Cognition 113 (2009) 189–204

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Beyond quantity: Individual differences in working memory and the ordinal understanding of numerical symbols

Ian M. Lyons, Sian L. Beilock *

Department of Psychology, 5848 South University Avenue, University of Chicago, Chicago, Illinois 60637, United States

ARTICLE INFO

Article history:
Received 19 November 2008
Revised 13 July 2009
Accepted 1 August 2009

Keywords:
Numerical cognition
Symbol learning
Working memory
Individual differences
Ordinal processing
Numerical symbols

ABSTRACT

In two different contexts, we examined the hypothesis that individual differences in working memory (WM) capacity are related to the tendency to infer complex, ordinal relationships *between* numerical symbols. In Experiment 1, we assessed whether this tendency arises in a learning context that involves mapping novel symbols to quantities by training adult participants to associate dot-quantities with novel symbols, the overall relative order of which had to be inferred. Performance was best for participants who were higher in WM capacity (HWMs). HWMs also learned ordinal information about the symbols that lower WM individuals (LWMs) did not. In Experiment 2, we examined whether WM relates to performance when participants are explicitly instructed to make numerical order judgments about highly enculturated numerical symbols by having participants indicate whether sets of three Arabic numerals were in increasing order. All participants responded faster when sequential sets (3–4–5) were in order than when they were not. However, only HWMs responded faster when non-sequential, patterned sets (1–3–5) were in order, suggesting they were accessing ordinal associations that LWMs were not. Taken together, these experiments indicate that WM capacity plays a key role in extending symbolic number representations beyond their quantity referents to include symbol–symbol ordinal associations, both in a learning context and in terms of explicitly accessing ordinal relationships in highly enculturated stimuli.

1. Introduction

Numerical representations can be thought of as having two interrelated aspects: a sense of quantity (about six apples; Dehaene, 1997) and relative order (six comes before seven and after five; Jacob & Nieder, 2008; Tzelgov & Ganor-Stern, 2004). Although these aspects are certainly intertwined (Fias, Lammertyn, Caessens, & Orban, 2007), the majority of previous research on symbolic number processing has focused on the quantitative aspect of symbolic numbers. However, mounting behavioral evidence indicates that relative order information plays a particularly important role in *symbolic* numerical representation (Roggeman, Verguts, & Fias, 2006; Turconi, Campbell, & Seron, 2006).

In the current work, we turn our attention to people's ordinal understanding of numerical symbols. In particular, we examined how a general cognitive capacity – namely, working memory (WM) – relates to individual differences in numerical understanding via variation in how relative order information is acquired and accessed. Because complex mathematical processes rely strongly on symbolic representations as well as WM (e.g., Ashcraft & Krause, 2007), understanding the interaction between ordinal processing and WM may provide important insight into why some excel in mathematics while others struggle.

1.1. Working memory and symbolic numerical representation

What are the cognitive and neural operations that support the learning and access of symbolic numbers? A prominent neuronal model suggests that Arabic numerals are understood by mapping them on to pre-existing

* Corresponding author. Tel.: +1 773 834 3713.
E-mail address: beilock@uchicago.edu (S.L. Beilock).

0010-0277/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2009.08.003

representations of approximate quantity (Verguts & Fias, 2004; see also Ansari, 2008; Dehaene & Changeaux, 1993; Nieder & Merten, 2007). Neuronal evidence provides some support for this view: Diester and Nieder (2007) trained monkeys to associate Arabic numerals with corresponding quantities of dots and used single neuron recordings to locate cells involved in this process. The authors found individual neurons that responded maximally to a specific quantity, whether that quantity was presented as dots or the corresponding Arabic numeral. Intriguingly, the majority of these 'association neurons' (neurons tuned to the same quantity regardless of format) were found in monkey prefrontal cortex.

Prefrontal cortices in monkeys and humans routinely activate in tasks with high WM demands (Kane & Engle, 2002; Wager & Smith, 2003) and individuals characteristically higher in constructs closely related to working memory (e.g., general fluid intelligence, or gF) activate these areas to a greater extent than lower gF individuals when performing cognitively demanding tasks (Gray, Chabris, & Braver, 2003). This suggests that higher-level cognitive capacities, such as WM, may play an important role in mapping numerical quantities onto abstract visual symbols. For this reason, it seems plausible that the symbol-mapping process and the representation of overlearned numerical symbols may vary as a function of one's WM capacity.

Working memory can be thought of as a short-term memory system involved in the control, regulation, and active maintenance of a limited amount of information with immediate relevance to the task at hand (Miyake & Shah, 1999). Previous work has shown that WM is critical to a wide range of mathematical tasks – such as carrying operations (Imbo, Vandierendonck, & Vergauwe, 2007), counting speed (Tuholski, Engle, & Baylis, 2001), mental arithmetic (LeFevre, DeStefano, Coleman, & Shanahan, 2004; Logie, Gilhooly, & Wynn, 1994), and even the selection of strategies used in completing complex arithmetic problems (Beilock & Decaro, 2007; for a review, see also Ashcraft & Krause, 2007).

In the current study, we suggest that, at the level of individual differences, *ordinal* understanding of symbolic numbers is strongly related to WM capacity. We examined this hypothesis in two different contexts. In Experiment 1, we assessed whether individual differences in WM capacity are related to the tendency to infer complex, ordinal relationships in a numerical symbol-learning task. In Experiment 2, we take our work a step further and ask whether the relation between WM and an individual's emphasis on ordinal associations generalizes to a task requiring explicit relative order judgments about highly enculturated numerical symbols: Arabic numerals.

2. Experiment 1

The ability to map numerical quantities onto symbols (e.g., Arabic numerals) and to subsequently retrieve this quantity information when presented with these numerical symbols at a later time is an important mathematical skill. Without this type of numerical symbol-mapping, important mathematical activities are outside one's reach. For example, cultures without distinct symbols or words for quantities exceeding the first few integers are only able to perform exact calculations by approximating the correct answer (Pica, Lemer, Izard, & Dehaene, 2004). Because numerical symbol-mapping carries implications for complex math skills like counting and arithmetic (Gallistel & Gelman, 1992), an understanding of how people vary in their acquisition and representation of numerical symbols is important for the development of educational practices and teaching tools that ensure high-level number understanding in all individuals.

Thus, in Experiment 1, we addressed two interrelated questions. First, we asked whether emphasis on order information when acquiring a set of novel numerical symbols is related to how well participants learn to use these novel symbols in numerical tasks. Given that relative order information plays an important role in symbolic number representation (Roggeman et al., 2006; Turconi et al., 2006), this seems plausible. Second, we asked whether individual differences in WM capacity dictate who emphasizes order in the first place. Specifically, we hypothesized that individuals higher in WM (HWMs) would be more likely to infer ordinal relations than lower working memory individuals (LWMs) (see reasons for this assumption below), and that this would lead to better overall performance for HWMs when acquiring a set of novel numerical symbols. To address these issues, we had participants learn to associate non-symbolic numerical quantities (in the form of dot-arrays presented too rapidly to count) with six novel visual shapes. Participants then performed a series of numerical tasks designed to shed light on their grasp of the symbols' numerical content.

How might individual differences in WM be related to the symbol-mapping process? On the one hand, one might assume that this process is simple enough to be constrained to brain regions canonically associated with numerical processes (Dehaene, Piazza, Pinel, & Cohen, 2003) and thus would not vary as a function of individual differences in more domain-general capacities. One the other hand, as mentioned above, Diester and Nieder (2007) found evidence of numerical tuning in neurons in monkey prefrontal cortex. In addition, Lyons and Ansari (2009) recently showed activation in a left prefrontal region also implicated in WM-dependent tasks when people learned to map quantities onto abstract symbols. Thus, an alternative view is that individual differences in WM capacity are a key driving force behind the speed and manner in which the numerical symbol-mapping process occurs. One possibility is that this occurs via WM-related differences in strategy selection.

It has recently been shown that variation in WM affects the strategies individuals employ to solve complex tasks. For instance, Beilock and DeCaro (2007) showed that when solving a multi-step math problem, the higher one's WM, the more likely one is to use a complex strategy (as opposed to a simpler 'short-cut' strategy). Likewise, Decaro, Thomas, and Beilock (2008) have shown that individual differences in WM impact how one learns complex categories, and they have suggested this is precisely because individuals higher in WM are more likely to use complex hypotheses to test category membership than their lower WM counterparts (for the role of WM-related differences in logical and hypothesis-testing processes more generally, see also De Neys, 2006; Evans, 2003; Stanovich & West, 2000).

If LWMs are less likely to hypothesis test than HWMs (Beilock & Decaro, 2007; Decaro et al., 2008; DeCaro, Carlson, Thomas, & Beilock, 2009), then while learning the quantities associated with abstract novel symbols, LWMs may use the simpler strategy of directly associating a quantity with the symbols' visual shape. This would leave the relative order of such symbols deducible, but only indirectly so, by referring back to the quantity each symbol represents. In this case, as the quantity of dots represented by a given symbol increases, numerical performance should decrease. Numerous studies have shown that representations of approximate, non-symbolic quantities (e.g., arrays of dots) become less precise as quantity increases (e.g., Buckley & Gillman, 1974; Moyer & Landauer, 1967).

HWMs, on the other hand, might be more inclined to go further. Specifically, as part of the learning process, HWMs might generate and test hypotheses regarding the symbol-set's relative order implied by symbol-quantity associations. In this case, if HWMs, in addition to learning the quantity of dots represented by each symbol, rely on a strategy that also emphasizes the relative order of the symbols during learning (whereby the quantity associated with a given symbol serves as a cue for learning the symbols' relative ordinal position), then the lowest and highest magnitude symbols may acquire special status as they represent the most extreme components of the symbol continuum (Leth-Steensen & Marley, 2000). Demonstrating the above mentioned WM differences would not only show that individual differences in a general cognitive capacity affect how we acquire numerical symbols, but also that strategies which emphasize ordinal relations among symbols play a key role in imbuing numerical representations with their full array of associative meaning.

It should be noted that previous studies have used novel training stimuli in simulating numerical and ordinal learning processes. Using both behavioral (Tzelgov, Yehene, Kotler, & Alon, 2000) and neuroimaging (Van Opstal, Verguts, Orban, & Fias, 2008) measures, researchers have shown that participants can make ordinal inferences about novel shapes that extend beyond the specific ordered pairs on which they were initially trained. In those studies, participants were shown pairs of novel shapes and received feedback after each judgment about which shape came later in terms of some unknown order (i.e., symbols were never associated with actual numerical quantities). By contrast, we were interested in whether the tendency to spontaneously infer ordinal relationships *over and above* simple symbol-quantity associations depends on individual differences in general cognitive abilities. Here it is crucial to note that the main goal of our task (from the participant's perspective) was to associate symbols with quantities, which allowed us to probe the relation between performance, WM, and the tendency to derive order-based strategies not explicitly required by the structure of the task.

To ensure that any observed differences in symbol-mapping as a function of WM capacity are indeed due to *strategic* differences related to emphasis of order information during symbol acquisition (rather than WM-related differences in the precision with which participants represent numerical quantities *per se*), we also examined whether WM capacity is related to representations of over-learned symbolic numbers (Arabic numerals) and non-symbolic approximate quantities (dot-arrays). Our goal was to show that HWMs are not simply better at discriminating symbolic and/or non-symbolic numerical ratios, but that HWMs' advantage in learning novel numerical symbol sets stems instead from strategic inferences of order-based information that LWMs are less likely to make.

2.1. Methods

2.1.1. Participants

Participants ($N = 51$; 34 female) were University of Chicago students (age: 18–36 years; $M = 22.5$ years) who took part in the experiment for course credit or monetary compensation.

2.1.2. Working memory measures

Individual differences in working memory were determined by taking the average of scores on two commonly used measures of WM: the automated reading- and operation-span tasks (aR-span and aO-span; Conway et al., 2005; Unsworth, Heitz, Schrock, & Engle, 2005). aR-span and aO-span were strongly correlated in our data [$r(49) = .630$, $p < .001$].

In the aR-span, participants judge the sensibility of sentences (e.g., "The only furniture Steve had in his first bowl was his waterbed.") and then remember a single letter. All participants perform 15 sentence-letter sequences, after which they are asked to recall the letters in the order presented. Sequences range from 3 to 7 letters in length. Three sequences of each length are presented to all participants, with the caveat that presentation order of the sequences themselves is randomized across participants. A participant's score for that sequence is accuracy $*$ length, where accuracy is whether all letters were recalled in the correct order (1 or 0) and length is the number of letters in the sequence in question. Scores for all sequences are then summed to determine the subject's final aR-span score (range: 0–75). aO-Span is identical to aR-span, with the exception that, instead of verifying the sensibility of sentences, participants verify whether simple math problems (e.g., $5 + 4 - 3 = 9$) have been correctly solved. A minimum level of accuracy must be maintained in the sentence/math-problem component of the task – 85% correct – for a participant's span-score to be considered valid. All working memory tasks were performed in a separate session that occurred within roughly one week of the main symbol-learning session (see below). Presentation order for the aO- and aR-span tasks was counterbalanced across participants.

It should be noted that in all primary analyses in Experiment 1 (and Experiment 2), WM was treated as a continuous variable. However, for ease of visualization and in certain analyses where it was deemed prudent in facilitating data interpretation, we also present results in terms of high and low WM groups. Groups were determined based on the upper and lower third of the average WM-span scores (in Experiment 1, HWMs: $n = 17$, $M = 62.50$, $SE = 1.48$, cutoff $T_U = 51$; LWMs: $n = 16$, $M = 21.81$, $SE = 2.45$, cutoff $T_L = 32$). For aO-span, Unsworth et al. (2005) report the following population norms: lower-quartile cutoff (Q_L) = 28, upper-quartile cutoff (Q_U) = 51. Thus, our cutoffs are overall in keeping with published norms.

2.1.3. Procedure

After giving informed consent, participants were told they were to learn to associate quantities of dots with novel visual shapes (or 'symbols'). Symbols were generated by randomly rearranging 16 20 pixel^2 composite red and white shapes in a 4×4 (80 pixel^2) grid. This method ensured that all symbols were novel and comprised of the same basic features, and that these features were equated in terms of quantitative dimensions (size, amount of 'redness', etc.).

Participants first learned to recognize the symbols in terms of their perceptual features. Participants studied 6 target symbols as 3 cm^2 laminated cards and were informed of their proper orientation. After 2 min, the symbols were removed and randomly intermixed with 18 distracter symbols (consisting of a randomly arranged composite of the same 16 features used to make the target symbols). Participants had 2 min to identify the 6 target symbols they had studied (with no false alarms). If this did not occur, the procedure was repeated (iterations to reach criterion: $M = 1.75$; $SE = .12$; $Max = 4$).

Participants next completed a computerized training session where they learned to associate 6 different dot-quantities (9, 18, 27, 36, 45 and 54 dots) with the 6 target symbols (Fig. 1). The number of dots associated with a given symbol was fixed throughout the experiment. Each symbol was presented with its respective dot-quantity 8 times (using 8 different arrays). This constituted 1 block. Once all 6 symbols had been presented (1 block for each symbol), participants were allowed to rest. The block sequence was then repeated (8 times in total). Across different sequences, blocks were presented in pseudo-randomized order to prevent participants from inferring the symbols' relative numerical order based on presentation order (see Fig. 2a for an overview of the symbol-training sequence.)

Each time a symbol was presented, it was located at the central fixation position. Eight separate arrays of dots (all of the same quantity) were briefly flashed nearby – one array at a time – each in one of 8 possible locations in a 3×3 grid of 256 pixel^2 tiles (with the symbol always occupying the center tile). Each tile subtended to a visual angle $\approx 4.6°$ per side. Each array was presented for 500 ms, thus ensuring that participants could only associate estimated approximate numerical magnitudes with each symbol, rather than count the exact number of dots. Importantly, the dot-arrays were counterbalanced across numerosities with respect to continuous parameters, including individual and aggregate dot-area, contour length (or group perimeter), and local density (or average minimum distance between dots). This was done to reduce the possibility that these factors might serve as indirect cues to the number of dots associated with each symbol. As no dot-array was ever repeated, pattern recognition could not provide a reliable cue for learning the associated dot-quantities. Stimuli were displayed at 1280×1024-resolution on a 19.1″ Dell flat-panel monitor located 1 m from the participant. At this distance, each symbol subtended to a visual angle $\approx 1.5°$.

After training, participants performed several tasks designed to gauge symbol learning. Individuals first performed two comparison tasks using the novel symbols (the *greater-than* and *ascending* tasks, see below). To assess individuals' ability to represent basic numerical quantities, participants next performed these comparison tasks with dot-arrays and Arabic numerals. Participants then performed a global ordering task using the novel symbols – designed to gauge their knowledge of the symbols' overall order. Finally, individuals were asked about any strategies they used to learn the novel symbols. Participants were then compensated for their time, thanked, and debriefed.

2.1.4. Tasks

Symbol comparison tasks. Individuals performed two numerical comparison tasks (*greater-than* and *ascending* tasks) using the recently acquired novel symbols. In both tasks, two symbols were presented simultaneously, one on either side of fixation (Fig. 2c). The symbols remained on the screen for 1000 ms, after which the screen went blank. This ensured that comparison stimuli were viewed for a maximum of 1000 ms regardless of time required to respond.

In the *greater-than* task, if the symbol that represented the greater quantity of dots was on the left side of the screen, participants pressed a key with their left-index-finger; if it was on the right, they pressed a second key with their right-index-finger. In the *ascending* task, if the symbols were in ascending order from *left to right* in terms of the quantities they represented, participants made a left-index-finger key press. If, instead, the symbols were in ascending order from *right to left*, they made a right-index-finger response.

For each of the above tasks, participants completed four blocks of 30 trials. All possible symbol-pair combinations

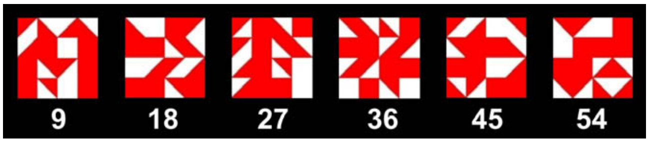

Fig. 1. This figure depicts the actual novel symbols used in Experiment 1. The Arabic numeral beneath each symbol represents that actual quantity of dots with which it was associated.

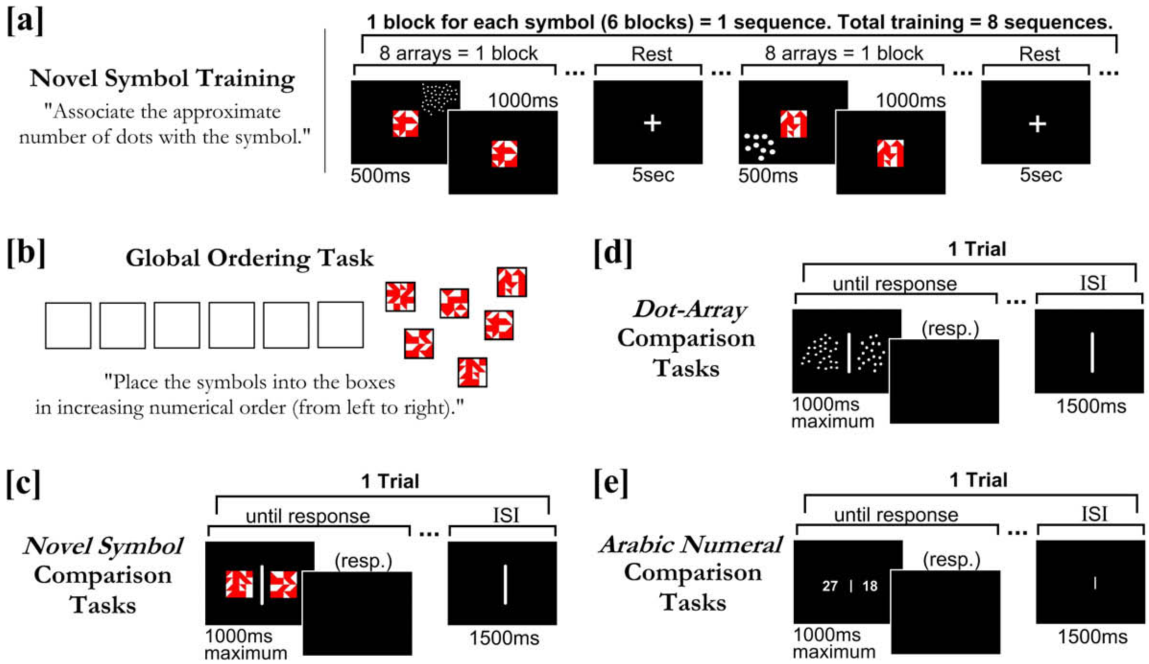

Fig. 2. This figure depicts Experiment 1 training and assessment tasks. Panel (a) shows the novel-symbol training procedure. Note that block-order was counterbalanced across presentation sequences. Panel (b) shows the global ordering task. Participants were handed copies of the symbols as small laminated cards (shuffled beforehand), and were asked to place the symbols in boxes printed on the provided piece of paper. No time limit was given for this task. Panels (c–e) show the comparison tasks for novel symbol, dot-array and Arabic numeral stimulus types, respectively. Note that stimulus presentation for the two types of comparison tasks (greater-than and ascending) was identical. Stimuli for the comparison tasks were presented for a maximum of 1000 ms. If participants responded before this time limit, the sequence immediately moved to the inter-stimulus-interval (ISI); otherwise, stimuli disappeared after 1000 ms, leaving a blank screen, which remained until subjects responded.

were repeated twice per block, once per side (i.e., within a block, a given symbol pair was presented once in left–right and once in right–left orientation). Across all blocks, each pair was presented a total of eight times. Task order was counterbalanced across participants.

Dot-array and Arabic numeral comparisons. Participants performed the same two comparison tasks described above (i.e., *greater-than* and *ascending* tasks), with the exception that (a) arrays of dots and (b) two-digit numbers written as Arabic numerals were used instead of the novel symbols (Fig. 2d–e). Note that the quantities used for dot and numeral comparisons were the same as those represented by the novel symbols. Dot-arrays were counterbalanced in terms of continuous parameters in the same manner as the novel-symbol training stimuli. To eliminate the possibility that pattern recognition could be used as a cue in making comparison judgments, all dot-stimuli were presented only once and did not overlap with the sample used in the novel-symbol training procedure.

Global ordering task. This task assessed participants' representations of the novel symbols' overall numerical order (see Fig. 2b). Participants were presented with a piece of paper containing six horizontally arranged boxes and given the same laminated cards (in a random order) depicting the six symbols used in the first symbol recognition phase. Participants were asked to arrange the six symbols in increasing order from left to right in the six boxes on the provided paper. There was no time limit for this task, and no feedback was provided.

2.2. Results

2.2.1. Novel symbols: global ordering task

Accuracy was coded in terms of whether participants placed the novel symbols in the correct positions. For example, if the numerically smallest symbol was placed in the position farthest to the left, that individual received a 1 for that symbol; otherwise they received a 0. Thus, chance performance for a given symbol was 16.7% correct, across subjects. Data were analyzed using a WM $\times$ 6(symbol: 9, 18, 27, 36, 45, 54 dots) mixed-design analysis of covariance (ANCOVA), with WM as a between-subjects continuous factor and symbol as a within-subjects discrete factor.

This analysis produced a significant main effect of symbol [$F(5, 245) = 6.92$, $p < .001$], in which there was an overall decrease in accuracy as the quantity a symbol represented increased. There was also a significant main effect of WM [$F(1, 49) = 17.78$, $p < .001$]. Overall, the higher one's WM, the more accurate was one's performance [$r(49) = .516$, $p < .001$]. Crucially, these main effects were qualified by a significant WM $\times$ symbol interaction [$F(5, 245) = 4.59$, $p < .001$]. There was no significant relation between WM and accuracy for the numerically smallest symbol [$r(49) = .133$, $p = .352$]. However, there was a strong positive relation between WM and accuracy at the numerically greatest symbol [$r(49) = .631$, $p < .001$]. As can be seen in Fig. 3, this correlation was driven by an upturn in accuracy for HWMs at the highest symbols

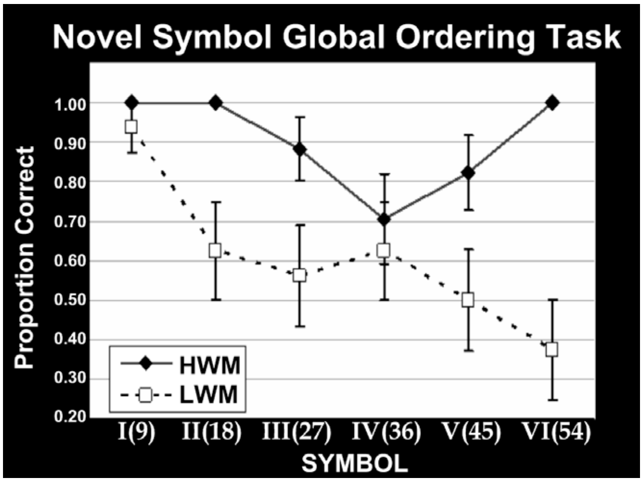

Fig. 3. This figure depicts cell means (proportion correct) on the Experiment 1 global ordering task. For the purposes of visualizing the data, the factor WM is presented as high (upper third) and low WM (lower third) groups. Symbols are noted using Roman numerals; the number in parentheses is the quantity of dots with which that symbol was associated. Error bars represent standard errors of the mean.

but a continued trend toward decreasing accuracy for LWMs.[1]

To better understand this interaction between symbol quantity and WM, we divided individuals into higher and lower WM groups based on the upper and lower third of the average WM-span scores. LWMs' ordering accuracy (as a function of increasing symbol magnitude) showed a significant linear contrast effect [$F(1, 15) = 13.97$, $p = .002$] but not a quadratic contrast effect ($F < 1$). HWMs showed the reverse pattern – a significant quadratic effect [$F(1, 16) = 6.18$, $p = .024$], and a marginal linear effect [$F(1, 16) = 3.43$, $p = .093$]. In other words, both LWMs and HWMs showed a significant effect of symbol quantity (i.e., accuracy significantly varied as function of symbol quantity). However, this contrast was significant for LWMs only when a linear function was assumed and, for HWMs, reached significance only when a quadratic function was assumed. These results are consistent with the hypothesis that LWMs' performance decreased as a function of numerical size – with systematically worsening performance as the quantity a symbol represented increased. HWMs' performance was best at the numerical endpoints (Fig. 3) – a pattern of results consistent with a strategy that emphasized the symbols' overall numerical order.

2.2.2. Novel symbols: comparison tasks

To assess WM differences in how participants represented the novel symbols in the comparison tasks, we performed separate WM × 5 (comparison ratio: 1:2, 2:3, 3:4, 4:5, 5:6) ANCOVAs on accuracy (proportion correct) and reaction time (RT) for correct trials. Analyses were limited to the five contiguous symbol pairs (e.g., 1:2 or 'first versus second symbol', 2:3 or 'second versus third' symbol, etc.) because we were specifically interested in WM differences at the ends of the ratio continuum (i.e., the largest – 1:2 – and smallest – 5:6 – ratios). In addition, this allowed for a constant numerical difference (or 'distance') of 9 at each ratio.

Greater-than task. "Which symbol represents the greater quantity?" In terms of accuracy, consistent with the global ordering task analysis above, there was a significant main effect of ratio [$F(4, 196) = 3.49$, $p = .009$], in which an overall decrease in accuracy was observed as the quantity a symbol represented increased. In addition, the main effect of WM reached significance [$F(1, 49) = 9.81$, $p = .003$], such that WM capacity was positively related to performance accuracy [$r(49) = .408$, $p = .003$]. There was also a significant WM × ratio interaction [$F(4, 196) = 3.45$, $p = .009$]. Although there was no significant relation between WM and accuracy for the 1:2 ratio [$r(49) = -.067$, $p = .642$], there was a strong positive relation between WM and accuracy at the 5:6 [$r(49) = .595$, $p < .001$]. As in the global ordering task above, this relationship was driven by an upturn in accuracy for HWMs at the highest symbols but a

[1] It is important to point out that the different levels of the factor symbol are not, strictly speaking, independent of one another, which may lead one to question the use of an ANCOVA to analyze these data, a method which relies on the assumption of independent samples. First, one of the main fears with respect to interrelatedness among observations, especially in a repeated-measures ANOVA, is that this will result in inflation of the Type I error rate (α). However, in the current case, even after correcting degrees of freedom using the Greenhouse–Geisser estimate of sphericity ($\varepsilon = .715$), both the main effect of symbol [$F(3.6, 175.1) = 6.92$, $p < .001$] and the WM × symbol interaction [$F(3.6, 175.1) = 4.59$, $p = .002$] remained significant. Second, using the Mann–Whitney non-parametric test for differences between two related groups (here we used our high and low WM groups), we found a difference between groups for the numerically largest [symbol VI: $Z = -3.85$, exact significance $p = .002$] but not the numerically smallest symbol [symbol I: $Z = -1.03$, exact significance $p = .763$]. This finding confirms our central hypothesis that increased reliance on an ordering strategy with higher working memory should have the largest effect on accuracy for the numerically largest symbol. Finally, it should be noted that a very similar pattern of results was found for the global ordering task and the greater-than comparison task. Comparing pairs in the latter task does not carry the same dependence that the permutation-constrained responses do in the global ordering task. Because the results for these two tasks were highly similar, it seems unlikely that the effects reported in the global ordering task arose due to the lack of complete independence between levels.

Table 1
Accuracy (proportion correct) and reaction time (ms) data for novel numerical symbols. Data are reported in terms of WM group (upper/lower thirds). Cells: mean *(SE)*.

(a) WM × ratio: greater-than comparison task (prop. correct)					
Ratio	1:2	2:3	3:4	4:5	5:6
HWMs	.882 *(.028)*	.912 *(.025)*	.853 *(.071)*	.716 *(.089)*	.980 *(.014)*
LWMs	.906 *(.033)*	.677 *(.092)*	.781 *(.074)*	.792 *(.077)*	.615 *(.056)*
(b) WM × ratio: greater-than comparison task (RT)					
Ratio	1:2	2:3	3:4	4:5	5:6
HWMs	1293 *(84)*	1508 *(102)*	1559 *(115)*	1552 *(108)*	1177 *(48)*
LWMs	1406 *(147)*	1790 *(158)*	1624 *(128)*	1652 *(140)*	1401 *(125)*
(c) WM × ratio: ascending comparison task (prop. correct)					
Ratio	1:2	2:3	3:4	4:5	5:6
HWMs	.931 *(.032)*	.922 *(.029)*	.833 *(.069)*	.755 *(.081)*	.843 *(.044)*
LWMs	.823 *(.047)*	.656 *(.067)*	.781 *(.056)*	.688 *(.071)*	.635 *(.043)*
(d) WM × ratio: ascending comparison task (RT)					
Ratio	1:2	2:3	3:4	4:5	5:6
HWMs	1443 *(156)*	1899 *(167)*	1977 *(153)*	1941 *(158)*	1438 *(75)*
LWMs	1750 *(159)*	2219 *(180)*	2179 *(194)*	2178 *(189)*	1875 *(132)*

continued trend toward decreasing accuracy for LWMs (Table 1a).

If accuracy decreases as comparison ratio approaches 1, this would be consistent with a direct association between approximate dot-quantity and symbol. However, if greater accuracy is seen for the symbol pairs involving the numerically largest symbols (i.e., the ratio 5:6), this would suggest participants used additional global-order information that privileged representation of symbols at the minimum and maximum edges of the set. Consistent with this notion, and in keeping with results from the global ordering task described above, when we divided individuals based on their WM scores into bottom and top one-thirds of the WM distribution, LWMs' accuracy (as a function of ratio) showed a marginally significant *linear* contrast effect [$F(1, 15) = 4.24$, $p = .057$] but no significant quadratic contrast effect ($F < 1$). HWMs showed the reverse pattern, with a significant *quadratic* [$F(1, 16) = 7.12$, $p = .017$] but non-significant linear ($F < 1$) contrast effect.

In terms of RT, the main effect of WM did not approach significance ($F < 1$). However, the main effect of ratio was significant [$F(4, 196) = 3.75$, $p = .006$]. The WM × ratio interaction was marginally significant [$F(4, 196) = 2.06$, $p = .070$]. As seen in Table 1b, RTs for both groups showed an increase as ratio approached 1, and a drop in RTs at the 5:6 ratio. It should be noted, however, that this drop was highly significant for HWMs [$t(16) = 4.03$, $p < .001$] but not LWMs [$t(15) = 1.89$, $p = .078$]. This finding suggests HWMs' accuracy increase reported above for the 5:6 ratio, was not due to a speed-accuracy trade-off.

Ascending task. "Are the two symbols ascending or descending?" In terms of accuracy, only the main effect of WM reached significance [$F(1, 49) = 16.55$, $p < .001$]. There was a positive relation between WM capacity and accuracy [$r(49) = .502$, $p < .001$]. A marginally significant effect of ratio was also observed, in which an overall decrease in accuracy was observed as the quantity a symbol represented increased [$F(4, 196) = 2.22$, $p = .068$]. In addition, despite the lack of a significant WM × ratio interaction [$F(4, 196) = 1.59$, $p = .180$], Table 1c reveals an overall pattern qualitatively similar to that seen for the greater-than comparison and global ordering tasks. Specifically, there was a decrease in accuracy for LWMs and an upturn for HWMs at ratios closest to 1. In terms of RT, only the main effect of ratio reached significance [$F(4, 196) = 3.43$, $p = .010$]. For all other effects, $F < 1$ (see condition means in Table 1d).

Thus, the data from the novel-symbol comparison tasks mirrored the results of the global ordering task. Overall, accuracy was highest for those with highest WM capacity. Moreover, while accuracy did not differ as a function of WM when comparing the numerically smallest symbols (i.e., the 1:2 ratio), comparisons involving the numerically largest symbols (5:6 ratio) yielded greater accuracy for those higher in WM capacity. This finding is consistent with the notion that HWMs are representing more than just the quantities associated with specific symbols, extending their representations of the novel symbols to include how they relate to each other in terms of overall numerical order. As we will see below, participants' strategy reports lend additional support to this interpretation of the data.

2.2.3. Strategy reports

Participants responded to two open-ended questions regarding their symbol-learning strategies: (1) "During training, how did you estimate the number of dots associated with each symbol? What strategies did you use?" (2) "What strategies did you use to complete the [novel symbol] comparison tasks?" For each question, if participants explicitly reported using a rank or ordering based strategy, then that question was scored as 1; otherwise, it was scored as 0. Scores for the two questions were summed to yield a range 0–2, with 0 indicating no report of ordering strategy, and 2 indicating ordering was used during both the learning and the comparison phases of the experiment. Results showed that the higher one's WM, the more likely one was to adopt (or at least report having adopted) an explicit ordering strategy [$r(49) = .385$, $p = .005$] in acquiring and using the novel numerical symbols. Note that if responses to Questions 1 (symbol learning) and 2 (symbol comparison) were related separately to WM, the results do not change [Question 1: $r(49) = .340$, $p = .015$; Question 2: $r(49) = .315$, $p = .024$].

If improved performance for the numerically largest symbols is indeed indicative of having made additional ordinal inferences, then ordering strategy reports should be related more specifically to better accuracy on the numerically larger items than the numerically smallest items. This was precisely what was found. Reported ordering strategy use was positively related to global-ordering accuracy on symbol VI [$r(49) = .354$, $p = .011$] but not symbol I [$r(49) = .125$, $p = .381$]. Similarly, for the greater-than and ascending comparison tasks, a positive relation was seen between ordering strategy and accuracy for the 5:6 [greater-than: $r(49) = .298$, $p = .034$; ascending: $r(49) = .242$, $p = .087$] but not the 1:2 comparison ratio [greater-than: $r(49) = -.031$, $p = .828$; ascending: $r(49) = .008$, $p = .956$].

2.2.4. Dot-array and Arabic numeral comparisons

Dot and Arabic numeral comparisons were examined separately in a manner similar to the novel symbols above, using WM × 5(ratio: 1:2, 2:3, 3:4, 4:5, 5:6) ANCOVAs.

Dot-comparisons. With respect to the greater-than comparison task, in terms of accuracy, only the main effect of ratio reached significance [$F(4, 196) = 9.66$, $p = .003$]. Accuracy generally decreased as ratio approached 1 (Table 2a). For RT, again only the main effect of ratio reached significance [$F(4, 196) = 14.96$, $p < .001$], with RTs increasing as ratio approached 1 (Table 2b).

Results for the ascending comparison task were similar to the greater-than task. In terms of accuracy, only the main effect of ratio reached significance [$F(4, 196) = 30.29$, $p < .001$], with accuracy decreasing as ratio approached 1 (Table 2c). In terms of RT, again only the main effect of ratio reached significance [$F(4, 196) = 26.88$, $p < .001$], with RTs increasing as ratio approached 1 (Table 2d).

Arabic numeral comparisons. With respect to the greater-than comparison task, in terms of accuracy, no effect reached significance (Table 3a). In terms of RT, the only effect to reach significance was the main effect of WM [$F(1, 49) = 8.48$, $p = .005$]: the higher one's WM, the faster one's overall RT [$r(49) = -.384$, $p = .005$] (Table 3b).

With respect to the ascending comparison task, in terms of accuracy, no effect reached significance (Table 3c). In terms of RT, there was a main effect of WM [$F(1, 49) = 8.20$, $p = .006$], such that the higher one's WM, the faster one's overall RT [$r(49) = -.379$, $p = .006$]. The effect of ratio was also significant [$F(4, 196) = 5.23$, $p = .001$], where RT increased as ratio approached 1 (Table 3d).

To summarize, dot-comparisons showed decrements in performance (decreasing accuracy and increasing RTs) as the ratio between numerosities approached 1. Crucially, the magnitude of this effect did not depend on individual differences in WM capacity. Comparisons involving Arabic numerals showed a qualitatively similar effect of ratio on performance. Moreover, as in the dot comparison, this effect of ratio was not dependent upon differences in participants' WM capacity.

2.3. Discussion

We examined the impact of individual differences in WM on numerical symbol-mapping and found that, overall, HWMs performed more accurately than LWMs. This performance difference was greatest at the numerically largest symbols, a pattern consistent with the view that HWMs were going beyond simple symbol-quantity associations and generating hypotheses about the symbol-set's overall rank order. In contrast, LWMs did not appear to place the additional emphasis on learning the symbols' global order that was exhibited by HWMs; rather, LWMs relied on simple symbol-quantity associations. Participants' strategy reports support this interpretation. The point at which a global-ordering strategy and a simple 'approximate-and-associate' strategy should yield divergent performance (i.e., the numerically largest symbol) is precisely the point at which explicitly reporting having used an ordering strategy most strongly related to both comparison and global-ordering accuracy.

Taken together, these data converge to support the view that, when acquiring a set of novel numerical symbols, HWMs' out-performance of LWMs is related to an added emphasis on learning the symbols' relative numerical order. Moreover, these data indicate that a greater understanding of symbols as an ordered set facilitates the numerical sym-

Table 2
Accuracy (proportion correct) and reaction time (ms) data for <u>dot-arrays</u>. Data are reported in terms of WM group (upper/lower thirds). Cells: mean *(SE)*.

(a) WM × ratio: greater-than comparison task (prop. correct)					
Ratio	1:2	2:3	3:4	4:5	5:6
HWMs	.990 *(.010)*	.990 *(.010)*	.980 *(.013)*	.863 *(.029)*	.853 *(.035)*
LWMs	.990 *(.010)*	.969 *(.017)*	.948 *(.025)*	.842 *(.036)*	.865 *(.031)*
(b) WM × ratio: greater-than comparison task (RT)					
Ratio	1:2	2:3	3:4	4:5	5:6
HWMs	585 *(27)*	670 *(35)*	721 *(44)*	781 *(41)*	773 *(45)*
LWMs	690 *(42)*	763 *(41)*	755 *(49)*	780 *(50)*	818 *(62)*
(c) WM × ratio: ascending comparison task (prop. correct)					
Ratio	1:2	2:3	3:4	4:5	5:6
HWMs	.941 *(.020)*	.990 *(.010)*	.961 *(.018)*	.831 *(.035)*	.760 *(.035)*
LWMs	.979 *(.014)*	.948 *(.020)*	.873 *(.029)*	.750 *(.037)*	.729 *(.052)*
(d) WM × ratio: ascending comparison task (RT)					
Ratio	1:2	2:3	3:4	4:5	5:6
HWMs	819 *(46)*	826 *(58)*	1007 *(73)*	1099 *(75)*	1185 *(80)*
LWMs	900 *(58)*	972 *(73)*	1084 *(74)*	1127 *(88)*	1207 *(78)*

Table 3
Accuracy (proportion correct) and reaction time (ms) data for Arabic numerals. Data are reported in terms of WM group (upper/lower thirds). Cells: mean *(SE)*.

(a) WM × ratio: greater-than comparison task (prop. correct)					
Ratio	1:2	2:3	3:4	4:5	5:6
HWMs	.931 *(.032)*	.951 *(.019)*	.961 *(.023)*	.945 *(.034)*	.882 *(.037)*
LWMs	.969 *(.017)*	.969 *(.023)*	.969 *(.017)*	.937 *(.054)*	.927 *(.050)*
(b) WM × ratio: greater-than comparison task (RT)					
Ratio	1:2	2:3	3:4	4:5	5:6
HWMs	664 *(28)*	640 *(27)*	690 *(26)*	697 *(27)*	704 *(23)*
LWMs	753 *(38)*	750 *(38)*	805 *(31)*	781 *(42)*	774 *(32)*
(c) WM × ratio: ascending comparison task (prop. correct)					
Ratio	1:2	2:3	3:4	4:5	5:6
HWMs	.912 *(.025)*	.971 *(.016)*	.863 *(.036)*	.922 *(.029)*	.912 *(.029)*
LWMs	.885 *(.062)*	.938 *(.052)*	.896 *(.043)*	.885 *(.064)*	.875 *(.058)*
(d) WM × ratio: ascending comparison task (RT)					
Ratio	1:2	2:3	3:4	4:5	5:6
HWMs	760 *(40)*	782 *(35)*	784 *(42)*	854 *(31)*	863 *(28)*
LWMs	923 *(57)*	892 *(59)*	958 *(44)*	976 *(55)*	1040 *(51)*

bol-mapping process. In other words, the extent to which relative order information is emphasized in symbolic numerical representations may impact the success with which these symbols are acquired in the first place.

Importantly, performance on comparison tasks involving neither dot-arrays nor Arabic numerals yielded an interaction between WM and comparison ratio. This suggests that the differences seen in acquisition stages (i.e., in the novel-symbol data) are not due to WM-related differences in the precision with which symbolic and non-symbolic quantities are represented *per se*. In addition, it is not the case that individuals lower in WM will *never* understand relative order information in numerical symbols. Rather, our results indicate that, *during the initial symbol-quantity mapping process*, LWMs are not as predisposed as their HWM counterparts to learn ordinal information.

It should be noted, however, that higher WM was related to faster response times in the Arabic numeral comparison tasks. On the one hand, this may simply be due to WM-related differences in general processing speed (see Engle, Tuholski, Laughlin, & Conway, 1999; Verhaeghen & Salthouse, 1997). On the other hand, HWMs may retain an important overall edge over LWMs when processing specifically numerical symbols. It is possible that the early inclination for HWMs to emphasize relative order information when mapping quantities to symbols may lead to a greater emphasis on ordered, relational numerical processes in overlearned Arabic numerals. That is, HWMs may have quicker access to ordinal information in overlearned numerical symbols in general. Thus, the facts '54 comes after 45' and '9 comes before 18' would both be accessed more rapidly by HWMs than LWMs because of this more efficient processing of relative numerical order. As a result, for overlearned numerical formats such as Arabic numerals, higher WM capacity should be related to faster symbolic numerical comparisons, regardless of ratio.

Of course, it is impossible to distinguish between the above two alternatives in Experiment 1, given that Arabic numeral comparisons represent one of the simplest numerical tasks one can perform with numerical symbols. In Experiment 2, however, participants performed an Arabic numeral task that involved a more explicit ordering component. In this way, we hoped to better assess whether individual differences in WM translate into differences in the efficiency with which ordinal information about well-learned Arabic numerals is accessed.

Finally, it is worth discussing a potential alternative explanation for why HWMs showed an accuracy increase at the numerically largest symbol. The symbols at the endpoints have competing associations in only one numerical direction, whereas the remaining 4 symbols are situated with symbol-quantity associations in both numerical directions (larger and smaller). In this way, one need not invoke ordinal inference on the part of HWMs to produce the U-shaped accuracy function displayed by this group. Rather, fewer competing associations at the endpoints may have lead to greater accuracy in retrieving the quantity associations for the smallest and largest symbols. However, if this were the case, one would expect LWMs to show a U-shape similar to that seen in HWMs, because LWMs' accuracy should also be boosted by less competing associations – indeed possibly more so than HWMs'. Given that our data do not reflect this, one might further push a strictly association-based account by postulating that LWMs' performance at the highest symbol was compromised by greater numerical overlap with neighboring symbols than the lowest symbol, as already described in our original hypothesis. Yet, it seems difficult to imagine why the purported impact of such representational overlap would be so much greater for LWMs than HWMs (our HWM group was perfect – 100% accuracy – whereas LWMs were barely above chance on the numerically largest symbol) without invoking another mechanism altogether. As already discussed, the most plausible mechanism in our minds is that HWMs adopted an order-based strategy and went on to infer that the asymmetry in competing associations for the numerically largest symbol *means that it is the numerically largest symbol*. Indeed, our strategy report data (explicit reports of using an order- or rank-based strategy) correlated both with WMC and performance on the numerically largest symbol, which lends empirical support to the notion that HWMs tended to go beyond symbol-quantity associations and infer and encode more complex symbol–symbol ordinal associations. Finally, as we shall see in Experiment 2 (which does not rely on any sort of training or learning mechanism within the experiment itself), the hypothesized tendency for HWMs to invoke more complex ordinal processing in numerical symbols extends to highly overlearned Arabic numerals as well.

3. Experiment 2

In Experiment 2, we turned to a second context to examine the question of how individual differences in WM shape one's ordinal understanding of numerical symbols. Namely, we examined this relationship using a task in which participants were explicitly asked to make order-related judgments about highly enculturated numerical symbols (Arabic numerals). To this end, a three-number ordering task was devised. Participants saw three one-digit Arabic numerals and were asked to judge whether *all three* numbers were in increasing order (from left to right). Not only does completing this type of task require accessing more ordinal information than the two-number comparison task used in Experiment 1, but a three-number task also allows for the manipulation of the accessibility of order information in terms of patterned relationships between the three digits. This enables us to directly test whether HWMs are accessing a richer set of ordinal associations than LWMs when processing overlearned symbolic numbers.

If only two digits are used (as with the comparison tasks used in Experiment 1), the question 'Are these digits in increasing order?' can easily be reduced to a simple magnitude comparison task. Using three digits makes this reduction more difficult because one must check the results of any pairwise comparisons against a rule that tells one whether all are in order or not. For example, in the case of {1, 2, 3}, one may compare 1 to 2, and then 2 to 3, but to answer the question, 'Are these digits in increasing order?' one needs to perform an additional operation. One could explicitly compare 1 to 2, 2 to 3, and 1 to 3, and then check this against the rule that if the right digit is greater in all three cases then the answer is always 'yes'; or one could skip the last comparison via explicit transitive inference ($2 > 1$, $3 > 2 \rightarrow 3 > 1 \rightarrow$ 'yes'). In all cases, one must check that the results of these comparisons are themselves properly ordered. In this way, we believe this task draws more heavily on representations of relative order information than a simple magnitude comparison task.

That said, a more efficient way of solving these problems would be to compare stimuli via direct retrieval to known ordered instances from memory. We hypothesized that HWMs have a richer set of such instances – perhaps as a result of being more inclined to have inferred and encoded such instances in the first place, as indicated by Experiment 1. To test this hypothesis, we asked participants to tell us whether sequences of three numbers were in increasing order (from left to right) and looked at their success in doing this (1) as a function of individual differences in WM, and (2) as a function of the type of relationship between the three numbers.

We assumed that a sequential pattern, such as {1, 2, 3} would contain strong ordinal associations between the sequential numbers (i.e., 1 to 2 and 2 to 3). This assumption is supported by recent evidence showing that participants are faster to compare two Arabic numerals when they are separated by a numerical distance of 1 than when this distance is greater than 1 (Turconi et al., 2006). This is a reversal of the standard distance effect in which one typically finds that the larger the numerical distance between compared numbers, the faster and more accurately one tends to respond (e.g., Moyer & Landauer, 1967). Importantly, reversal of the distance effect is seen only when the two numbers are in left–right increasing order (e.g., 7–8, and not 8–7; Turconi et al., 2006), perhaps due to increased experience with sequential digits given frequent repetition of the integer count sequence (Gallistel & Gelman, 1992). In other words, numerical symbols are strongly ordinal in their representations, but asymmetrically so.

In terms of the current paradigm, this means that presentation of the *sequential* numbers {1, 2, 3} should lead to automatic activation of the ordinal relationship between the numbers in this set, biasing initial responses to be 'yes, this set is in order', even when stimuli may not be. In this way, participants should be much faster in saying that {1, 2, 3} is in order compared to saying that {2, 1, 3} is not in order. We hypothesized that this difference would hold regardless of WM because, by adulthood, all participants should be highly familiar with the integer count sequence. In other words, subtracting faster 'yes' from slower 'no' RTs should yield a large positive value, the magnitude of which should not depend on WM capacity.

By contrast, we hypothesized that a numerically *skewed* set such as {1, 6, 8} should contain far weaker ordinal associations between constituent numbers, as it is not part of the canonical count sequence. Thus, the difference between saying 'no' {6, 1, 8} is not in order and 'yes' {1, 6, 8} is in order should be far less than the 'no' minus 'yes' difference for the sequential stimuli {1, 2, 3}. In this way, the weak ordinal associations in skewed trials should lead to minimal differences between 'no' and 'yes' responses, regardless of WM capacity.

Crucially, given that in Experiment 1 we found evidence that HWMs are more likely to go beyond simple associations and infer more complex ordinal relationships in symbolic numerical stimuli, we hypothesized that a third type of ordered set might fall somewhere in between these two extremes. Arithmetically *balanced* patterns (i.e., where the arithmetic mean equals the median) such as {2, 4, 6} or {3, 6, 9}, though perhaps not as obviously ordered as sequential numbers (e.g., Jou, 2003, Experiment 2), might nonetheless contain ordinal associations stronger than those in skewed sets. This might be due to familiarity with odd/even counting routines or the fact that interval-widths are of constant (i.e., linear) magnitude.

With respect to individual differences in WM capacity, then, we hypothesized that only HWMs' richer set of ordinal associations would lead to a bias in seeing balanced sets as in order. As a result, HWMs should show a positive difference for 'no' minus 'yes' balanced trials that LWMs do not. To state this last prediction about balanced trials another way, although highly routinized procedures like the integer count sequence may bias everyone to 'see' sequential sets of symbolic numbers as in order, a deeper bias to process symbolic numbers ordinally may vary with individual differences in WM capacity, such that HWMs

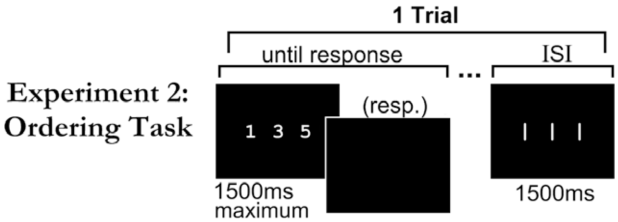

Fig. 4. Experiment 2 ordering task: example trial. Stimuli were presented for a maximum of 1500 ms. If participants responded before this time limit, the sequence immediately moved to the inter-stimulus-interval (ISI); otherwise, stimuli disappeared after 1500 ms, leaving a blank screen, which remained until subjects responded.

are more likely to show such a bias when assessing order information in balanced sets.

The above result would be consistent with the central finding in Experiment 1, namely that those higher in WM are more likely to go beyond immediate associations and infer less obvious ordinal relations in symbolic numerical stimuli. In Experiment 2, we look for evidence in *Arabic numerals* that, beyond simple associations between sequential numbers, HWMs have established strong ordinal associations between non-sequential numerical patterns as well. Such a result would provide important evidence that WM-related differences in numerical symbol acquisition, if left unchecked, may have implications for more complex ordinal processing in well-learned symbolic mathematics.

3.1. Methods

3.1.1. Participants

Participants ($N = 30$; 17 female) were University of Chicago students (age: 18–31 years; $M = 21.3$ years) drawn from the same population as the first experiment. As in Experiment 1, individual differences in WM were determined by the average score on the aO- and aR-span tasks. These measures were highly correlated [$r(28) = .539$, $p = .002$]. As in Experiment 1, WM was primarily analyzed as a continuous measure; however, data are also presented in terms of high and low WM groups (HWMs: $n = 10$, $M = 64.30$, $SE = 1.32$, $T_U = 56$; LWMs: $n = 10$, $M = 25.00$, $SE = 3.21$, $T_L = 32$).

3.1.2. Procedure

In a single testing session, participants completed 3 tasks in total: aO-span, aR-span and the ordering task. All participants performed the three tasks in this order, with the exception that the order of the two span tasks was counterbalanced across participants. For all tasks, stimuli were displayed at 1280 × 1024-resolution on a 19.1″ Dell flat-panel monitor approximately 1 meter from the participant.

Ordering task. On each trial, participants saw three, 1-digit Arabic numerals (range: 1–9) arranged horizontally. Stimuli were in white 24-point Courier font presented on a black background. The distance between the left and rightmost numerals subtended to approximately 6° of visual angle. Numerals were separated from one another by ~1.5° of visual angle.

Participants' task was to indicate whether all three numerals were in increasing order from left to right. If all three digits were in increasing order, participants were to press the 'Z' key on a standard keyboard with their left hand. Otherwise (if any two numeral pairs were in decreasing order from left to right), participants were to press the 'M' key with their right hand.

For each trial, the three one-digit Arabic numerals were presented on the screen for a maximum of 1500 ms. If participants had not responded by this time, the screen went blank until they did. Whenever a response was detected, three white vertical pipes appeared on the screen for a fixed inter-stimulus interval of 1500 ms. An example trial is shown in Fig. 4.

Three stimulus types were used: sequential, balanced and skewed. *Sequential* stimuli, when properly ordered (i.e., on 'yes' trials), formed a three-digit segment of the integer count sequence – e.g., {1, 2, 3}, {3, 4, 5}, {7, 8, 9}. *Balanced* stimuli, when properly ordered, formed a pattern with constant intervals between adjacent numbers. Another way to think about balanced trials is that they were comprised of non-sequential sets where the median and mean numbers were equal – e.g., {2, 4, 6}, {3, 6, 9}, {1, 5, 9}. *Skewed* stimuli, when properly ordered, formed a pattern with unequal intervals between adjacent numbers. In other words, the median and mean of skewed sets were never equal – e.g., {1, 6, 8}, {4, 8, 9}, {1, 2, 7}. For each subject, skewed sequences were randomly drawn from the set of all possible skewed sequences. In the case of ordered ('yes') stimuli, there are 68 possible skewed combinations. Each participant saw 4 of these during practice and the remaining 64 during the main experiment. The set of 4 used in practice was selected randomly for each participant. The results of this random selection did not co-vary with working memory capacity. Examples of each condition for both 'yes' and 'no' trials are provided in Table 4.

When only 'yes' trials are considered for digits 1–9, there are 7 possible *sequential* combinations (see Table

Table 4
Example stimuli for the Experiment 2 ordering task: examples are provided for each condition and each (correct) response type (i.e., 'Yes, in order' and 'No, not in order').

Sequential		Balanced		Skewed	
Yes	No	Yes	No	Yes	No
1 2 3	2 1 3	1 3 5	5 1 3	1 2 8	1 8 2
2 3 4	4 2 3	2 4 6	4 2 6	1 6 9	6 1 9
3 4 5	5 3 4	3 5 7	3 7 5	2 6 7	7 2 6
4 5 6	4 6 5	4 6 8	6 8 4	2 8 9	8 9 2
5 6 7	6 5 7	5 7 9	5 9 7	3 5 8	5 8 3
6 7 8	8 6 7	1 4 7	4 1 7	3 6 7	3 7 6
7 8 9	7 9 8	2 5 8	8 2 5	4 7 9	9 4 7
		3 6 9	6 9 3	5 6 8	5 8 6
		1 5 9	1 9 5	6 7 9	7 6 9

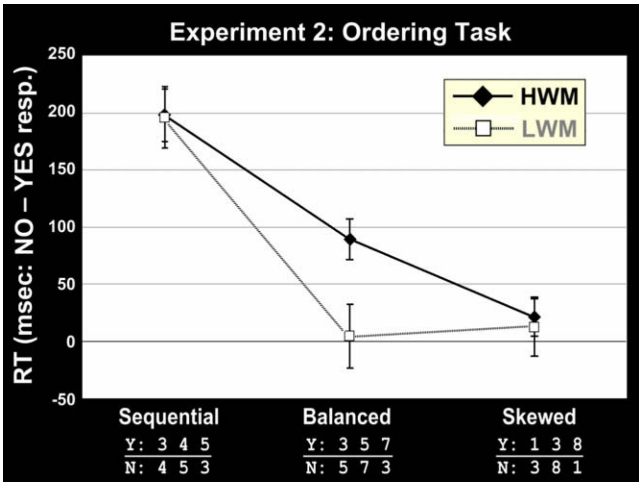

Fig. 5. This figure depicts cell means for RT data in the Experiment 2 ordering task. Means represent average yes-trial RTs subtracted from average no-trial RTs for each condition. Example stimuli (1 for each response type) are provided for each condition (see also Table 4 for additional examples). Error bars represent standard errors of the mean.

4). Each combination was repeated 4 times (total of 28 'yes' responses). There were also 28 'no' trials for the sequential condition. Note that 'no' trials did not repeat: for each set [e.g., {1, 2, 3}], four mixed-order combinations are possible [excluding the all-decreasing case: {1, 3, 2}, {2, 1, 3}, {2, 3, 1}, {3, 1, 2}]. There were 9 possible *balanced* combinations (Table 4). Again each combination was repeated 4 times (total of 36 'yes' and 36 non-repeating 'no' trials). Because an overall null effect was predicted for skewed stimuli (for the difference between 'yes' and 'no' trials, as well as across individual differences in WM), a greater number of skewed trials was used: 64 'yes' and 64 'no' non-repeating trials.

Participants began by performing 12 practice trials to familiarize themselves with the task. They then moved on to the main experiment where they performed 4 blocks of 64 trials each, with a short break in between each block. Each block contained 14 sequential, 18 balanced, and 32 skewed trials. An equal number of 'yes' and 'no' trials were included for each stimulus type in each block.

3.2. Results

We began by calculating difference scores for each participant for each stimulus type. For RTs, this was done by subtracting a participant's average response time for 'yes' trials from 'no' trials for that stimulus type (incorrect trials were excluded from the RT analysis). Thus, three difference scores were calculated for each participant. These difference scores were then entered into a WM × 3(type: sequential, balanced, skewed) ANCOVA, with WM as a between-subjects continuous factor and type as a within-subjects discrete factor.

With respect to RT, the main effect of WM was not significant [$F(1, 28) = 1.26$, $p = .272$]. In other words, the overall tendency to respond faster on 'yes' than 'no' trials did not vary as a function of WM capacity. However, there was a significant main effect of type [$F(2, 56) = 30.21$, $p < .001$]. *Sequential* trials showed a large positive difference between 'no' and 'yes' trials [$M = 198$ ms, $SE = 15$; one-sample, two-tailed t-test against 0 ms: $t(29) = 13.31$, $p < .001$]. *Balanced* trials showed a smaller but still significant effect of response [$M = 43$ ms, $SE = 16$; $t(29) = 2.68$, $p = .012$]. *Skewed* trials showed no difference overall between yes and no trials [$M = 8$ ms, $SE = 12$; $t(29) = .63$, $p = .532$].[2] Crucially, this pattern depended on WM capacity [WM × type interaction: $F(2, 27) = 9.40$, $p < .001$]. There was a strong positive correlation between WM and difference scores in the balanced condition [$r(28) = .542$, $p = .002$] but not between WM and difference scores in the other two conditions [sequential: $r(28) = -.049$, $p = .798$; skewed: $r(28) = -.056$, $p = .771$]. Cell means are depicted in terms of WM groups in Fig. 5.

To verify the hypothesized order-related differences between our chosen stimulus types and to further understand the specificity of the stimulus types' relation with WM, it is useful to consider the data in terms of raw yes and no RT scores (see Table 5).[3] If we take sequential trials as the extreme case where order information is strongly represented and skewed trials as the opposite extreme where order information is largely absent from participants' representations, then we would expect participants to be much faster to verify that sequential stimuli are in order relative to skewed stimuli. Conversely, we would expect participants to be far slower to correctly reject a sequential set that is out of order relative to an unordered skewed set. This is precisely what we found. Across participants, there was a significant difference in RTs between sequential and skewed

[2] The lack of response bias for Skewed stimuli serves as an important manipulation check. The ordering task employed here does not inevitably lead to a kind of confirmation bias, such that yes trials are always faster than no trials. Only when one is sensitive to inherent order information in the stimulus set is one faster to confirm that numbers are in the correct order.

[3] If, rather than using difference scores, Yes and No responses are treated as two levels of an additional factor Response, the crucial three-way interaction (WM × type × response) is equivalent to the two-way (WM × type) interaction already described. See also Table 5 for cell means.

Table 5
Response times (correct trials only) for Experiment 2. Means are given for 'yes' and 'no' trials separately. WM groups represent upper/lower thirds. Cells: mean *(SE)*.

	Sequential	Balanced	Skewed
Experiment 2: ordering task (response times)			
HWMs			
No	1057 *(48)*	981 *(40)*	955 *(40)*
Yes	861 *(34)*	892 *(13)*	945 *(33)*
LWMs			
No	1131 *(57)*	1015 *(48)*	1028 *(50)*
Yes	936 *(45)*	1010 *(51)*	1021 *(47)*

Table 6
Error rates (proportion incorrect) for Experiment 2. Means are given for 'yes' and 'no' trials separately. WM groups represent upper/lower thirds. Cells: mean *(SE)*.

	Sequential	Balanced	Skewed
Experiment 2: ordering task (prop. incorrect)			
HWMs			
No	.129 *(.032)*	.073 *(.025)*	.047 *(.017)*
Yes	.009 *(.018)*	.028 *(.012)*	.052 *(.016)*
LWMs			
No	.139 *(.022)*	.037 *(.011)*	.048 *(.008)*
Yes	.014 *(.008)*	.034 *(.011)*	.040 *(.007)*

conditions both when stimuli were correctly ordered ['yes' trials: $t(29) = -8.89$, $p < .001$] and were not correctly ordered ['no' trials: $t(29) = 13.31$, $p < .001$]. Crucially, this difference was in opposite directions for the two trial types (note the sign-change of the two *t*-statistics). For neither response type was this difference (sequential RTs–skewed RTs) related to WM capacity ['yes' trials: $r(28) = -.132$, $p = .487$; 'no' trials: $r(28) = .030$, $p = .876$].

We hypothesized that balanced stimuli should fall somewhere in between these two extreme cases, *but only for those higher in WM capacity*. That is, the higher one's WM, the more likely one should be to solve balanced problems by retrieving ordered representations of these stimuli (i.e., more like sequential stimuli). Those lower in WM, by contrast, should solve balanced trials more as though they were skewed trials. To test this, we looked at the difference between balanced and skewed RTs (balanced RTs–skewed RTs). If we are correct, this difference should be *negatively* related to WM for yes (ordered) trials (higher WM is related to a larger negative difference), and *positively* related to WM for no (non-ordered) trials (higher WM is related to a larger positive difference). This is exactly what we found ['yes' trials: $r(28) = -.411$, $p = .024$; 'no' trials: $r(28) = .429$, $p = .018$].[4] In short, HWMs appeared to respond to balanced trials more like sequential stimuli, and LWMs appeared to respond to balanced trials more like skewed stimuli. This finding is consistent with our central hypothesis that higher WM capacity is related to more complex representations of order information in symbolic numerical stimuli.

It is important to point out that 'yes' stimuli were repeated in the sequential and balanced conditions (1 instance of each stimulus per block) because of the limited number of possible stimulus combinations. This was not the case in the skewed condition. To assess whether stimulus repetition may have played a role in generating the results reported above, a WM $\times$ 3(type: sequential, balanced, skewed) $\times$ 4(block: 1, 2, 3, 4) ANCOVA was run. There were no main effects or interactions involving block [all $ps > .17$].

For response accuracy, data were analyzed as error rates (proportion incorrect). This was done for ease of interpretation, such that, after subtracting 'yes' error rates from 'no' error rates, response curves should be qualitatively similar to those in the RT analysis (i.e., with larger positive values for conditions where greater ordinal bias is present). As with the RT analysis, difference scores were entered into a WM $\times$ 3(type: sequential, balanced, skewed) ANCOVA. The main effect of WM did not reach significance ($F < 1$). As with the RT data, the main effect of type was significant [$F(2, 56) = 6.37$, $p = .003$], such that the sequential trials showed a large positive difference between 'no' and 'yes' trials ($M = .119$, $SE = .013$; $t(29) = 7.73$, $p < .001$), balanced trials less so ($M = .023$, $SE = .008$; $t(29) = 2.84$, $p = .008$), and skewed trials showed no difference overall ($M = .008$, $SE = .008$; $t(29) = 0.99$, $p = .328$). Though the WM $\times$ type interaction did not reach significance [$F(2, 56) = 2.43$, $p = .118$], a comparison of Tables 5 and 6 reveals that the overall pattern of results was similar to that seen for the RT data. There was a trend toward a positive correlation between WM and difference scores for the balanced stimuli [$r(28) = .270$, $p = .148$] but not between WM and difference scores for the other stimulus types [sequential: $r(28) = .082$, $p = .667$; skewed: $r(28) = -.099$, $p = .603$]. Given the similar patterns seen in the RT and accuracy data, results are unlikely to have been driven by a speed-accuracy trade-off.

3.3. Discussion

In Experiment 2, we hypothesized that the tendency seen in Experiment 1 for HWMs to go beyond simple associations and infer relative order information when acquiring novel numerical symbols should extend to deeper ordinal understanding in overlearned Arabic numerals. We found a bias in all participants to indicate sets of sequential numerals were in order, even when some of the numeral pairs were not (sequential stimuli). Thus, even LWMs do appear to establish strong ordinal associations, at least for stimuli in which those associations are directly related to an overlearned count sequence. This is consistent with our interpretation of the dot-array and Arabic numeral comparison data from Experiment 1 (Tables 2 and 3): higher WM capacity does not appear to be necessary for learning to emphasize ordinal associations in numerical stimuli *per se*.

Crucially, however, when ordinal relations between stimuli were patterned but less immediately obvious (balanced stimuli), only HWMs were faster and more accurate at recognizing balanced sets were in order relative to skewed sets. This latter result is also consistent with the

[4] Results point to the same conclusions if one considers the relation between Balanced and Sequential RTs rather than Balanced and Skewed RTs.

data from Experiment 1: the tendency to go beyond simple direct associations and infer deeper relative ordinal relationships between symbolic numerical stimuli depends on individual differences in WM capacity. That is, perhaps due to the tendency to adopt more demanding, cognitively complex learning strategies (Beilock & Decaro, 2007; De Neys, 2006; Decaro et al., 2008, 2009; Evans, 2003; Stanovich & West, 2000), HWMs are more likely to infer (Experiment 1) and retrieve (Experiment 2) more complex ordinal associations between symbolic numbers.

It should be noted that previous work has shown a tendency for HWMs to be more efficient at inhibiting unwanted associations (Conway & Engle, 1994). Thus, it may be that HWMs are more efficient than LWMs at inhibiting the ordinal associations between digits on 'no' trials. If this were the case, it would suggest that, if anything, we are underestimating the emphasis that HWMs place on ordinal information. This is because, if LWMs were not able to exercise inhibitory control to the same extent as HWMs, this would presumably result in a slowdown of LWMs' 'no' responses, thereby exacerbating the difference between 'no' minus 'yes' trials in all conditions. In this way, an inhibitory advantage for HWMs would simply work against our ability to detect a positive relation between WM capacity and difference scores for the balanced stimuli (imagine the balanced condition becoming more positive for LWMs – i.e., closer to HWMs – in Fig. 5). Moreover, whether LWMs struggle more so than HWMs with inhibiting automatic ordinal associations in the balanced condition may be irrelevant for interpreting LWMs' performance because LWMs appear to have weak ordinal associations in the context of balanced stimuli to begin with (indeed in the current data, the LWM split showed no such bias at all – Fig. 5).

To summarize, the data from Experiment 2 are consistent with Experiment 1 in showing that WM capacity is strongly related to a greater tendency to process symbolic numerical stimuli in terms of deeper ordinal relationships. Experiment 2 also expands upon Experiment 1 in showing that this tendency applies not only to the nature of inferences made in a symbol-learning context, but extends to explicit judgments made about complex ordinal relations in overlearned Arabic numerals as well.

4. General discussion

In two experiments we show that individual differences in WM capacity predict participants' predisposition to go beyond simple, direct associations and infer deeper ordinal relationships among symbolic numerical stimuli. In our first experiment, we show that this tendency allows for more accurate mapping of numerical content onto novel symbols. Experiment 2 reveals that this tendency extends to HWMs' sensitivity to more complex order information in overlearned symbolic stimuli (Arabic numerals). Because complex mathematical process often rely strongly on symbolic representations, understanding individual differences in how order information is represented may yield important insight into why HWMs typically outperform LWMs on a wide range of numerical tasks (e.g., Ashcraft & Krause, 2007).

In Experiment 1, we set out to examine whether higher WM capacity would facilitate more complex inferences about how numerical symbols were ordinally related to one another. In doing so, we simulated the process of mapping symbols onto approximate (pre-verbal) quantities that has been postulated by several prominent models of symbolic number representation (Dehaene & Changeaux, 1993; Gelman & Gallistel, 1978; Verguts & Fias, 2004). To this effect, the data from Experiment 1 may extend our understanding of this hypothesized mapping process in two important ways. First, we show that individual differences in a domain-general cognitive capacity – working memory – play an important role in determining participants' success in using these newly acquired symbols in a series of simple numerical tasks (i.e., higher WM was positively related to greater accuracy across the comparison and global ordering tasks). Second, we show that this advantage is due at least in part to HWMs' tendency to go beyond direct symbol-quantity associations and infer ordinal associations between the symbols themselves. We thus provide evidence that both domain-general cognitive processes and relative order information play key roles in determining successful numerical symbol-mapping.

The view that WM plays an important role in symbol-mapping is consistent with neural evidence showing involvement of prefrontal cortical structures in mapping quantity information onto abstract symbols, both at the neuronal level in monkeys (Diester & Nieder, 2007) and in terms of neuroimaging data in humans (Lyons & Ansari, 2009). Importantly, however, prefrontal activity does not by itself imply involvement of WM-related processes. By directly assessing the role of individual differences in WM capacity in the numerical symbol-mapping process, the current work serves as an important bridge between neural and behavioral evidence by showing that WM does indeed play a central role in numerical symbol-mapping. This implies that a comprehensive account of numerical symbol acquisition requires an understanding of the interaction between models of working memory and symbolic numerical representation.

A potentially interesting question is whether our results generalize to *any* set of inherently ordered stimuli (light wavelength, luminance, sound pitch, etc.). In our view, the results of Experiment 1 – that individual differences in WM shape the strategies individuals use to learn stimuli and the relations among these stimuli – is potentially wide ranging. Support for this idea comes from our dot-comparison data, where high and low WMs did not differ in their ability to discriminate non-symbolic quantities – that is, the impetus to emphasize order information seems strategy-based, and is not due to anything inherent to the stimuli themselves. In that sense, we would have no problem if using non-enumerable stimuli (such as luminosity) showed the same pattern of results, since both continuous and discrete magnitudes can be placed on a single axis that implies order. In the current study, we chose stimuli with enumerable magnitudes because learning the meaning of symbolic numbers is an important case where inferring relative order appears to be an important step in moving beyond simple symbol-magnitude associations – a step that might be advantageous for mathematics learning in general.

The fact that higher WM predicted better acquisition of ordinal relationships among symbols due (at least in part) to differences in strategy use may signal a point of pedagogical intervention. That is, if all individuals, regardless of WM capacity, are made explicitly aware of the importance of emphasizing order information in numerical symbols, LWMs may acquire numerical symbol-mappings as rapidly and accurately as HWMs. Similarly, the current research may inform developmental theories of numerical symbol acquisition by suggesting that inference of relative order information is an important component of this process (see also Lyons & Ansari, 2009; Tzelgov et al., 2000). However, we must urge strong caution in drawing too close an analogy between the training method we employed in Experiment 1 using adult participants and the developmental processes by which children actually map symbols onto numerical quantities. Nevertheless, recent work by Booth and Siegler (2008) has shown that increasing linearity in children's numerical estimates can lead to improvements in arithmetic ability. One possibility is that development of a linear representation of numerical magnitudes is facilitated by greater understanding of ordinal relations that bind the number line together (see Opfer & Siegler, 2007, for a similar suggestion). In this way, it seems reasonable to speculate that explicit instruction in strategies that emphasize associations *between* numerical symbols might facilitate linear reshaping of numerical representations in general.

Due to the novel and artificial nature of the training stimuli used in Experiment 1, it is important to show that WM relates to differences in numerical order processing in highly familiar stimuli that are of direct ecological relevance to participants. In Experiment 2, we found that only those higher in WM capacity showed a response bias for balanced stimuli. Our interpretation of these data is that the tendency to go beyond simple associations and infer more complex ordinal relationships (as was seen in Experiment 1) is something that applies to highly enculturated numbers as well. If HWMs are more likely to notice ordinal relationships, they may be more likely to encode these associations. Repeated encoding in this manner in turn may lead to more efficient recognition of these ordinal relationships. Though such an account is admittedly a speculative one at present, further work may elucidate the role that WM and understanding ordinal associations play in developing a full complement of symbolic numerical skills.

Interestingly, we found that all subjects, regardless of WM capacity, were biased to see sequential sets of numerals as in order, even when one or more numerals pairs was in fact switched. This finding is consistent with the view that direct ordinal associations between adjacent symbolic integers are indeed a central component of symbolic numerical representations. However, the lack of response bias for skewed trials across all participants suggests that, even for HWMs, there is a limit to the ordinal associations that are readily accessible. This suggests it is not the case that ordinal relations in all possible three-item combinations (at least for the numerals 1–9) are automatically accessed, even for overlearned numerical symbols in literate adults (see Tzelgov & Ganor-Stern, 2004, for related discussion). Instead, we found evidence for this sort of retrieval-based access to order information only in sequential stimuli (for all participants) and balanced stimuli for HWMs in particular.

In conclusion, we suggest a new and potentially important source of individual variability in processing symbolic numbers: the interaction between ordinal processing and working memory capacity. In two different contexts, we provide evidence that WM plays a critical role in processing order information in numerical symbols: those higher in WM are more likely to go beyond the actual quantity a symbol represents and focus on the symbol's relative numerical order – a fact that generalizes to include efficient retrieval of more complex ordinal relationships among Arabic numerals. In short, WM differences are an important component of symbolic numerical processing, especially when the relative numerical order of stimuli is relevant to the task at hand. An important next step – albeit outside the purview of the current paper – would be to examine whether differences in ordinal processing of symbolic numbers may serve as a yet untapped source of predicting mathematical skills. Because complex mathematical process often rely strongly on symbolic representations as well as working memory (e.g., Ashcraft & Krause, 2007), understanding such individual differences may provide important insight into how and why some excel in mathematics while others struggle. Thus, our findings may set the stage for a new line of research aimed at understanding the interplay between ordinal processing, working memory, and numeracy.

Acknowledgement

Research supported by NSF CAREER Grant DRL-0746970 to Sian Beilock.

References

Ansari, D. (2008). Effects of development and enculturation on number representation in the brain. *Nature Reviews Neuroscience, 9*(4), 278–291.

Ashcraft, M. H., & Krause, J. A. (2007). Working memory, math performance, and math anxiety. *Psychonomic Bulletin and Review, 14*(2), 243–248.

Beilock, S. L., & Decaro, M. S. (2007). From poor performance to success under stress: working memory, strategy selection, and mathematical problem solving under pressure. *Journal of Experimental Psychology: Learning, Memory and Cognition, 33*(6), 983–998.

Booth, J. L., & Siegler, R. S. (2008). Numerical magnitude representations influence arithmetic learning. *Child Development, 79*(4), 1016–1031.

Buckley, P. B., & Gillman, C. B. (1974). Comparisons of digits and dot patterns. *Journal of Experimental Psychology, 103*(6), 1131–1136.

Conway, A. R., & Engle, R. (1994). Working Memory and Retrieval: A resource-dependent inhibition model. *Journal of Experimental Psychology: General, 123*(4), 154–173.

Conway, A. R., 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 and Review, 12*(5), 769–786.

De Neys, W. (2006). Dual processing in reasoning: Two systems but one reasoner. *Psychological Science, 17*(5), 428–433.

DeCaro, M. S., Carlson, K. D., Thomas, R. D., & Beilock, S. L. (2009). When and how less is more: Reply to Tharp & Pickering. *Cognition, 111*(3), 391–403.

Decaro, M. S., Thomas, R. D., & Beilock, S. L. (2008). Individual differences in category learning: Sometimes less working memory capacity is better than more. *Cognition, 107*(1), 284–294.

Dehaene, S. (1997). *The number sense: How the mind creates mathematics*. New York: Oxford University Press.

Dehaene, S., & Changeaux, J. P. (1993). Development of elementary numerical abilities – A neuronal model. *Journal of Cognitive Neuroscience, 5*(4), 390–407.

Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology, 20*(3–6), 487–506.

Diester, I., & Nieder, A. (2007). Semantic associations between signs and numerical categories in the prefrontal cortex. *PLoS Biology, 5*(11), e294.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General, 128*(3), 309–331.

Evans, J. St. B. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences, 7*(10), 454–459.

Fias, W., Lammertyn, J., Caessens, B., & Orban, G. A. (2007). Processing of abstract ordinal knowledge in the horizontal segment of the intraparietal sulculs. *Journal of Neuroscience, 27*(33), 8952–8956.

Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition, 44*(1–2), 43–74.

Gelman, R., & Gallistel, C. R. (1978). *The child's understanding of number*. Cambridge, MA: Harvard University Press.

Gray, J. R., Chabris, C. F., & Braver, T. S. (2003). Neural mechanisms of general fluid intelligence. *Nature Neuroscience, 6*(3), 316–322.

Imbo, I., Vandierendonck, A., & Vergauwe, E. (2007). The role of working memory in carrying and borrowing. *Psychological Research, 71*(4), 467–483.

Jacob, S. N., & Nieder, A. (2008). The ABC of cardinal and ordinal number representations. *Trends in Cognitive Sciences, 12*(2), 41–43.

Jou, J. (2003). Multiple number and letter comparison: Directionality and accessibility in numeric and alphabetic memories. *The American Journal of Psychology, 116*(4), 543–579.

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin and Review, 9*(4), 637–671.

LeFevre, J.-A., DeStefano, D., Coleman, B., & Shanahan, T. (2004). Mathematical cognition and working memory. In J. I. Campbell (Ed.), *Handbook of mathematical of cognition*. New York: Psychology Press.

Leth-Steensen, C., & Marley, A. A. (2000). A model of response time effects in symbolic comparison. *Psychological Review, 107*(1), 62–100.

Logie, R. H., Gilhooly, K. J., & Wynn, V. (1994). Counting on working memory in arithmetic problem solving. *Memory and Cognition, 22*(4), 395–410.

Lyons, I. M., & Ansari, D. (2009). The cerebral basis of mapping nonsymbolic numerical quantities onto abstract symbols: An fMRI training study. *Journal of Cognitive Neuroscience, 21*(9), 1720–1735.

Miyake, A., & Shah, P. (1999). *Models of working memory: mechanisms of active maintenance and executive control*. New York: Cambridge University Press.

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgments of numerical inequality. *Nature, 215*(109), 1519–1520.

Nieder, A., & Merten, K. (2007). A labeled-line code for small and large numerosities in the monkey prefrontal cortex. *Journal of Neuroscience, 27*(22), 5986–5993.

Opfer, J. E., & Siegler, R. S. (2007). Representational change and children's numerical estimation. *Cognitive Psychology, 55*(3), 169–195.

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science, 306*(5695), 499–503.

Roggeman, C., Verguts, T., & Fias, W. (2006). Priming reveals differential coding of symbolic and non-symbolic quantities. *Cognition, 105*(2), 380–394.

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences, 23*(5), 645–726.

Tuholski, S. W., Engle, R. W., & Baylis, G. C. (2001). Individual differences in working memory capacity and enumeration. *Memory and Cognition, 29*(3), 484–492.

Turconi, E., Campbell, J. I., & Seron, X. (2006). Numerical order and quantity processing in number comparison. *Cognition, 98*(3), 273–285.

Tzelgov, J., & Ganor-Stern, D. (2004). Automaticity in processing ordinal information. In J. I. Campbell (Ed.), *Handbook of mathematical of cognition*. New York: Psychology Press.

Tzelgov, J., Yehene, V., Kotler, L., & Alon, A. (2000). Automatic comparison of artificial digits never compared: Learning linear ordering relations. *Journal of Experimental Psychology: Learning, Memory and Cognition, 26*(1), 103–120.

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods, 37*(3), 498–505.

Van Opstal, F., Verguts, T., Orban, G. A., & Fias, W. (2008). A hippocampal-parietal network for learning an ordered sequence. *NeuroImage, 40*(1), 333–341.

Verguts, T., & Fias, W. (2004). Representation of number in animals and humans: a neural model. *Journal of Cognitive Neuroscience, 16*(9), 1493–1504.

Verhaeghen, P., & Salthouse, T. A. (1997). Meta-analyses of age-cognition relations in adulthood: estimates of linear and nonlinear age effects and structural models. *Psychological Bulletin, 122*(3), 231–249.

Wager, T. D., & Smith, E. E. (2003). Neuroimaging studies of working memory: a meta-analysis. *Cognitive, Affective and Behavioral Neuroscience, 3*(4), 255–274.