weights. To control the influence of RT (van Dijck et al., 2009; Wood et al., 2008), the differences in RT between the baseline and load conditions
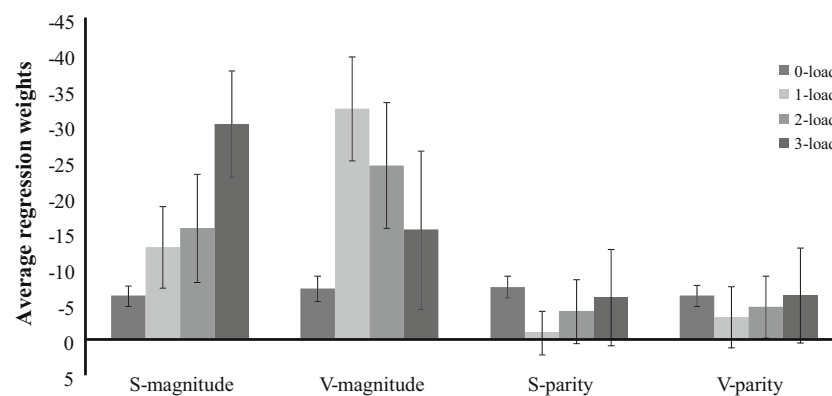
Fig. 4 Sizes of the SNARC effects (*β* values) and their significance in all four experimental conditions. All the sizes for the basic task (0-load) in these four conditions were significant. The SNARC effects in the S-parity and V-parity conditions were abolished when under the working memory load, regardless of the amount of load (1-load, 2-load, or 3-load). In contrast, the magnitudes of the SNARC effects were all significant, regardless of the amount of load, in the S-magnitude condition; in the V-magnitude conditions, the SNARC effects were significant in the 1-load and 2-load tasks, but they became nonsignificant under the 3-load condition

(mean of RT load – RT baseline) served as a covariate variable. The results revealed a main effect of type of task, $F(1, 117) = 13.836$, $p < .0001$, $\eta^2 = .106$, as well as significant interactions between amount and type of load, $F(3, 351) = 3.215$, $p < .05$, $\eta^2 = .027$, and amount of load and type of task, $F(3, 351) = 3.029$, $p < .05$, $\eta^2 = .025$. All the other main effects and interactions were nonsignificant.

Considering that the effects of the four conditions were different, and that the triple interaction effect was not significant, we performed another repeated measures ANOVA on the regression weights with the amount of load as a within-subjects variable and the condition (S-magnitude, S-parity, V-magnitude, or V-parity) as a between-subjects variable, with the difference in RTs between the baseline and the load conditions as a covariate variable. The results revealed that the main effect of the conditions reached significance, $F(3, 117) = 4.677$, $p < .01$, $\eta^2 = .107$. The interaction between the conditions and the amount of load also reached significance, $F(9, 351) = 2.709$, $p < .01$, $\eta^2 = .065$. Simple-effects analysis suggested that, under the 0-load baseline, there was no difference between the four conditions (all $ps > .05$), but in the S-magnitude condition, the size of the SNARC effect under the 3-load condition was larger than under 0-load ($p < .0001$). Under the V-magnitude condition, the size of the 1-load condition was also larger than under 0-load ($p < .01$). However, in the parity tasks, no significant differences emerged between the load conditions (1-load, 2-load, and 3-load) and the 0-load condition (all $ps > .05$), for both the S-parity and V-parity conditions.

In sum, our results showed that the SNARC effects under parity judgment and magnitude comparison were stable (as indicated by the basic tasks), whereas the requirements for working memory resources under these two tasks were different (as indicated by the load tasks). The SNARC effects under parity judgment were all abolished, regardless of the type and the amount of load, whereas the SNARC effects under magnitude comparison became stronger as the amount of spatial load increased, but the SNARC effect under the verbal load condition became stronger under the 1-load condition, and the difficulty decreased as the amount of load increased further, disappearing under the 3-load. Additionally, although the RTs of digital judgment and the *SD*s of the SNARC effect increased with the amount of load, the SNARC effect did not always increase, as well.

Experiment 2

The results of Experiment 1 showed that the SNARC effect in parity judgment disappeared altogether, but that the effect increased in magnitude comparison, regardless of the type or amount of load. However, this begs the question of whether the pattern was caused by the intervening stimuli or the act of switching between two tasks. For this reason, in Experiment 2, whether the same interference could be observed in the interval task and the switch task was tested further. During these two tasks, we included no working memory load manipulation, but the interval stimuli and task switching were as in Experiment 1. Thus, if the interval stimuli or the task switching caused the differences between the trends in the changes under load from those with 0-load both in the parity and magnitude tasks, the difference would still be observable in Experiment 2. However, if the working memory load caused these differences, then the SNARC effects in these tasks would be similar to each other in Experiment 2.

Method

Experimental design

To simplify and clarify the design, the condition that combined the comparison of the magnitude task with the spatial interval stimuli or the spatial judgment task was here named the *S-magnitude-control* condition; the magnitude comparison task combined with the verbal interval stimuli or the verbal judgment task was named the *V-magnitude-control* condition; the parity task combined with the spatial interval stimuli or the spatial judgment task was named the *S-parity-control* condition; and the parity task combined with the verbal interval stimuli or the verbal judgment task was named the *V-parity-control* condition.

Participants

In all, 25 participants took part in the S-magnitude-control condition (18–27 years old, with a mean age of 21 years; 13 female, 12 male), 26 in the V-magnitude-control condition (18–29 years old, with a mean age of 22 years; 14 female, 12 male), 27 in the S-parity-control condition (19–30 years old, with a mean age of 23 years; 13 female, 14 male), and 26 in the V-parity-control condition (20–30 years old, with a mean age of 24 years; 15 female, 11 male). All of these participants were native Chinese speakers with normal or corrected-to-normal vision.

Materials and procedure

Each participant in each of the experimental groups (S-magnitude-control, V-magnitude-control, S-parity-control, or V-parity-control) completed three tasks, including the basic task, the interval task (number judgment task with spatial or verbal stimuli interval), and the switch task (switching between the number judgment task and the spatial/verbal judgment task). The order in which the participants performed these three tasks was randomized.

Basic task This was the same as Experiment 1. All 128 trials were completed in two blocks in both the parity task and the magnitude judgment task.

Interval task In this task, interstimulus intervals were inserted into each number judgment trial. Each trial started with the digital judgment task, which required participants to press "A" or "L" to indicate a number's magnitude (for the S-magnitude-control and V-magnitude-control conditions) or parity (for the S-parity-control and V-parity-control conditions), just as in the basic tasks. The maps between magnitude (or parity) and response keys were counterbalanced across these two blocks. Then the interval stimuli (one or two squares located in eight possible locations in the S-magnitude-control and S-parity-control conditions; one or two letters located in the center of the screen in the V-magnitude-control and V-parity-control conditions) were presented for several seconds (random among 500, 800, 1,000, 1,500, 2,000, 4,000, and 6,000 ms) and did not require a response. The times 2,000, 4,000, and 6,000 ms were consistent with the times that participants had needed to wait in Experiment 1; 500, 800, 1,000, and 1,500 ms represented the times that participants might have needed if they pressed the key in Experiment 1. Then the screen went blank for 500 ms, until the beginning of the next trial. All 128 trials were completed in two blocks. Preceding each trial block, participants completed six practice trials to become familiar with the procedure.

Switch task In this task, a spatial or verbal judgment task was added to each number judgment trial, so participants had to alternate back and forth between the number judgment task and the spatial or verbal judgment task.

Each trial started with the digital judgment task, just as in the basic and interval tasks. Then the spatial (in the S-magnitude-control and S-parity-control conditions) or the verbal (in the V-magnitude-control and V-parity-control conditions) judgment task took place. Participants had to judge whether two squares were in a vertical or horizontal line or whether two letters had the same pronunciation. If the answer was "yes," they needed to press the space bar with their thumbs; if the answer was "no," they did not need to press any key, but rather needed to wait for 2 s. Then the screen went blank for 500 ms, until the next trial started. All 128 trials were completed in two blocks in both kinds of basic tasks. Preceding each trial block, participants completed six practice trials to familiarize themselves with the procedure.

Figures 5 and 6 summarize the main procedures used in Experiment 2.

Results and discussion

The error rates for the number-judging tasks under each set of conditions were low (all less than 5%), so only the correct

trials were considered for further analysis. The RT data were trimmed to three standard deviations. The proportions of the remaining number-judging data under these conditions were all above 95%. The error rates of the spatial and verbal judgment tasks were all low (less than 5%) in the switch task. The mean RTs and standard errors of the means (SEMs) were calculated for the remaining number-judging data; see Fig. 7.

Basic task As is shown in Fig. 8, all the SNARC effects in the basic task were significant [S-magnitude-control condition: $M = -7.59$, $SD = 16.67$, $t(24) = -2.28$, $p < .05$; V-magnitude-control condition: $M = -7.33$, $SD = 17.43$, $t(25) = -2.143$, $p < .05$; S-parity-control condition: $M = -5.46$, $SD = 8.17$, $t(26) = -3.47$, $p < .01$; V-parity-control condition: $M = -6.58$, $SD = 10.86$, $t(25) = -3.09$, $p < .01$], and there were no differences among them [$F(3, 100) = 0.127$, $p > .05$].

Interval task A 2 (Type of Task) × 4 (Condition) repeated measures ANOVA was performed on the mean RTs, with the type of task (basic task and interval task) as a within-subjects variable and the four conditions (S-magnitude-control, S-parity-control, V-magnitude-control, and V-parity-control) as a between-subjects variable. The results revealed a main effect of the type of task, $F(1, 99) = 300.69$, $p < .0001$, $\eta^2 = .752$. All the other main effects and interactions were nonsignificant. The results indicated that the RTs for the interval tasks were longer than those for the basic tasks.

As is shown in Fig. 8, all the SNARC effects in the interval task were also significant [S-magnitude-control condition: $M = -9.18$, $SD = 21.03$, $t(24) = -2.18$, $p < .05$; V-magnitude-control condition: $M = -7.85$, $SD = 16.55$, $t(25) = -2.42$, $p < .05$; S-parity-control condition: $M = -8.26$, $SD = 11.35$, $t(26) = -3.78$, $p < .01$; V-parity-control condition: $M = -6.51$, $SD = 12.29$, $t(25) = -2.69$, $p < .05$]. A 2 (Type of Task) × 4 (Condition) repeated measures ANOVA was conducted on the regression weights, and none of the main effects or interactions were significant ($ps > .05$). The results indicated that the stimulus interval influenced the RTs but did not influence the SNARC effect.

Switch task A 2 (Type of Task) × 4 (Condition) repeated measures ANOVA on the mean RTs was conducted with the type of task (basic and switch task) as a within-subjects variable and the four conditions (S-magnitude-control, S-parity-control, V-magnitude-control, and V-parity-control) as a between-subjects variable. The results revealed a main effect of type of task, $F(1, 99) = 508.53$, $p < .0001$, $\eta^2 = .837$. All the other main effects and interactions were nonsignificant. The results indicated that the RTs for the switch tasks were longer than those for the basic tasks.

As is shown in Fig. 8, all the SNARC effects in the switch task were significant [S-magnitude-control condition, $M = $
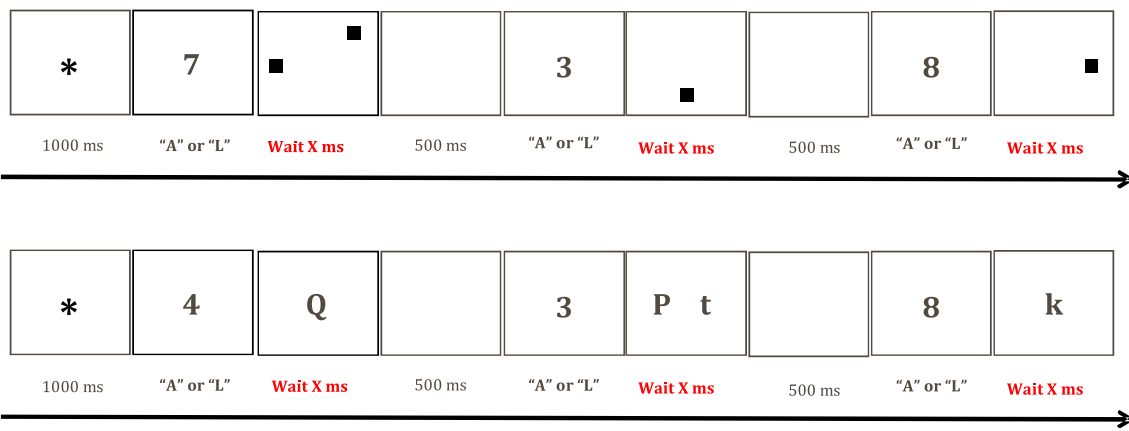
Fig. 5 Schematic illustrations of the experimental procedures for the interval task under all four sets of conditions. These two graphs show the procedures for the S-magnitude-control/S-parity-control conditions and the V-magnitude-control/V-parity-control conditions. For each task, each block began with a fixation (*, 1,000 ms). The interstimulus interval separating trials was 500 ms of blank screen. X represents that the presentation time was random chosen from among 500, 800, 1,000, 1,500, 2,000, 4,000, and 6,000 ms

−10.39, $SD = 24.81$, $t(24) = −2.09$, $p < .05$; V-magnitude-control condition, $M = −8.05$, $SD = 19.07$, $t(25) = −2.15$, $p < .05$; S-parity-control condition, $M = −6.59$, $SD = 15.22$, $t(26) = −2.25$, $p < .05$; V-parity-control condition, $M = −10.33$, $SD = 15.17$, $t(25) = −3.47$, $p < .01$]. A 2 (Type of Task) × 4 (Condition) repeated measures ANOVA was performed on the regression weights, and none of the main effects or interactions were significant (ps > .05).

Experiment 2 ruled out two possible interference effects. First, intervening stimuli did not cause the difference between parity and magnitude judgments, as indicated by the fact that the SNARC effects were the same for the interval and basic tasks. Second, switching between two tasks did not cause the difference, either, because the SNARC effect was the same in the switch task as in the basic task. Taking into account the results of both Experiments 1 and 2, we can conclude that what primarily affected the differences in the SNARC effects during these tasks was the working

memory load imposed on the mechanisms responsible for the number and spatial associations.

General discussion

The effect of working memory load on the processing of the numerical–spatial representations underlying the SNARC effect was here investigated by imposing different amounts of verbal or spatial working memory load during each single-number parity judgment or magnitude comparison trial. Because it was not possible to confirm whether spatial working memory resources were needed by the magnitude comparison task in our present study (the discussion will be continued in more detail in the sixth paragraph), the results of the present work could not definitely indicate whether there was a dissociation between the type of working memory load and the type of digital judgment task, as was described by Herrera
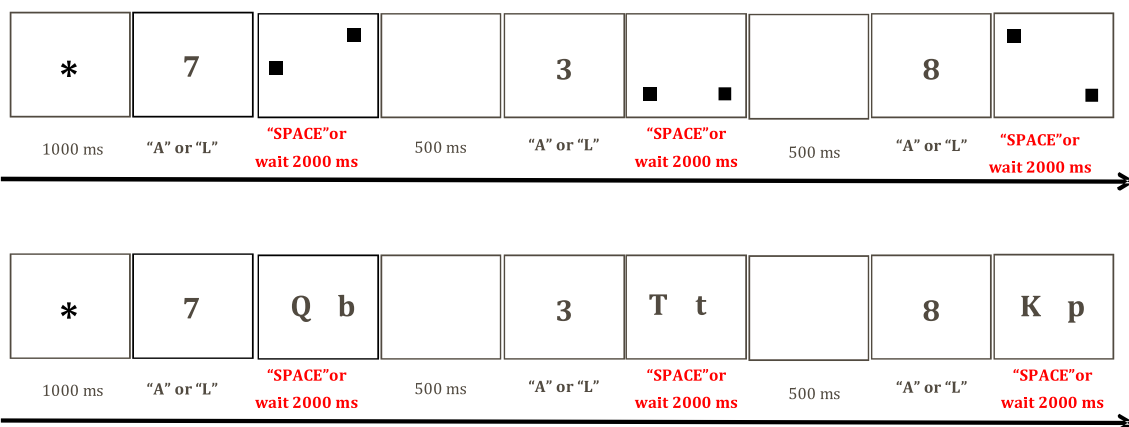
Fig. 6 Schematic illustrations of the experimental procedures for the switch task under all four sets of conditions. These two graphs are the procedures for the S-magnitude-control/S-parity-control and V-magnitude-control/V-parity-control conditions, in that order. For each task, each block began with a fixation (*, 1,000 ms). The interstimulus interval separating trials was 500 ms of blank screen
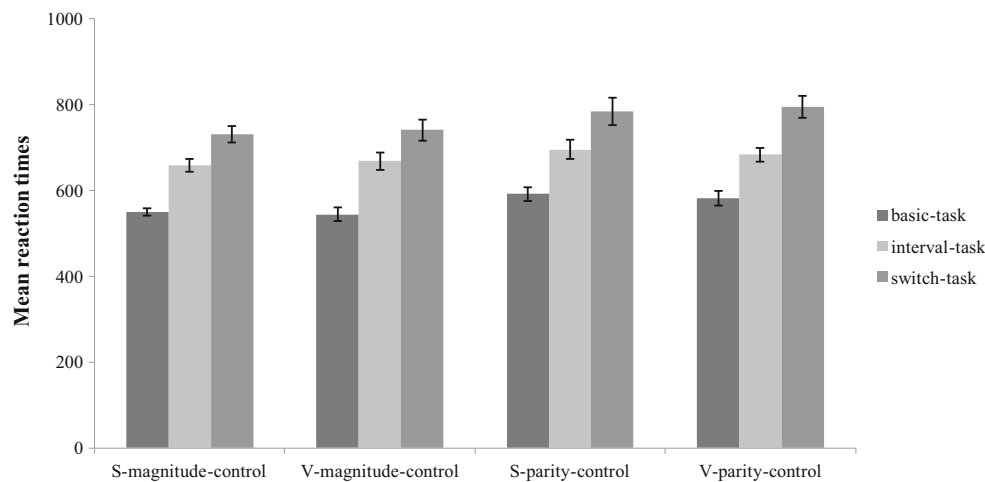
Fig. 7 Mean reaction times (in milliseconds) for the number judgment task under these four control conditions. *Error bars* represent the *SEM*s

et al. (2008) and van Dijck et al. (2009). Our findings did show a dissociation between the amount of working memory load and the type of digital judgment task, though. The SNARC effect under parity judgment demonstrated greater sensitivity to working memory resources than did comparisons of magnitude.

The factors that may influence the different levels of working memory resources engaged in the parity judgment and magnitude comparison tasks remain uncertain. Thus, it was first necessary for us to rule out possible confounding factors. First, we had to confirm that the differences were not due to the interval or the task switching, because the SNARC effects in the interval and switch tasks were all the same as in the basic tasks, under both the parity judgment and magnitude comparison tasks in Experiment 2. Second, we found that this was not due to the changes in RTs, because there was a notable dissociation of the load effects on the RTs of number judgments and on the slope values of the SNARC effect. For example, the RTs of number judgments increased as the amount of load increased, whereas the sizes of the SNARC effect did not always increase across these levels. Third, we found that

this effect was not due to a change of interpersonal variability. Although individual differences are also an important factor that affects the SNARC effect (Georges, Hoffmann, & Schiltz, 2014a; Hoffmann, Pigat, & Schiltz, 2014; Viarouge, Hubbard, & McCandliss, 2014), this was not the reason for the difference between parity and magnitude tasks in the present study. The spatial working memory load condition here served as an example. The *SD*s of the participant's regression weights with the 3-load task were larger than those with the 1-load condition in both tasks, but the SNARC effect disappeared from the parity task but not from the magnitude task.

One possible factor worth considering, regarding the difference between the parity and magnitude tasks, is what the determining factor for the different levels of working memory resources engaged might be. The difference between these two tasks may come from either of two stages, or from both. One is at the response selection stage, as was indicated in several studies that found that the to-be-discriminated alternative responses must be represented in working memory (Ansorge, & Wühr, 2004; Gevers, Verguts, Reynvoet, Caessens, & Fias, 2006). For comparisons of magnitude, the
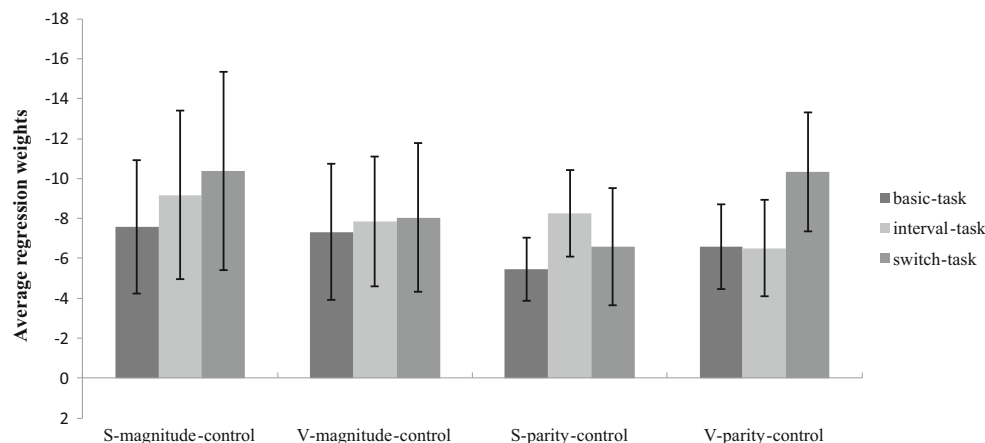
Fig. 8 Sizes of the SNARC effects (β values), which were significant under all four control conditions

same responses were associated with numbers that were smaller or larger than the referent, but the responses alternated with each number in parity judgment. Because more alternative responses were needed in the parity than in the magnitude task (van Dijck et al., 2009), this made it easy to infer that more working memory resources were needed for the parity than for the magnitude task. The other stage was in the processing of magnitude information. Obviously, magnitude comparison drew directly from the magnitude information, which may have helped activate "the mental number line," but this was not the case for parity judgment, in which accessing the magnitude information required a switch from judging parity (Bae, Choi, Cho, & Proctor, 2009). Further studies will be needed to point out the exact stage at which the spatial–numerical associations consume working memory resources.

The design of the present study may also have increased the sensitivity to the amount change in working memory load in the parity task. Here, participants needed to update the working memory load for every single number trial, but in previous studies participants had maintained working memory load across a sequence of numbers (16 number trials; Herrera et al., 2008; van Dijck et al., 2009). This difference in study designs may mean that even small changes in working memory load can impact the SNARC. The SNARC effect under parity judgment being sensitive to even very low loads may be supported by the studies that have used parity judgment to show that the SNARC effect can be reduced just after incongruent trials (Fischer et al., 2010; Pfister et al., 2013). Other studies that have used the parity task also showed no SNARC effect when three sequences were maintained in descending order (e.g., 5–4–3; Lindemann, Abolafia, Pratt, & Bekkering, 2008). The present study is the first one to explicitly show that working memory resources are largely needed by the SNARC effect in the parity task.

The fact that the SNARC effect under magnitude comparison was influenced differently by different types and amounts of load requires that these factors be taken into more detailed consideration. First, it must be acknowledged that the SNARC effect under magnitude comparison also requires working memory resources, because it disappeared in the 3-load verbal condition. The assumption of autonomy rendered the system immune to the influence of any other task being executed concurrently (Palmeri, 2002). The conclusion that verbal working memory resources are needed for magnitude tasks was consistent with the findings reported by Gevers and colleagues (2010), who observed that verbal–spatial coding was the dominant factor driving the SNARC effects in both the parity judgment and magnitude comparison tasks.

Second, our results showed that, in the magnitude comparison task, the SNARC effect increased with increasing amounts of spatial working memory load. This indicated that spatial working memory resources might not be needed in the magnitude task. It was also possible that spatial working memory resources might be needed less than verbal working memory resources, or that they would not be needed to the same extent as in the parity task, because the spatial working memory load used in the present study probably did not load working memory as extensively as in the study by van Dijck et al. (2009), who manipulated sequence length according to participants' working memory spans. No matter which of these is the case, all scenarios indicate that spatial working memory is not needed by the SNARC effect under magnitude comparison tasks to the same extent as under the parity judgment task. This leaves the question of why the SNARC effect became more pronounced as the amount of spatial load increased in the magnitude comparison task. It may be that the spatial material acted as a cue to activate brain regions associated with spatial operations (ventral intraparietal cortex). This area overlapped with the horizontal aspect of the intraparietal sulcus, the part of the brain used to process numbers (Dehaene, Spelke, Pinel, Stanescu, & Tsivkin, 1999; Van Opstal, Santens, & Ansari, 2012). In this way, increasing the spatial load leads to an increasingly closer relationship between numbers and space.

In conclusion, the present results indicate that working memory is needed for both parity and magnitude judgment tasks, but the amounts and types of the working memory resources needed are different, as indicated by the differences in the SNARC effect. The results of the present work confirmed that the mechanisms underlying the SNARC effects created during different tasks were not uniform. This is the first study to indicate that different amounts of working memory resources are needed for the association of numbers and space in magnitude comparison and parity judgment tasks. These findings remind us that the type of task is also a key element in the exploration of the nature of the SNARC effect. However, additional direct methods and neuroimaging studies will be needed to determine the mechanisms underlying these tasks.

References

Adachi, I. (2014). Spontaneous spatial mapping of learned sequence in chimpanzees: Evidence for a SNARC-like effect. *PLoS ONE, 9,* e90373. doi:10.1371/journal.pone.0090373

Ansorge, U., & Wühr, P. (2004). A response-discrimination account of the Simon effect. *Journal of Experimental Psychology: Human Perception and Performance, 30,* 365–377. doi:10.1037/0096-1523.30.2.365

Ardila, A. (2014). Working memory. In R. R. Heredia & J. Altarriba (Eds.), *Foundations of bilingual memory* (pp. 223–234). New York: Springer.

Bächtold, D., Baumüller, M., & Brugger, P. (1998). Stimulus–response compatibility in representational space. *Neuropsychologia, 36,* 731–735.

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology, 63,* 1–29. doi:10.1146/annurev-psych-120710-100422

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47–90). New York: Academic Press. doi:10.1016/S0079-7421(08)60452-1

Bae, G. Y., Choi, J. M., Cho, Y. S., & Proctor, R. W. (2009). Transfer of magnitude and spatial mappings to the SNARC effect for parity judgments. *Journal of Experimental Psychology Learning Memory and Cognition, 35,* 1506–1521. doi:10.1037/a0017257

Ben Nathan, M., Shaki, S., Salti, M., & Algom, D. (2009). Numbers and space: Associations and dissociations. *Psychonomic Bulletin & Review, 16,* 578–582. doi:10.3758/PBR.16.3.578

Chow, M., & Conway, A. R. A. (2015). The scope and control of attention: Sources of variance in working memory capacity. *Memory & Cognition, 43,* 325–339.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12,* 769–786. doi:10.3758/BF03196772

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition, 44,* 1–42. doi:10.1016/0010-0277(92)90049-N

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General, 122,* 371–396. doi:10.1037/0096-3445.122.3.371

Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., & Tsivkin, S. (1999). Sources of mathematical thinking: Behavioral and brain-imaging evidence. *Science, 284,* 970–974.

Dodd, M. D., Van der Stigchel, S., Leghari, M. A., Fung, G., & Kingstone, A. (2008). Attentional SNARC: There's something special about numbers (let us count the ways). *Cognition, 108,* 810–818.

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,* 95–110.

Fias, W., van Dijck, J. P., & Gevers, W. (2011). How is number associated with space? The role of working memory. In S. Dehaene & E. Brannon (Eds.), *Space, time and number in the brain: Searching for the foundations of mathematical thought (Attention and Performance XXIV* (pp. 133–148). San Diego: Academic Press.

Fischer, M. H., Castel, A. D., Dodd, M. D., & Pratt, J. (2003). Perceiving numbers causes spatial shifts of attention. *Nature Neuroscience, 6,* 555–556.

Fischer, M. H., & Fias, M. H. (2005). Spatial representation of numbers. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 43–54). New York: Psychology Press.

Fischer, M. H., Mills, R. A., & Shaki, S. (2010). How to cook a SNARC: Number placement in text rapidly changes spatial–numerical associations. *Brain and Cognition, 72,* 333–336.

Fried, P. J., Rushmore, R. J., Moss, M. B., Valero-Cabré, A., & Pascual-Leone, A. (2014). Causal evidence supporting functional dissociation of verbal and spatial working memory in the human dorsolateral prefrontal cortex. *European Journal of Neuroscience, 39,* 1973–1981.

Georges, C., Hoffmann, D., & Schiltz, C. (2014a). *Cognitive style influences number–space associations. Poster presented at the Meeting of the Belgian Association for Psychological Science*. Belgium: Leuven.

Georges, C., Hoffmann, D., & Schiltz, C. (2014b). *Different number-processing tasks entail qualitatively different SNARC effects. Poster presented at the Workshop on the Educational Neuroscience of Mathematics*. Germany: Tübingen.

Georges, C., Schiltz, C., & Hoffmann, D. (2015). Task instructions determine the visuospatial and verbal–spatial nature of number–space associations. *Quarterly Journal of Experimental Psychology, 68,* 1895–1909. doi:10.1080/17470218.2014.997764

Gevers, W., Reynvoet, B., & Fias, W. (2003). The mental representation of ordinal sequences is spatially organized. *Cognition, 87,* B87–B95.

Gevers, W., Santens, S., Dhooge, E., Chen, Q., Van den Bossche, L., Fias, W., & Verguts, T. (2010). Verbal-spatial and visuospatial coding of number–space interactions. *Journal of Experimental Psychology: General, 139,* 180–190.

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,* 32–44. doi:10.1037/0096-1523.32.1.32

Ginsburg, V., van Dijck, J. P., Previtali, P., Fias, W., & Gevers, W. (2014). The impact of verbal working memory on number–space associations. *Journal of Experimental Psychology Learning Memory and Cognition, 40,* 976–986.

Halligan, P. W., Fink, G. R., Marshall, J. C., & Vallar, G. (2003). Spatial cognition: Evidence from visual neglect. *Trends in Cognitive Sciences, 7,* 125–133. doi:10.1016/S1364-6613(03)00032-9

Herrera, A., Macizo, P., & Semenza, C. (2008). The role of working memory in the association between number magnitude and space. *Acta Psychologica, 128,* 225–237.

Hoffmann, D., Pigat, D., & Schiltz, C. (2014). The impact of inhibition capacities and age on number–space associations. *Cognitive Processing, 15,* 329–342.

Hubbard, E. M., Piazza, M., Pinel, P., & Dehaene, S. (2005). Interactions between number and space in parietal cortex. *Nature Reviews Neuroscience, 6,* 435–448. doi:10.1038/nrn1684

Hubbard, E. M., Piazza, M., Pinel, P., & Dehaene, S. (2009). Numerical and spatial intuitions: A role for posterior parietal cortex. In L. Tommasi, M. A. Peterson, & L. Nadel (Eds.), *Cognitive biology: Evolutionary and developmental perspectives on mind, brain, and behavior* (pp. 221–246). Cambridge: MIT Press.

Jaeggi, S. M., Buschkuehl, M., Perrig, W. J., & Meier, B. (2010). The concurrent validity of the *N*-back task as a working memory measure. *Memory, 18,* 394–412.

Lindemann, O., Abolafia, J. M., Pratt, J., & Bekkering, H. (2008). Coding strategies in number space: Memory requirements influence spatial–numerical associations. *Quarterly Journal of Experimental Psychology, 61,* 515–524.

Macmillan, N. A., & Creelman, C. D. (1990). Response bias: Characteristics of detection theory, threshold theory, and "nonparametric" indexes. *Psychological Bulletin, 107,* 401–413. doi:10.1037/0033-2909.107.3.401

McCabe, D. P., Roediger, H. L., III, McDaniel, M. A., Balota, D. A., & Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: Evidence for a common executive attention construct. *Neuropsychology, 24,* 222–243.

Nystrom, L. E., Braver, T. S., Sabb, F. W., Delgado, M. R., Noll, D. C., & Cohen, J. D. (2000). Working memory for letters, shapes, and locations: fMRI evidence against stimulus-based regional organization in human prefrontal cortex. *NeuroImage, 11,* 424–446.

Palmeri, T. J. (2002). Automaticity. In L. Nadel (Ed.), *Encyclopedia of cognitive science* (pp. 390–401). London: Nature Publishing.

Pfister, R., Schroeder, P. A., & Kunde, W. (2013). SNARC struggles: Instant control over spatial–numerical associations. *Journal of Experimental Psychology Learning Memory and Cognition, 39,* 1953–1958.

Postle, B. R., D'Esposito, M., & Corkin, S. (2005). Effects of verbal and nonverbal interference on spatial and object visual working memory. *Memory & Cognition, 33,* 203–212. doi:10.3758/BF03195309

Restle, F. (1970). Speed of adding and comparing numbers. *Journal of Experimental Psychology, 83,* 274–278. doi:10.1037/h0028573

Ristic, J., Wright, A., & Kingstone, A. (2006). The number line effect reflects top-down control. *Psychonomic Bulletin & Review, 13,* 862–868. doi:10.3758/BF03194010

Rugani, R., Vallortigara, G., Priftis, K., & Regolin, L. (2015). Number–space mapping in the newborn chick resembles humans' mental number line. *Science, 347,* 534–536.

van Dijck, J. P., & Fias, W. (2011). A working memory account for spatial–numerical associations. *Cognition, 119,* 114–119. doi:10.1016/j.cognition.2010.12.013

van Dijck, J. P., Gevers, W., & Fias, W. (2009). Numbers are associated with different types of spatial information depending on the task. *Cognition, 113,* 248–253. doi:10.1016/j.cognition.2009.08.005

Van Opstal, F., Santens, S., & Ansari, D. (2012). The numerate brain: Recent findings and theoretical reviews on the neurocognitive foundations of number processing. *Frontiers in Human Neuroscience, 6,* 201. doi:10.3389/fnhum.2012.00201

Viarouge, A., Hubbard, E. M., & McCandliss, B. D. (2014). The cognitive mechanisms of the SNARC effect: An individual differences approach. *PLoS ONE, 9,* e95756. doi:10.1371/journal.pone.0095756

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,* 489–525.

Zorzi, M., Priftis, K., & Umiltà, C. (2002). Brain damage: Neglect disrupts the mental number line. *Nature, 417,* 138–139. doi:10.1038/417138a