Is $26 + 26$ smaller than $24 + 28$? Estimating the approximate magnitude of repeated versus different numbers

Pom Charras · Garvin Brod · Juan Lupiáñez

Published online: 28 October 2011
© Psychonomic Society, Inc. 2011

Abstract It has recently been suggested that regardless of the dimension at hand (i.e., numerosity, length, time), similar operational mechanisms are involved in the comparison process based on approximate magnitude representation. One piece of evidence for this hypothesis lies in the presence of similar behavioral effects for any comparison (i.e., the distance effect). In the case of length comparison, the comparison process can be biased by summation toward either an underestimation or an overestimation: The sum of equal-size stimuli is underestimated, whereas the sum of different-size stimuli is overestimated. Relying on the hypothesis that similar operational mechanisms underlie the comparison process of any magnitude, we aim at extending these findings to another magnitude dimension. A number comparison task with digit numbers was used in the two experiments reported presently. The objective was to investigate whether summation also biases magnitude representation of numerical and symbolic information. The results provided evidence that the summation bias can also apply to numerical magnitude comparison, since the sum of repeated numbers ($26 + 26$) was underestimated whereas the sum of different numbers ($24 + 28$) was overestimated. We propose that these effects could be accounted for by a heuristic linking cognitive effort and magnitude estimation.

Keywords Approximate magnitude · Comparison process · Numerical cognition

In a numerical comparison task in which participants have to judge, for example, whether a numerosity is inferior or superior to 5, reaction times (RTs) are longer and accuracy is lower when responding to numerosity “4” than to “1” (see Fig. 1a). This effect, known as the *distance effect*, refers to the fact that the longer the time required to judge whether two magnitudes differ the shorter the distance between the two magnitudes (Dehaene, 1996; Moyer & Landauer, 1967). Interestingly, numerical judgments based on symbolic information (i.e., digits) show a similar distance effect, revealing that numerical comparisons based on symbolic and nonsymbolic information might involve similar mechanisms and rely on analogous magnitude representation (Butterworth, 1999; Dehaene, 1997; Piazza, Pinel, Le Bihan, & Dehaene, 2007; Shuman & Kanwisher, 2004). At a neuronal level, numerical comparisons with symbolic and nonsymbolic information entail the activation of the horizontal segment of the intraparietal sulcus (Piazza et al., 2007). On the basis of this behavioral and neuroimaging evidence, a common magnitude system for number and numerosity has been proposed (Dehaene, 1996; Gallistel & Gelman, 1992, 2000; Piazza, Izard, Pinel, Le Bihan, & Dehaene, 2004). This system is supposed to underlie the transformation of discrete events into analogous representation of numerical magnitude (for a review, see Brannon, 2006; see also Pinel, Dehaene, Riviere, & Le Bihan, 2001; Pinel, Piazza, Le Bihan, & Dehaene, 2004). Symbolic numerical information would be automatically transformed into analogous magnitude representation (Dehaene, Dehaene-Lambertz, & Cohen, 1998). Some authors have suggested that this magnitude system could be involved in dimensions other than number, such as space and time, for instance (see Hubbard,

P. Charras · G. Brod · J. Lupiáñez
University of Granada,
Granada, Spain

P. Charras (✉)
Departamento de Psicología Experimental
y de Fisiología del Comportamiento,
Campus de Cartuja, S/N.,
C.P. 18071, Granada, Spain
e-mail: pomcharras@ugr.es

Piazza, Pinel, & Dehaene, 2005; Kaufmann et al., 2005; Walsh, 2003). Although it still remains controversial to declare the existence of a shared magnitude representation system (Bueti & Walsh, 2009; Cappelletti, Freeman, & Cipolotti, 2009; Cohen Kadosh et al., 2005; Pinel et al., 2004), it seems commonly admitted that the mechanisms used for numerical magnitude comparison are similar to those used for physical comparison (Cantlon, Platt, & Brannon, 2009; Gallistel & Gelman, 1992, 2000). One piece of evidence accounting for this claim comes from the presence of the distance effect in any comparison task, regardless of the dimension (i.e., time, brightness, size, length). Indeed, in a line length comparison task, the more similar two stimuli appear, the slower responses are and the more errors are made (Fias, Lammertyn, Reynvoet, Dupont, & Orban, 2003; Fulbright, Manson, Skudlarski, Lacadie, & Gore, 2003; Johnson, 1939). Therefore, the distance effect seems to be an indicator of magnitude processing for numerical but also physical comparisons.

In the length comparison literature, it has recently been shown that line bisection modulates length estimation (see Charras & Lupiáñez, 2009, 2010; see also Künnapas, 1955; Wolfe, Maloney, & Tam, 2005). Length estimation of a line bisected into two equal-size parts is underestimated, whereas, in contrast, a line bisected into two different-sized parts is overestimated (Charras & Lupiáñez, 2009, 2010). In these tasks, a two-line configuration composed of a vertical and a horizontal line was briefly presented on every trial. The participant's task was to determine which of the two lines was the longest (the horizontal or the vertical one). The two lines always shared a junction, since they crossed each other. Therefore, in order to perform the comparison task, participants had somehow to sum up the two either equal

(symmetrically bisected lines) or different (asymmetrically bisected lines) parts of the line. Bisection into two equal versus different-sized parts clearly affected the point at which the lines were perceived as equal to each other. Bisection into two equal versus different-sized parts clearly affects the distribution of the distance effect. Indeed, by definition, the distance effect is an index of magnitude estimation, performance being worse at the closest distance. It predicts that if estimation is not biased (neither underestimation nor overestimation), performance is worse when the comparative stimuli are truly closer in magnitude to the standard stimulus, as is shown by the maximum error rate in Fig. 1a. However, if the comparative stimuli are underestimated (as with lines bisected into two equal-sized parts), the standard stimulus will appear close in magnitude to comparative stimuli that are superior in magnitude. Schematically, the error peak reflecting the point of perceived shorter distance (i.e., equality) will be displaced toward superior comparative stimuli (see Fig. 1b). In opposition, if the comparative stimuli are overestimated (as with lines bisected into two different-sized parts), the standard stimulus will appear close in magnitude to comparative stimuli that are inferior in magnitude, leading the error peak (i.e., the point of perceived shorter distance or equality) to be shifted toward inferior comparative stimuli (Fig. 1b).

On the basis of the fact that the distance effect is similar regardless of the dimension that is compared (from continuous and nonsymbolic to discrete and symbolic information), the objective of the present study was to test to what extent the summation biases observed in the magnitude representation of length (under and overestimations; Charras & Lupiáñez, 2009, 2010) could also be observed in the magnitude representation of numerical information. Therefore, we tested

Fig. 1 **a** Representation of the percentage of errors (%) in a number comparison task as a function of distance between standard (number 5 presently) and comparative stimulus (from 1 to 4 and 6 to 9) for digit numbers. **b** The graphs represent the percentage of errors (%) in a line length comparison task as a function of distance between standard and comparative stimuli for lines divided into two equal-sized parts and for lines divided into two different-sized parts. Note that in these two graphs, the four left-sided lines are physically shorter than the standard

stimulus (a non-bisected line), whereas the four right-sided lines are physically longer. The increase in errors between the closest inferior and superior distances, causing a deviation towards the right, refers to an underestimation of comparative stimuli as compared to the standard stimulus (a non-bisected line); in opposition, when errors decrease between the closest inferior and superior distances, producing a displacement towards the left, it mirrors an overestimation (see Charras & Lupiáñez, 2009, 2010)

whether in number processing the sum of two equal parts is biased toward an underestimation (i.e., $26 + 26$) and, conversely, whether the sum of two different parts is biased toward an overestimation (i.e., $24 + 28$).

Behavioral parallels between line and number processing have previously been demonstrated in the literature. For instance, left unilateral neglect, which is characterized by a rightward spatial bias, affects both line and number processing because these two dimensions show similar spatial organizations (Rossetti et al., 2004; Zorzi, Priftis, & Umiltà, 2002). In the present study, our objective was to highlight a new parallel between line and number processing that was not based on their similar spatial organizations but rather on their commonalities in coding and comparing magnitudes (i.e., the distance effect; for a review, see Cantlon et al., 2009; Cohen Kadosh, Lammertyn, & Izard, 2008). Henceforth, in the following experiments, we explored whether summation biases could be observed in a number comparison task—that is, a task based on symbolic and discrete information. We matched the comparison task used with lines (Charras & Lupiáñez, 2009, 2010) to those used in the present study. Participants performed a number comparison task with Arabic digits in which the comparative stimulus could be an addition. Participants were asked to rapidly add the two numbers and to judge whether the sum was inferior or superior to a standard stimulus. Two types of addition were presented: In the repeated number condition, stimuli were repeated number additions (RN, e.g., $26 + 26$), whereas in the different number condition, participants had to add two different numerals (DN, e.g., $24 + 28$). We aimed at testing whether the sum “ $26 + 26$ ” would be underestimated (i.e., judged as inferior to the standard stimulus 52), and the sum “ $24 + 28$ ” overestimated (i.e., judged as superior to the standard stimulus 52). The distance effect was analyzed and used as an index of the comparison efficiency.

If, as was expected, RN additions were underestimated in the present experiments, the error peak reflecting the perceived shorter distance or equality would have been displaced toward superior stimuli in the RN condition. Consequently, we should have observed better accuracy for inferior than for superior stimuli (see Fig. 1b). In contrast, if DN additions were overestimated, we should have observed better accuracy for superior stimuli than for inferior ones (see Fig. 1b).¹

¹ Since RN additions might be easier to solve than DN additions, we had in mind not to compare RN with DN additions, but rather to look at the planned comparisons between inferior and superior RN additions and between inferior and superior DN additions. However, it is important to underline that difficulty should not produce a distance effect displacement; rather, it should enhance the distance effect, in the sense that performance should be lower in general for DN than for RN additions.

Experiment 1

In the present experiment, two distances between standard and comparative stimuli were used (four units and two units) that were either inferior or superior ($n - 4$, $n - 2$, $n + 2$, $n + 4$). This manipulation aimed at highlighting a distance effect—that is, better performance for a four-unit distance than for a two-unit distance (either inferior or superior). And, as was explained above, our objective was to test whether performance drops for RN additions between inferior and superior distances ($n - 2 > n + 2$), in opposition to DN additions, in which we expected performance to improve between inferior and superior distances ($n - 2 < n + 2$).

Method

Participants Sixteen undergraduate students (one male) from the University of Granada participated voluntarily in this experiment in exchange for course credits. The mean age was 20 years old (range from 18 to 24). All participants reported to have normal or corrected-to-normal vision and all were naive as to the purpose of the experiment. The experiments reported presently were conducted in accordance with the ethical guidelines laid down by the Department of Experimental Psychology, University of Granada.

Apparatus and stimuli Participants entered a dimly lit room and sat down in front of a computer screen at a distance of about 60 cm. A PC computer running E-Prime software (Schneider, Eschman, & Zuccolotto, 2002) controlled the presentation of the stimuli, timing operations, and data collection. A fixation point, a silver colored dot (font, Courier New and bold; size, 18 pt), was displayed at the center of the screen and was removed during the intertrial delay. Digit numbers were presented using Arabic numerals and always appeared above fixation. They were written in Courier New font with a point size of 18. To differentiate standard and comparative stimuli, which were only separated by an empty black background display, the standard stimulus was written in green, and the comparative stimulus in white. The participants' task was to judge whether the comparative stimulus was inferior to the standard stimulus, in which case, they were to press the “c” key with their left hands, or superior, in which case they were to press the “b” key with their right hands (two forced choices response). Given that responses to numerical magnitude are spatially associated, being much easier to respond “inferior” with the left hand and superior with the right one (Snarc effect, see Dehaene, Bossini, & Giraux, 1993; Gevers, Lammertyn, Notebaert, Verguts, & Fias, 2006), response keys were not counterbalanced between participants.

The standard stimulus was a two-digit number whose value was 48, 50, 52 or 54, and varied from trial to trial within the same block.

The comparative stimulus could be a single number (SN) or an addition. More precisely, two types of additions were manipulated: repeated number additions (RN, i.e., $24 + 24$) and different number additions (DN, i.e., $22 + 26$). Independently of the comparative stimulus type (SN, RN, and DN), the standard-comparative stimuli distance (S–C) was of “-4 units,” “-2 units,” “+2 units,” or “+4 units”. For instance, using the standard stimulus 54, the comparative stimulus could assume one of the following values: 50, 52, 56, or 58 ($n - 4$, $n - 2$, $n + 2$ or $n + 4$). The set of stimuli used in the present study is detailed in Table 1.

In the DN condition, the numbers’ location in the additions was controlled, so that the smaller number was displayed on the left of the “plus” sign on half of the trials, and on the right on the other half. To equate the total of trials between the three types, SN and RN comparative stimuli were presented twice. As was mentioned above, the standard stimulus had four different values so as to prevent participants from learning the sum of additions and also to weaken target–response associations. For instance, the comparative stimulus “ $26 + 26$ ” triggered the response “superior” when compared to 50, but the response “inferior” when compared to 54. The standard stimulus manipulation could obviously not eliminate the advantage of practice, but at least it reduced it. The same trial was repeated only six times.

Procedure Every trial began with the appearance of a fixation point for a variable period of time (between 1,000 and 1,500 ms). Then, the standard stimulus was presented for 100 ms, after which it disappeared. After a variable duration of 900 or 1,100 ms (black screen), the comparative stimulus appeared for 500 ms. The fixation point was removed during response time. Participants had 2,500 ms to judge whether the comparative stimulus was inferior or superior to the previously presented standard stimulus. If no response was emitted during this time, a no-response feedback sentence (“too late”) was displayed on the screen. Feedback was provided for incorrect responses during practice; the word “NO!” written in red appeared on the screen when the response was incorrect.

Design The experiment had a 3 (type of comparative stimulus) $\times 4$ (S–C distance) design with the two factors manipulated within participants. Comparative stimuli could be an SN, an RN, or a DN. The S–C Distance varied between $n - 4$, $n - 2$, $n + 2$, $n + 4$.

Participants performed 20 practice trials before beginning the experimental block composed of 288 trials (24

Table 1 Stimuli used in Experiments 1 and 2

	Criterion	Distance	SN	RN	DN
EXPERIMENT 1	48	$n - 4$	44	$22 + 22$	$20 + 24$
		$n - 2$	46	$23 + 23$	$21 + 25$
		$n + 2$	50	$25 + 25$	$23 + 27$
		$n + 4$	52	$26 + 26$	$24 + 28$
	50	$n - 4$	46	$23 + 23$	$21 + 25$
		$n - 2$	48	$24 + 24$	$22 + 26$
		$n + 2$	52	$26 + 26$	$24 + 28$
		$n + 4$	54	$27 + 27$	$25 + 29$
	52	$n - 4$	48	$24 + 24$	$22 + 26$
		$n - 2$	50	$25 + 25$	$23 + 27$
		$n + 2$	54	$27 + 27$	$25 + 29$
		$n + 4$	56	$28 + 28$	$26 + 30$
	54	$n - 4$	50	$25 + 25$	$23 + 27$
		$n - 2$	52	$26 + 26$	$24 + 28$
		$n + 2$	56	$28 + 28$	$26 + 30$
		$n + 4$	58	$29 + 29$	$27 + 31$
EXPERIMENT 2	45	$n - 27$	18	$9 + 9$	$7 + 11$
		$n - 19$	26	$13 + 13$	$11 + 15$
		$n - 11$	34	$17 + 17$	$15 + 19$
		$n - 7$	38	$19 + 19$	$17 + 21$
		$n - 5$	40	$20 + 20$	$18 + 22$
		$n - 3$	42	$21 + 21$	$19 + 23$
		$n - 1$	44	$22 + 22$	$20 + 24$
		$n + 1$	46	$23 + 23$	$21 + 25$
		$n + 3$	48	$24 + 24$	$22 + 26$
		$n + 5$	50	$25 + 25$	$23 + 27$
		$n + 7$	52	$26 + 26$	$24 + 28$
		$n + 11$	56	$28 + 28$	$26 + 30$
		$n + 19$	64	$32 + 32$	$30 + 34$
		$n + 27$	72	$36 + 36$	$34 + 38$
	53	$n - 27$	26	$13 + 13$	$11 + 15$
		$n - 19$	34	$17 + 17$	$15 + 19$
		$n - 11$	42	$21 + 21$	$19 + 23$
		$n - 7$	46	$23 + 23$	$21 + 25$
		$n - 5$	48	$24 + 24$	$22 + 26$
		$n - 3$	50	$25 + 25$	$23 + 27$
		$n - 1$	52	$26 + 26$	$24 + 28$
		$n + 1$	54	$27 + 27$	$25 + 29$
		$n + 3$	56	$28 + 28$	$26 + 30$
		$n + 5$	58	$29 + 29$	$27 + 31$
		$n + 7$	60	$30 + 30$	$28 + 32$
		$n + 11$	64	$32 + 32$	$30 + 34$
		$n + 19$	72	$36 + 36$	$34 + 38$
		$n + 27$	80	$40 + 40$	$38 + 42$

SN Single Number, RN repeated numbers addition, DN different numbers addition

per experimental condition). Every 48 trial, a short break was allowed.

Results and discussion

Data from two participants were removed from the analysis since they did not reach 80% accuracy in the SN condition. We submitted accuracy (% of correct responses) and mean RTs to repeated measures ANOVAs, with the factors 3 (type of comparative stimulus) $\times$ 4 (S–C distance).

The analyses showed a main effect of type of comparative stimulus on accuracy and RTs, $F(2, 26) = 26.51$, $p < .001$, $\eta^2 = .67$, and $F(2, 26) = 89.90$, $p < .001$, $\eta^2 = .87$ (see Fig. 2), revealing that performance was better for SN than for RN, $F(1, 13) = 23.92$, $MSE = 47$, $p < .001$, for accuracy, and $F(1, 13) = 85.88$, $MSE = 67515$, $p < .001$, for RTs; and for RN as compared with DN, $F(1, 13) = 18.85$, $MSE = 111$, $p < .001$, for accuracy, and $F(1, 13) = 27.61$, $MSE = 22,061$, $p < .001$, for RTs. We also found the expected effect of S–C distance, $F(3, 39) = 12.29$, $p < .001$, $\eta^2 = .49$, for accuracy, and $F(3, 39) = 6.52$, $p = .001$, $\eta^2 = .33$, for RTs. To more specifically analyze the distance effect, the $n - 4$ and $n + 4$ conditions were grouped and compared to the $n - 2$ and $n + 2$ conditions. The planned comparisons showed that, as was

expected, performance was better for the four-unit than for the two-unit condition, indicating a clear distance effect, $F(1, 13) = 10.89$, $MSE = 63.89$; $p = .005$; $F(1, 13) = 28.62$, $MSE = 4159$, $p < .001$, respectively, for accuracy and RTs; see Table 2).

More importantly, the interactions between type of comparative stimulus and S–C distance were also significant, $F(6, 78) = 9.21$, $p < .001$, $\eta^2 = .41$, for accuracy, and $F(6, 78) = 5.65$, $p < .001$, $\eta^2 = .30$, for RTs. To test whether, as was predicted, performance was better for $n - 2$ than for $n + 2$ distance in the RN condition but was worse for $n - 2$ than for $n + 2$ in the DN condition, we conducted planned comparisons. We first observed that there was no difference between $n - 2$ and $n + 2$ in the SN condition (F s < 1 for accuracy and RTs). For RN additions, planned comparisons showed that performance to $n - 2$ and $n + 2$ was not statistically different, $F(1, 13) = 1.08$, $MSE = 35.89$, $p = .32$, and $F < 1$ for accuracy and RTs, whereas for DN additions, a superior distance advantage ($n - 2 < n + 2$) was observed on accuracy and RTs, $F(1, 13) = 20.01$, $MSE = 149.05$, $p < .001$, and $F(1, 13) = 15.59$, $MSE = 14,848$, $p = .002$ respectively).

Therefore these results are partly in line with our predictions, given that we did observe a DN overestimation, but no RN underestimation. A Stroop-like effect could have led to the lack of RN underestimation. Indeed, the standard stimulus was always a two-digit single number (“52”), whereas additions were composed of two numbers of two digits and a plus sign (“26 + 26”). Consequently, the physical size was much bigger for comparative than for standard stimuli. And, as was described in the literature, the discrepancy between physical size and numerical value can produce a Stroop-like effect (Pinel et al., 2004; Tzelgov, Meyer, & Henik, 1992). Thus, additions might be overestimated in the RN and DN conditions because stimuli constituting the comparative stimuli were physically larger than those constituting the standard stimulus. This Stroop-like effect might have accentuated the DN overestimation, but eliminated the RN underestimation. To test this assumption, we conducted Experiment 2 in which the standard stimulus was not visually presented.

Experiment 2

The objective of Experiment 2 was similar to that of Experiment 1, in the sense that we expected an underestimation of RN and an overestimation of DN. As in Experiment 1, a number comparison task was used in which comparative stimuli could be SN, RN, or DN. However in Experiment 2, the standard stimulus was not visually presented. Participants were instructed as to what was the standard stimulus at the beginning of the experiment (which was kept constant

Fig. 2 Representation of errors (%) and mean RTs (ms) for the three types of comparative stimulus as a function of the S–C distance in Experiment 1

Table 2 Means (*M*) and standard deviations (*SD*) of accuracy (in percentages) and reaction times (RTs, in milliseconds) as a function of type of comparative stimulus and S–C distance in Experiments 1 and 2

			<i>n</i> - 4		<i>n</i> - 2		<i>n</i> + 2		<i>n</i> + 4							
			<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>						
Experiment 1	Acc (%)	SN	98.57	1.98	96.78	3.42	97.42	3.71	97.64	4.86						
		RN	91.50	8.22	89.92	8.79	92.28	7.50	91.35	6.15						
		DN	83.64	14.08	66.85	17.93	87.5	11.93	92.50	8.18						
	RTs (ms)	SN	675	234	674	217	660	173	642	180						
		RN	1,044	247	1,155	245	1,162	218	1,110	216						
		DN	1,289	198	1,391	179	1,209	254	1,172	246						
			<i>n</i> - 27		<i>n</i> - 19		<i>n</i> - 11		<i>n</i> - 7		<i>n</i> - 5		<i>n</i> - 3		<i>n</i> - 1	
			<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Experiment 2	Acc (%)	98.9	3.61	97.7	4.83	96	5.8	91.9	11.9	94.2	7.35	87.7	11.87	84.2	19.72	
		98.2	5.93	97.7	4.83	93.5	10.9	97.1	6.63	94.5	8.94	87	18.75	74.5	19.42	
		98.2	5.93	96.2	8.32	94.6	8.85	85.7	16.93	78.7	16.48	67.8	20.76	67.1	24.81	
	RTs (ms)	602	124	611	140	641	148	691	177	673	148	713	151	772	219	
		699	128	714	134	778	214	844	231	886	223	871	198	1041	294	
		686	128	745	147	889	214	983	321	1,043	365	1,122	390	1,129	382	
			<i>n</i> + 1		<i>n</i> + 3		<i>n</i> + 5		<i>n</i> + 7		<i>n</i> + 11		<i>n</i> + 19		<i>n</i> + 27	
			<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Experiment 2	80.5	12.18	88.3	13.34	93	10.75	96	7.07	97.6	5.04	97.1	6.63	99.4	2.62		
	76.4	24.51	78.9	19.55	88.1	15.47	92.7	12.95	92.7	11.16	96	7.07	97.1	6.63		
	78.4	21.96	84.4	14.58	86.6	18.32	90.1	15.82	94.8	8.35	95.3	11.44	98.2	4.56		
	797	204	711	156	662	152	587	98	601	122	592	112	573	92		
	1061	332	1046	321	892	241	758	168	783	175	741	169	696	153		
	1082	345	1068	342	974	308	891	262	881	240	790	220	766	188		

throughout the experiment). Two standard stimuli were used (45 vs. 53, manipulated between subjects; see Table 1). Note that most of the superior additions for the standard stimulus 45 became the inferior additions for the standard stimulus 53. Since our objective was to compare inferior to superior additions separately for RN and DN additions, this manipulation canceled out possible issues about the digit numbers that were used (i.e., compatibility effect; Nuerk, Weger, & Willmes, 2001, 2004). And, finally, to obtain a more pronounced distance effect, we used a larger scale of stimuli (from *n* - 27 to *n* + 27 in 14 steps).

Method

Participants Twenty-one undergraduate students from the University of Granada participated voluntarily in this exper-

iment in exchange for course credits. The participants' age ranged from 19 to 41 years (mean age 21).

Apparatus and stimuli A fixation point, a silver colored dot (font, Courier New and bold; size, 25), was displayed at the center of the screen. Digit numbers were Arabic numerals (font, Courier New and bold; size, 15) that always appeared above fixation. The participants' task was to judge whether the comparative stimulus was either inferior to the standard stimulus, in which case they were to press the "v" key with their right index fingers, or superior, in which case they were to press the "b" key with their right middle fingers (Spanish keyboard, in which the "v" is to the left of the "b"). As in the previous experiment, comparative stimuli could be either a SN, or a sum of RN or DN. However, in this experiment, SN stimuli were presented separately in the starting and final blocks. RN and DN stimuli were mixed within the same block and were presented between the

starting and final blocks. Independently of the type (SN, RN, DN), the S–C distance was of “27 units,” “19 units,” “11 units,” “7 units,” “5 units,” “3 units,” or “1 unit,” either inferior or superior to the standard stimulus. To equate the material between RN and DN conditions, two standard stimuli were used (either 45 or 53) but were counter-balanced between participants (see Table 1)—that is, each participant received only one standard stimulus.

Procedure Every trial began with the appearance of a fixation point for a period of 1,000 ms and was followed by the comparative stimulus for 500 ms. During the response time delay, the fixation point was removed. Participants had to judge whether the comparative stimulus was inferior or superior to the standard stimulus (either 45 or 53) within a time window of 3 s.

Design The experiment had a 3 (type of comparative stimulus) $\times$ 14 (S–C distance) design with the two factors manipulated within participants. Participants performed 20 practice trials before beginning the experimental block composed of 336 trials (eight repetitions of each trial).²

Results and discussion

RTs faster than 300 ms and longer than 2,700 ms were removed from data analysis (1.45% of the trials). We then conducted repeated measures ANOVAs for accuracy and RTs with the factors type (SN, RN, DN) and S–C distance (14 levels).³ The main effects of type and S–C distance were significant, $F(2, 40) = 17.63$, $p < .001$, $n^2 = .47$, and $F(13, 260) = 21.10$, $p < .001$, $n^2 = .51$, respectively, for accuracy; $F(2, 40) = 33.70$, $p < .001$, $n^2 = .63$, and $F(13, 260) = 36.22$, $p < .001$, $n^2 = .64$, for RTs. The type effect showed that SN were more easily solved than RN, $F(1, 20) = 5.53$, $MSE = .02$, $p = .02$, for accuracy, and $F(1, 20) = 41.26$, $MSE = 121326$, $p < .001$, for RTs, which were more easily solved than DN, $F(1, 20) = 13.73$, $MSE = .01$, $p = .001$,

for accuracy, and $F(1, 20) = 14.76$, $MSE = 78,232$, $p = .001$ for RTs. The main effect of S–C distance refers to the distance effect: Performance is better for the farthest than for the closest S–C distances. To explicitly test the effect of distance on performance, we performed planned comparisons for accuracy and RTs in which the three farthest conditions (n27, n19, n11; for both inferior and superior) were grouped and compared to the three closest conditions (n5, n3, n1; for both inferior and superior). The planned comparisons confirmed that performance is better for the three farthest than for the three closest S–C distances, $F(1, 20) = 99.88$, $MSE = .03$, $p < .001$, for accuracy, and $F(1, 20) = 84.93$, $MSE = 97,046$, $p < .001$, for RTs.

Importantly, the interactions between the two factors were also highly significant, $F(26, 520) = 2.91$, $p < .001$, $n^2 = .13$, for accuracy, and $F(26, 520) = 3.49$, $p < .001$, $n^2 = .15$, for RTs. We performed planned comparisons for the two closest inferior conditions ($n - 3$, $n - 1$) versus the two closest superior conditions ($n + 1$, $n + 3$) separately for SN, RN, and DN. No difference was observed for the SN condition between inferior and superior distances, $F_s < 1$, *n.s.* for accuracy and RTs). In the RN condition, although accuracy was statistically similar for inferior ($n - 3$ and $n - 1$) and superior ($n + 3$ and $n + 1$) distances ($F_s < 1$), a significant effect was observed on RTs, $F(1, 20) = 4.20$, $MSE = 47,238$, $p = .05$. Responses were faster for inferior than for superior distances. In contrast, for the DN condition, accuracy was greater for superior distances than for inferior ones, $F(1, 20) = 8.05$, $MSE = .05$, $p = .01$, whereas there was no significant difference in RTs, $F(1, 20) = 1.11$, $MSE = 49,146$, $p = .30$, *n.s.*

Although these results suggest a RN underestimation and a DN overestimation, we further analyzed the data by computing the point of subjective equality (PSE⁴) for every type of comparative stimulus. We then conducted a repeated measures ANOVA with the factor type (SN, RN, DN). A main effect of type was found, $F(2, 40) = 6.79$, $p = .003$, $n^2 = .25$ (see Fig. 3). Planned comparisons showed that the PSE of DN additions was statistically different from that of RN additions, $F(1, 20) = 15.59$, $MSE = 3.76$, $p < .001$, confirming that DN stimuli were overestimated as

² Note that this experiment initially included two sessions, one involving a comparison task and the other one an estimation task. The estimation task session is not reported in the present article. Task order was counterbalanced between participants.

³ Even though two criterions were used—45 and 53—we grouped the data of these two groups to control for possible issues with the material. The use of two criterions enabled us to reduce discrepancies between RN and DN problems, or between inferior or superior problems. Note, however, that we tested for a possible effect of the criterion on performance, and we found no effect on accuracy ($F < 1$) and only a tendencial effect on RTs, indicating that responses to the criterion 53 were slower than those to the criterion 45.

⁴ PSE is the estimated value at which the comparison number is judged as equal to the standard. To compute the PSE for each participant, we collected the percentage of “superior” response in every S–C distance conditions. In the $n - 27$ condition, we obtained a “superior” response percentage near 0%, whereas in the $n + 27$ condition, the “superior” response percentage was about 100%. Then, we used an inverse normal cumulative distribution to transform the “superior” response percentage into a score. These scores for every S–C distance condition made it possible to calculate for each participant the slope, by dividing the score by the S–C distance, and the intercept. Once the slope and the intercept were determined, the PSE could be calculated by the following formula: $PSE = (0 - \text{intercept}) / \text{slope}$. Note that 0 in z score is equal to 50%.

Fig. 3 Representation of errors (%) and mean RTs (in milliseconds) for the three types of comparative stimulus as a function of the S–C distance in Experiment 2

compared with RN stimuli. Single mean t tests indicated that the PSE in the SN condition was not different from 0, $t(20) = -0.97$, $p = .35$, indicating that SN stimuli were well estimated. However, the PSE of DN additions (-1.30) tended to be inferior to 0, revealing an overestimation, $t(20) = -1.96$, $p = .06$, whereas the PSE of RN additions ($+1.06$) was superior to 0, suggesting an underestimation, $t(20) = 2.11$, $p = .05$.

General discussion

Nowadays, converging evidence has led us to think that humans might have a general system to represent and compare magnitudes approximately regardless of the dimension (for a review, see Cantlon et al., 2009). In everyday life, we constantly have to estimate distance, time, and numerosity to make decisions according to approximate magnitude representations—for example, “Which is the shortest way to get to my office?” “Do I have time to buy the newspaper before my train leaves?” It has been argued that the comparison process of any continuous variables shows common signatures, such as the distance and size effects (for a review, see Buetti & Walsh, 2009; Cantlon et al., 2009; Cohen Kadosh et al., 2008; Hubbard et al., 2005; Walsh, 2003).

In the present investigation, we wanted to explore how summation biases magnitude estimation as indexed by the distribution of the distance effect. Recent studies have reported that line length comparisons are biased toward an underestimation when summing two identical-sized parts and toward an overestimation when the two parts are different in size (Charras & Lupiáñez, 2009, 2010). In the present study, we further explored whether summation also affects the representation of approximate magnitude when information is conveyed by symbolic and discrete information—that is, digit numbers. Instead of using a continuous and nonsymbolic dimension, a number comparison task was used in which participants had to rapidly judge whether the sum of two large numbers under Arabic digit notation was inferior or superior to a standard stimulus. Large numbers were used in order to avoid retrieving the sum from memory and hence encouraging the representation of approximate magnitude.

First, the results of the two experiments have clearly provided evidence that RN additions are more easily and accurately solved than DN additions, as indexed by a larger percentage of errors and longer RTs. In the literature of numerical cognition, it is well established that performance at adding two single digits is characterized by a reliable advantage in both latency and accuracy for two repeated-number additions as compared with two different-number additions (i.e., $8 + 8$, called a “tie problem,” vs. $9 + 7$, called a “nontie problem,” see Ashcraft, 1995; Groen & Parkman, 1972; Lefevre, Sadesky, & Bisanz, 1996). The fact that repeated and different number additions are not uniform in difficulty is referred to as the *tie effect* (Campbell, 1999; Campbell & Graham, 1985; Campbell & Gunter, 2002; Clapp, 1924; Lefevre & Liu, 1997). So far, to our knowledge, the tie effect was considered to be specific of small numbers. Indeed, this advantage was supposed to rely on direct memory retrieval (Lefevre, Shanahan, & DeStefano, 2004). However, it seems hard to believe that in the present experiments sums were retrieved from memory.⁵ Another account for the tie effect stresses encoding facilitation (Blankenberger, 2001). And, by definition, it assumes that problem size should not affect the tie effect. The general advantage of RN as compared to DN, observed in the present study, directly questions the nature of the tie effect, since large-number additions appear to be also influenced by

⁵ Some of the additions used here were very easy to solve (i.e., $20 + 20$), and their sums might indeed have been retrieved from memory. But, in general, additions were much more difficult and time pressure of the experiments did not allow slow strategies. Moreover, in Experiment 2, an easy addition as $20 + 20$ for RN and $26 + 30$ for DN was counterbalanced by a difficult addition in the other standard stimulus condition (i.e., $24 + 24$ for RN, and $22 + 26$ for DN). Even though it is possible that the two ways of processing have been involved in parallel, the presence of a general distance effect argues for a processing mainly based on approximate magnitude estimation rather than on retrieval of memory facts.

number repetition. But, besides the general advantage for RN, our results show that the type of comparative stimulus (RN vs. DN) highly affects the distribution of the distance effect (see Fig. 3). Indeed, summation biased the global representation of approximate magnitude toward an underestimation when they were identical, and toward an overestimation when the numbers to be added were different.

In this context, it is worth noting that recently McCrink, Dehaene, and Dehaene (2007) reported that computation of additions yields to an overestimation, in opposition to subtractions that produce an underestimation; this is referred to as the *operational momentum effect* (OM). On the basis of evidence supporting close associations between space and number, and more specifically on the spatially organized mental representation of numerical magnitude, these authors have suggested that addition processing is biased toward the right on the mental number line, mirroring an overestimation, whereas subtraction process is biased toward the left, congruent with an underestimation (see also Knops, Viarouge, & Dehaene, 2009; McCrink & Wynn, 2009; Pinhas & Fischer, 2008). This account implies that the OM effect is determined exclusively by the operator. However, our results provide evidence that the nature of the operands might also be critical (see also Cordes, Gallistel, Gelman, & Latham, 2007). One could consider that the under- and overestimation effects reported presently and the OM effect are additive, both acting on approximate calculation during additions. Whereas the nature of operation (addition, in contrast with subtraction) induces an overestimation bias, the nature of the operands to be added (RN vs. DN) could induce either an overestimation (in the case of DN) or an underestimation (in the case of RN). If operands (RN vs. DN) and operator (addition) biases acted independently on performance, we should get an enhanced overestimation of DN, since both biases would act in the same direction, and a reduced or absent underestimation of RN, since the underestimation induced by the operands would be compensated by the overestimation induced by the operator. The overall pattern of results observed in our two experiments seems to accommodate this hypothesis, although further investigation will be needed to specifically test this proposal.

At this point, one might wonder why the comparison process is biased toward an underestimation for sums of repeated stimuli and toward an overestimation for sums of different stimuli. In the following section, we have attempted to provide a tentative account. On the basis of the fact that the two stimuli in the repeated-number condition are by definition identical, we hypothesize that they might be processed as a single (repeated) stimulus. We propose that their processing requires less computation/cognitive effort, since there is no need to build a separate representation for each addend, as compared with DN additions, for which each operand needs to

be represented. A heuristic linking cognitive effort and magnitude estimation could then account for the RN and DN effects. The easiness of representing RN additions could be due to the fact that they can be represented as a single repeated stimulus (i.e., plus a duplication sign). Although further research is needed to better understand the effect of stimulus repetition on magnitude estimation, it seems already uncontroversial that repeated stimuli are not processed in the same way as different stimuli. For instance, Xu and Chun (Xu, 2007, 2009; Xu & Chun, 2007, 2009), in accordance with the Kahneman's object file theory (Kahneman, Treisman, & Gibbs, 1992), have recently argued that object perception runs through two phases: object individuation and object identification. Object individuation, which leads to "open" an object file, is limited to four events and is implemented in the inferior intraparietal sulcus (IPS); whereas object identification relies on the superior IPS. As a consequence, object perception should be impaired after bilateral parietal lesion, consistent with what can be observed in patients suffering from Balint's syndrome. Xu (2009) also proposed that stimulus repetition has a special status, and she provided evidence that stimuli that were repeated on the same visual display were considered as distinct objects in the object individuation phase, but as a single object in the object identification phase. We propose to extend these notions to the case of magnitude estimation. If a stimulus is repeated, then the ease in processing identical stimuli or the cognitive gain in representing these objects may produce a magnitude underestimation as compared with different objects.

Obviously further research is needed to claim that the over- and underestimation effects are a new common signature that affects the comparison process. Nevertheless, the present article is, to our knowledge, the first piece of evidence supporting the idea of a general bias due to stimulus summation in the representation of approximate magnitude. It seems now necessary to gather more evidence supporting the modulation of the over- and underestimation effects over the magnitude comparison process and to elaborate a robust theoretical account supporting these effects.

Author Note This research was supported by the Spanish "Ministerio de Educación y Ciencia" (by predoctoral Grant AP-2006-3911 to P.C., and research Grant PSI2008-03595PSIC to G.B.). Please direct correspondence concerning this article to: P. Charas, Departamento de Psicología Experimental y Fisiología del Comportamiento, Facultad de Psicología, Universidad de Granada, Campus Universitario de Cartuja S/N, 18071, Granada, Spain (e-mail: pomcharas@ugr.es).

References

- Ashcraft, M. H. (1995). Cognitive psychology and simple arithmetic: A review and summary of new directions. *Mathematical Cognition*, 1, 3–34.

- Blankenberger, S. (2001). The arithmetic tie effect is mainly encoding-based. *Cognition*, 82, 15–24.
- Brannon, E. M. (2006). The representation of numerical magnitude. *Current Opinion in Neurobiology*, 16, 222–229.
- Bueti, D., & Walsh, V. (2009). The parietal cortex and the representation of time, space, number and other magnitudes. *Philosophical Transactions of the Royal Society B*, 364, 1831–1840.
- Butterworth, B. (1999). *The Mathematical Brain*. London: Macmillan.
- Campbell, J. I. (1999). Division by multiplication. *Memory & Cognition*, 27, 791–802.
- Campbell, J. I., & Graham, D. J. (1985). Mental multiplication skill: Structure, process, and acquisition. *Canadian Journal of Psychology*, 39, 338–366.
- Campbell, J. I., & Gunter, R. (2002). Calculation, culture, and the repeated operand effect. *Cognition*, 86, 71–96.
- Cantlon, J. F., Platt, M. L., & Brannon, E. M. (2009). Beyond the number domain. *Trends in Cognitive Sciences*, 13, 83–91.
- Cappelletti, M., Freeman, E. D., & Cipolotti, L. (2009). Dissociations and interactions between time, numerosity and space processing. *Neuropsychologia*, 47, 2732–2748.
- Charras, P., & Lupiáñez, J. (2009). The relevance of symmetry in line length perception. *Perception*, 38, 1428–1438.
- Charras, P., & Lupiáñez, J. (2010). Length perception of horizontal and vertical bisected lines. *Psychological Research*, 74, 196–206.
- Clapp, F. L. (1924). *The number combinations: Their relative difficulty and frequency of their appearance in textbooks*. Madison: Bureau of Educational Research.
- Cohen Kadosh, R., Henik, A., Rubinsten, O., Mohr, H., Dori, H., van de Ven, V., ... Linden, D. E. (2005). Are numbers special?: The comparison systems of the human brain investigated by fMRI. *Neuropsychologia*, 43, 1238–1248.
- Cohen Kadosh, R., Lammertyn, J., & Izard, V. (2008). Are numbers special? An overview of chronometric, neuroimaging, developmental and comparative studies of magnitude representation. *Progress in Neurobiology*, 84, 132–147.
- Cordes, S., Gallistel, C. R., Gelman, R., & Latham, P. (2007). Nonverbal arithmetic in humans: Light from noise. *Perception & Psychophysics*, 69, 1185–1203.
- Dehaene, S. (1996). The organization of brain activations in number comparison: Event-related potentials and the additive-factors method. *Journal of Cognitive Neuroscience*, 8, 47–68.
- Dehaene, S. (1997). *The number sense: How the mind creates mathematics*. New York: Oxford University Press.
- Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and numerical magnitude. *Journal of Experimental Psychology: General*, 122, 371–396.
- Dehaene, S., Dehaene-Lambertz, G., & Cohen, L. (1998). Abstract representations of numbers in the animal and human brain. *Trends in Neuroscience*, 21, 355–361.
- Fias, W., Lanunertyn, J., Reynvoet, B., Dupont, P., & Orban, G. A. (2003). Parietal representation of symbolic and nonsymbolic magnitude. *Journal of Cognitive Neuroscience*, 15, 47–56.
- Fulbright, R. K., Manson, S. C., Skudlarski, P., Lacadie, C. M., & Gore, J. C. (2003). Quantity determination and the distance effect with letters, numbers, and shapes: A functional MR imaging study of number processing. *American Journal of Neuroradiology*, 24, 193–200.
- Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition*, 44, 43–74.
- Gallistel, C. R., & Gelman, R. (2000). Non-verbal numerical cognition: From reals to integers. *Trends in Cognitive Sciences*, 4, 59–65.
- Gevers, W., Lammertyn, J., Notebaert, W., Verguts, T., & Fias, W. (2006). Automatic response activation of implicit spatial information: Evidence from the SNARC effect. *Acta Psychologica*, 122, 221–233.
- Groen, G. J., & Parkman, J. M. (1972). A chronometric analysis of simple arithmetic. *Psychological Review*, 79, 329–343.
- Hubbard, E. M., Piazza, M., Pinel, P., & Dehaene, S. (2005). Interactions between number and space in parietal cortex. *Nature Reviews Neuroscience*, 6, 435–448.
- Johnson, D. M. (1939). Confidence and speed in the two-category judgment. *Archives of Psychology*, 241, 1–52.
- Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, 24, 175–219.
- Kaufmann, L., Koppelstaetter, F., Delazer, M., Siedentopf, C., Rhomberg, P., Golaszewski, S., ... Ischebeck, A. (2005). Neural correlates of distance and congruity effects in a numerical Stroop task: an event-related fMRI study. *NeuroImage*, 25, 888–898.
- Knops, A., Viarouge, A., & Dehaene, S. (2009). Dynamic representations underlying symbolic and nonsymbolic calculation: Evidence from the operational momentum effect. *Attention, Perception, & Psychophysics*, 71, 803–821.
- Künnapas, T. M. (1955). An analysis of the "vertical horizontal illusion". *Journal of Experimental Psychology*, 49, 134–140.
- LeFevre, J., Sadesky, G. S., & Bisanz, J. (1996). Selection of procedures in mental addition: Reassessing the problem size effect in adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 216–230.
- Lefevre, J. A., & Liu, J. (1997). The role of experience in numerical skill: Multiplication performance in adults from Canada and China. *Mathematical Cognition*, 3, 31–62.
- Lefevre, J. A., Shanahan, T., & DeStefano, D. (2004). The tie effect in simple arithmetic: An access-based account. *Memory & Cognition*, 32, 1019–1031.
- McCrink, K., Dehaene, S., & Dehaene-Lambertz, G. (2007). Moving along the number line: Operational momentum in nonsymbolic arithmetic. *Perception & Psychophysics*, 69(8), 1324–1333.
- McCrink, K., & Wynn, K. (2009). Operational momentum in large-number addition and subtraction by 9-month-olds. *Journal of Experimental Child Psychology*, 103, 400–408.
- Moyer, R. S., & Landauer, T. K. (1967). The time required for judgments of numerical inequality. *Nature*, 215, 1519–1520.
- Nuerk, H. C., Weger, U., & Willmes, K. (2001). Decade breaks in the mental number line? Putting the tens and units back in different bins. *Cognition*, 82, 25–33.
- Nuerk, H. C., Weger, U., & Willmes, K. (2004). On the perceptual generality of the unit-decade compatibility effect. *Experimental Psychology*, 51, 72–79.
- Piazza, M., Izard, V., Pinel, P., Le Bihan, D., & Dehaene, S. (2004). Tuning curves for approximate numerosity in the human intraparietal sulcus. *Neuron*, 44, 547–555.
- Piazza, M., Pinel, P., Le Bihan, D., & Dehaene, S. (2007). A magnitude code common to numerosities and number symbols in human intraparietal cortex. *Neuron*, 53, 293–305.
- Pinel, P., Dehaene, S., Riviere, D., & Le Bihan, D. (2001). Modulation of parietal activation by semantic distance in a number comparison task. *NeuroImage*, 14, 1013–1026.
- Pinel, P., Piazza, M., Le Bihan, D., & Dehaene, S. (2004). Distributed and overlapping representations of number, size, and luminance in parietal cortex during comparative judgements. *Neuron*, 41, 983–993.
- Pinhas, M., & Fischer, M. H. (2008). Mental movements without magnitude? A study of spatial biases in symbolic arithmetic. *Cognition*, 109, 408–415.
- Rossetti, Y., Jacquin-Courtois, S., Rode, G., Ota, H., Michel, C., & Boisson, D. (2004). Does action make the link between number and space representation? Visuo-manual adaptation improves number bisection in unilateral neglect. *Psychological Science*, 15, 426–430.

- Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-prime user's guide*. Pittsburg: Psychology Software Tools, Inc.
- Shuman, M., & Kanwisher, N. (2004). Numerical magnitude in the human parietal lobe; tests of representational generality and domain specificity. *Neuron*, 44, 557–569.
- Tzelgov, J., Meyer, J., & Henik, A. (1992). Automatic and intentional processing of numerical information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 166–179.
- Walsh, V. (2003). A theory of magnitude: Common cortical metrics of time space and quantity. *Trends in Cognitive Sciences*, 7, 483–488.
- Wolfe, U., Maloney, L. T., & Tam, M. (2005). Distortions of perceived length in the frontoparallel plane: Tests of perspective theories. *Perception & Psychophysics*, 67, 967–979.
- Xu, Y. (2007). The role of the superior intraparietal sulcus in supporting visual short-term memory for multifeature objects. *Journal of Neuroscience*, 24, 11676–11686.
- Xu, Y. (2009). Distinctive neural mechanisms supporting visual object individuation and identification. *Journal of Cognitive Neuroscience*, 21, 511–518.
- Xu, Y., & Chun, M. M. (2007). Visual grouping in human parietal cortex. *Proceedings of the National Academy of Sciences*, 104, 18766–18771.
- Xu, Y., & Chun, M. M. (2009). Selecting and perceiving multiple visual objects. *Trends in Cognitive Sciences*, 13, 167–174.
- Zorzi, M., Priftis, K., & Umiltà, C. (2002). Neglect disrupts the mental number line. *Nature*, 417, 138–139.