Journal of Numerical Cognition
jnc.psychopen.eu | 2363-8761

Empirical Research

Verbal Working Memory Load Dissociates Common Indices of the Numerical Distance Effect: Implications for the Study of Numerical Cognition

Erin A. Maloney*[a], Nathaniel Barr[b], Evan F. Risko[c], Jonathan A. Fugelsang[c]

[a] School of Psychology, University of Ottawa, Ottawa, Canada. **[b]** Sheridan College, Oakville, Canada. **[c]** Department of Psychology, University of Waterloo, Waterloo, Canada.

Abstract

In four experiments, we explore the role that verbal WM plays in numerical comparison. Experiment 1 demonstrates that verbal WM load differentially impacts the two most common variants of numerical comparison tasks, evidenced by distinct modulation of the size of the numerical distance effect (NDE). Specifically, when comparing one Arabic digit to a standard, the size of the NDE increases as a function of increased verbal WM load; however, when comparing two simultaneously presented Arabic digits, the size of the NDE decreases (and here is eliminated) as a function of an increased verbal WM load. Experiment 2, using the same task structure but different stimuli (physical size judgments), provides support for the notion that this pattern of results is unique to tasks employing numerical stimuli. Experiment 3 demonstrates that the patterns observed in Experiment 1 are not an artifact of the stimulus pairs used. Experiment 4 provides evidence that the differing pattern of results observed between Experiment 1 and Experiment 2 are due to differences in stimuli (numerical vs. non-numerical) rather than to other differences between tasks. These results are discussed in terms of current theories of numerical comparison.

Keywords: numerical distance effect, working memory, mental number line, numerical representation

Journal of Numerical Cognition, 2019, Vol. 5(3), 337–357, https://doi.org/10.5964/jnc.v5i3.155

Received: 2017-11-23. Accepted: 2019-01-22. Published (VoR): 2019-12-20.

*Corresponding author at: 6022 Vanier Hall, 136 Jean Jacques Lussier, Ottawa, Ontario, Canada K1N 6N5. E-mail: erin.maloney@uottawa.ca

Working memory (WM) has been demonstrated to play a central role in various types of mathematical processing (Ashcraft & Kirk, 2001; LeFevre, DeStefano, Coleman, & Shanahan, 2005; Raghubar, Barnes, & Hecht, 2010). However, the vast majority of research in this area has focused on WM's role in the application of various mathematical operations, such as addition (e.g., Kalaman & LeFevre, 2007) and subtraction (Seyler, Kirk, & Ashcraft, 2003). A complementary, but less researched, question regards the role that WM plays in basic numerical processing. In the present investigation, we explore the role of verbal WM in one of the most fundamental of numerical processes - the comparison of numerical magnitudes.

Numerical Comparison and the Numerical Distance Effect

The ability to compare two numbers is a fundamental numerical ability that has been the subject of much empirical research (McCloskey, 1992). Indeed, McCloskey (1992) argued that the ability to select the larger of two quantities is *the* criterion for the understanding of numbers. While a great deal of research exists highlighting both the parallels and differences between comparing symbolic quantities and non-symbolic quantities, (e.g., Maloney, Risko, Preston, Ansari, & Fugelsang, 2010; Holloway & Ansari, 2008), the present paper focuses on the comparison of symbolic quantities. This ability to select the larger of two quantities appears to represent a fundamental building block of higher level mathematics as it is related to one’s overall math ability (Holloway & Ansari, 2008).

The most often studied phenomenon in numerical comparison is the numerical distance effect (NDE). When comparing two numerical quantities, participants are faster and more accurate at indicating which of two numbers is larger when the numerical distance separating the two numbers is relatively large (e.g., 2 vs. 9), compared to when it is small (e.g., 8 vs. 6; Dehaene, Dupoux, & Mehler, 1990; Moyer & Landauer, 1967). To date, numerical comparison has been predominantly assessed by two tasks. In a comparison to a standard task, participants are shown one number and are asked to indicate whether that number is higher or lower than a standard (e.g., 5). In a simultaneous presentation task, participants see two numbers presented at the same time and are asked to indicate the larger (or the smaller) of the numbers. These two tasks elicit NDEs that are comparable in magnitude and, as such, have been used interchangeably in the literature. Since its discovery, the NDE has taken a prominent place in studies on numerical cognition, both at the behavioral level (e.g., Dehaene, Dupoux, & Mehler, 1990), and at the neural level (e.g., Holloway & Ansari, 2008) as it is held to provide a critical clue as to how numbers are represented in the mind.

One of the most prominent accounts of the NDE in numerical comparison posits that the NDE is due to the placement of numbers along an internal mental number line (Restle, 1970). This theory presumes an analogue system in which numerical values correspond to spatial coordinates on a mental number line. The location of a number on the mental number line is not exact, but is represented as a distribution around the true location of that number on the line. Thus, the closer two numbers are on the number line, the more their representations will overlap. Numerical comparison in this model can be achieved by comparing the relative location of the activated representations on the mental number line (e.g., Pinel et al., 2001; Restle, 1970). According to this theory, as the distance between two digits decreases the difficulty of discriminating their location on the mental number line increases. This manifests as longer response times and lower accuracies for close digit pairs compared to far digit pairs. This explanation of the NDE is typically referred to as the *representational overlap view* and in its various forms represents the dominant theory of the NDE (e.g., Cohen Kadosh et al., 2005; Kaufmann & Nuerk, 2005; Piazza, Izard, Pinel, Le Bihan, & Dehaene, 2004; Turconi, Campbell, & Seron, 2006).

In contrast to the *representational overlap view*, Verguts, Fias, and Stevens (2005; see also Van Opstal, Gevers, De Moor, & Verguts, 2008) have stressed the role of stimulus response associations in generating the NDE. According to the Verguts et al. (2005) *model of exact small number representation* (which only addresses small numbers as other confounding factors arise with larger numbers; see Brysbaert, 1995), each digit is differentially associated with the responses “smaller than” and “larger than.” Within this model, as digit magnitude increases, the association between that digit and the “larger than” response increases and the association with the “smaller than” response decreases. Likewise, as digit magnitude decreases, the association between that digit and the “smaller than” response increases and the association with the “larger than” response decreases. When two numbers are presented, the model chooses, for example the larger number, via a process of response competition wherein, over time, the digit with the stronger association will be more likely to be selected. According to this model, the NDE results from a greater amount of response competition between two digits with similar strength of associations with the relevant response. For example, when presented with two numerically close numbers, both numbers activate the response “larger than” to a similar degree making it harder for the system to discriminate between these two numerosities. This increased response competition manifests in longer response times (RTs) and lower accuracies for close relative to far digit pairs.

More recently, a third account of the NDE has been put forth. Van Dijck and Fias (2011) suggest that during tasks that require the comparison of numbers, participants spontaneously make use of the inherent ordinal structure of the number system and systematically map the numbers onto an internal number line that is held in WM (this differs from the above account which suggests a mental number line held in long-term memory). This account suggests that it is from the fact that the positions in this temporary store are systematically associated to space, the NDE effect naturally follows.

The Numerical Distance Effect and Working Memory

Neither the representational overlap account nor the response competition account make explicit predictions about the potential contribution of verbal WM to the NDE. This is likely due in large part to the aforementioned preponderance of studies examining the role of WM in higher level mathematical processing rather than in lower level numerical processing. That said, evidence that WM is involved in other low-level numerical processes suggests that there is good reason to expect that WM may play an important role in numerical comparison. For example, Lyons and Beilock (2011) investigated the role that trait-level WM capacity plays in inferring ordinality from both novel symbols and Arabic digits. Participants that were defined as high WM capacity outperformed their low WM capacity counterparts both in learning ordinal information with novel symbols and in extracting ordinality information from sequential sets of Arabic digits. Although this study did not employ a numerical comparison task, ordinality may be an important part of numerical comparison (e.g., Restle, 1970). Furthermore, standard versions of comparison tasks require the temporary storage and manipulation of information (a hallmark of a task that requires WM; Baddeley & Hitch, 1974). The clearest example of this idea is in the comparison to a standard task. In this task, participants must maintain the standard (often the digit 5) in WM throughout the task. Lastly, evidence from research in math anxiety suggests that WM could contribute to numerical comparison. Specifically, in the context of a numerical comparison task, Maloney, Ansari, and Fugelsang (2011) demonstrated that math-anxious individuals, who are thought to suffer from a reduced verbal WM capacity when doing number-related tasks (e.g., Ashcraft & Kirk, 2001), had larger NDEs than their non-math-anxious peers on both the comparison to a standard task and the simultaneous comparison task.

Van Dijck and Fias’ (2011) account of the NDE does make specific claims about the role that verbal WM should play in the NDE. Specifically, given their theory that the mental number line is constructed in WM in response to the need to compare numbers within a task, verbal WM resources are thus required for the construction of the mental number line. Therefore, if one were to impose a verbal WM load, then we should see no (or a reduced) NDE arise. Herrera, Macizo, and Semenza (2008) presented evidence inconsistent with the prediction that WM is involved in numerical comparison. They had participants perform lower or higher than five judgments (comparison to a standard) under varying WM loads and demonstrated no effect of WM load on the NDE. However, while, the comparison to a standard task and a simultaneous presentation task both elicit NDEs of roughly the same magnitude, using these two tasks interchangeably may not be a good practice. Research by Maloney et al. (2010) demonstrates that these two tasks vary in terms of their split-half reliability and, more surprisingly, that a person’s NDE on one of these tasks is not predictive of the size of their NDE that arises in the other task, suggesting that these two tasks may be tapping into different representations or processes. Thus, the fact that Herrera et al. (2008) found no effect of WM on the comparison to a standard task does not mean that we should assume no effect of WM on the simultaneous presentation task. Against this background we set out to further investigate the role of verbal WM in numerical comparison using both the simultaneous comparison task and the comparison to a standard task.

Experiment 1

The aim of Experiment 1 was to examine the relation between WM and the NDE in two of the most popular tasks used to measure the NDE (i.e., simultaneous comparison and comparison to a standard). The motivation to use both simultaneous comparison and comparison to a standard, rather than the more typical approach of using only one, is based on research demonstrating that the NDEs in these tasks may not be measuring the same underlying process. Specifically, Maloney et al. (2010) demonstrated no correlation between the NDE across the simultaneous comparison and comparison to a standard tasks.

Thus, in Experiment 1, participants completed both the comparison to a standard task and the simultaneous presentation task with Arabic digit stimuli under two blocked WM load conditions (zero load and high load). The WM manipulation was verbal in nature and consisted of either 6 filler stimuli (in the zero load condition) or 6 letters (in the high load condition). We elected to use a verbal WM load as it is thought to be an integral component in at least one other basic numerical process (counting; Healy & Nairne, 1985; Nairne & Healy, 1983). We thus felt that the use of a verbal WM load would increase the probability of detecting a WM by NDE interaction if one does, in fact, exist. In the simultaneous presentation task, participants saw two simultaneously presented Arabic digits and identified the larger of the two. In the comparison to a standard task, participants were presented with one Arabic digit and identified whether the presented digit was larger than or smaller than the number 5.

Method

Participants

Fifty-three undergraduate students from a Canadian university participated and were granted experimental credit towards a course. Four participants failed to complete both tasks leaving forty-nine participants.

Apparatus, Stimuli and Procedure

The data were collected on a Pentium 4 PC computer running E-Prime 1.1 (Schneider, Eschman, & Zuccolotto, 2001). Stimuli were displayed on a 19” monitor. All stimuli were presented in 18 point Times New Roman font. Participants completed both tasks and tasks were presented in a counterbalanced order.

In the simultaneous comparison task, each trial began with a fixation point that remained on the screen for 500 ms. In the “zero” WM load condition, a display containing 6 hash marks (######) was presented for 3 seconds. In the “high” WM load condition a display containing 6 letters was presented for 3 seconds. Letters were chosen from the set of F, H, J, K, L, N, P, Q, R, S, T, Y as used in Unsworth, Redick, Heitz, Broadway, and Engle (2009). Participants were told that they could ignore the hash marks but would need to remember the letters and report them at the end of the trial. After 3 seconds, a display containing two Arabic digits was presented. Numbers ranged from 1-4 and from 6-9. The numerical distance between the stimuli was 1 or 4.[i] Participants were told to identify which of the two numbers was numerically larger by pressing the “A” key if the left number was larger, and the “L” key if the right number was larger. Following their judgment, participants were prompted to type in the letters presented at the beginning of the trial. There were two blocks each with 160 trials. The two WM blocks were counterbalanced across participants.

In the comparison to a standard task each trial began with a fixation point that remained on the screen for 500 ms. Participants were given the same working memory loads as in the simultaneous comparison task. After a 3-second display of the to-be-remembered letters, a display containing one Arabic digit was presented. Numbers were 1, 4, 6, or 9. Thus, the numerical distance between the stimuli and 5 was 1 or 4. Participants were told to identify whether the number was lower than 5 or higher than 5 by pressing the “A” key if the number was lower and the “L” key if the number was higher. Following their judgment, participants were prompted to type in the letters presented at the beginning of the trial. There were two blocks each with 160 trials. Stimuli composition was identical for both blocks (but appeared in a randomized order).

Results

Two participants were removed because their accuracy was below 20% on the WM task, meaning that they did not have enough trials to calculate a reliable score in those cells. This left us with 47 participants. RTs and errors were analyzed across participants with Task, Numerical Distance, and WM load as within-subject factors. Trials on which participants responded incorrectly to the WM load (12.0%) and trials on which there was an incorrect response in the number comparison task (3.5%) were removed prior to RT analysis. Remaining RTs were submitted to an outlier removal procedure using a 2.5 standard deviation cutoff in each cell resulting in the removal of an additional 3.8% of the data. Table 1 depicts the mean RT and errors as a function of Task, Numerical Distance, and WM load, and Figure 1 displays the Task, WM by Numerical Distance interaction in terms of the NDE.

Table 1

Response Times (in ms) as a Function of Working Memory Load and Numerical Distance (Percent Errors in Brackets) for Both Tasks in Experiment 1

	Comparison to a Standard		**Simultaneous Comparison**	
WM Load	**Distance 1**	**Distance 4**	**Distance 1**	**Distance 4**
Low Load	631 (4.1)	572 (2.9)	617 (2.1)	567 (3.6)
High Load	899 (4.0)	819 (2.1)	790 (3.5)	775 (5.0)

A 2 (Task: comparison to a standard vs. simultaneous comparison) x 2 (WM load: zero vs. high) x 2 (Numerical Distance: 1 vs. 4) ANOVA conducted on the RT data yielded no main effect of Task, *F*(1, 46) = 2.8, *MSE* = 60180.3, *p* = .10, η_p^2 = .06. There was a main effect of WM load, *F*(1, 46) = 81.4, *MSE* = 57874.4, *p* < .001, η_p^2 = .64, such that responses were slower under a high WM load, and a main effect of Numerical Distance, *F*(1, 46) = 113.6, *MSE* = 2150.0, *p* < .001, η_p^2 = .71, such that responses were faster for distances of 4 than for distances of 1. The Task x Load interaction was marginal, *F*(1, 46) = 4.03, *MSE* = 26442.7, *p* = .05, η_p^2 = .08, whereby the effect of load was larger in the comparison to a standard task than the simultaneous comparison task. There was also a Task x Numerical Distance interaction, *F*(1, 46) = 9.9, *MSE* = 3289.4, *p* = .003, η_p^2 = .18, such that the NDE was larger in the comparison to a standard task than the simultaneous comparison task.

There were no other significant two-way interactions. Importantly, there was also a Task x WM load x Numerical Distance interaction, *F*(1, 46) = 11.1, *MSE* = 1734.1, *p* = .002, η_p^2 = .19. The three-way interaction was the product of two WM load x Distance interactions that took opposite forms as a function of Task. Specifically, for the comparison to a standard task, the size of the NDE increased from the zero load (59 ms) to the high load (80 ms), whereas for the simultaneous comparison task, the size of the NDE decreased from the zero load (50 ms) to the high load (14 ms).

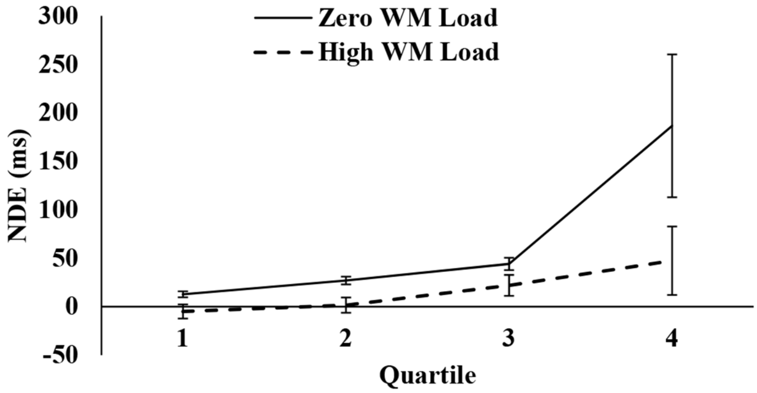

Figure 1. The numerical distance effect under zero and high working memory loads in Experiment 1 calculated as RTs at distance 1 minus RTs at distance 4. Error bars represent +/- 1 standard error of the mean.

A parallel analysis on the error data yielded a main effect of Numerical Distance, *F*(1, 46) = 18.1, *MSE* = 7.3, *p* < .001, η_p^2 = .28, whereby participants were more accurate in the Distance 4 condition. The interaction between Load and Distance was significant, *F*(1, 46) = 4.9, *MSE* = 7.4, *p* = .03, η_p^2 = .10. The pattern of means that make up these interactions qualitatively mirrored those in RT. There were no other significant main effects or interactions.

Discussion

In Experiment 1, we assessed the role of verbal WM in numerical comparison by examining the conjoint effects of verbal WM load and numerical distance in the context of two classic numerical comparison tasks – the comparison to a standard and simultaneous numerical comparison tasks. Critically, the influence of verbal WM load was qualitatively different across these two variants. Specifically, in the simultaneous comparison task we found an interaction between verbal WM load and the NDE whereby the size of the NDE was *smaller* under a high verbal WM load than under no verbal WM load. Conversely, in the comparison to a standard task the NDE was *larger* under a high verbal WM load than under no verbal WM load. This dissociation provides further evidence that the two tasks used most often in the field of numerical cognition in studies of the NDE may be indexing different underlying processes (i.e., they are differentially affected by verbal WM load). While Maloney et al. (2010) suggested that this may be the case (due to the lack of a correlation between the NDEs produced in each task), there has yet to be any strong experimental evidence to substantiate this claim before now. Furthermore, this task-driven dissociation is not currently predicted by any theory of numerical comparison. In a similar vein, another important result from Experiment 1 was the elimination of the NDE in the high load condition of the simultaneous comparison task. The elimination of the NDE is particularly interesting in the context of claims that the NDE is impervious to elimination (Dehaene, 2011). Indeed, the lack of an NDE while individuals compare two numbers is difficult to reconcile with the notion that the NDE results from how numbers are represented (information that presumably has to be accessed to complete the comparison). However, the elimination of the NDE under verbal WM load is consistent with the account put forth by van Dijck and Fias (2011) who claim that the mental number line is constructed in WM. The fact that a verbal WM load dissociates the NDEs that arise in these two comparison tasks, however, is not predicted by any model of the NDE.

The increase in the NDE as a function of increased verbal WM load in the comparison to a standard task is inconsistent with the findings of Herrera et al. (2008) who reported that the NDE in the comparison to a standard task was unaffected by increased WM loads. It is likely that the WM manipulation in the Herrera et al. (2008) experiment did not tax participant’s WM capacity to the same degree as that used here and as such they did not detect the interaction. One could argue that one way to account for the overadditive interaction between verbal WM and the NDE is to do so by recourse to a scaling argument. While there are cases whereby the size of an effect of interest does increase as overall RTs increase (as they do in the high verbal WM load condition here), this argument is rather weak for the current data as evidenced by the underadditive interaction in the simultaneous presentation task (i.e., slower RTs and a smaller NDE).

Experiment 1 has yielded a complex pattern of results relating WM load to the NDE in two different numerical comparison tasks. Specifically, increasing verbal WM load decreases the size of the NDE in the simultaneous presentation task, but increases the size of the NDE in the comparison to a standard task. The entire pattern of results is not predicted by any existing account of numerical comparison and demonstrates that the contribution of WM to numerical comparison is not in any way straightforward (at least provided current assumptions in the numerical cognition literature). This complex pattern thus provides a unique opportunity for novel insights into numerical comparison in general and the NDE in particular. In Experiment 2 we begin the attempt to decompose this complex pattern of results.

Experiment 2

Numerical comparison tasks consist of two independent components: the “numerical” aspect (i.e., the stimuli) and the “comparison” aspect. In Experiment 2 we set out to test the hypothesis that the dissociation between simultaneous presentation and comparison to a standard arises because of differences in the structure of the two comparison tasks, rather than to the numerical aspects of the task per se. Specifically, we wanted to know if the difference arises because of a differential contribution of verbal WM to comparing a stimulus to a standard and comparing two simultaneously presented stimuli. Thus we kept the task instructions the same but changed the to-be-compared stimuli. Specifically, rather than compare Arabic digits, participants compared squares which varied in physical size. Comparing the size of shapes has been shown to elicit a distance effect (Fulbright, Manson, Skudlarski, Lacadie, & Gore, 2003) and some researchers have even argued that the distance effect observed when comparing numbers is due to the same mechanism as the distance effect observed when comparing the physical size of objects (see Cohen Kadosh, Lammertyn, & Izard, 2008; Walsh, 2003). Through the use of these stimuli, we are able to directly assess whether the numerical components of the task are at all responsible for the pattern of the results. If it is only the nature of the tasks themselves (and not the stimuli), then one would predict that a similar pattern of results would emerge to that observed in Experiment 1.

Thus, in Experiment 2, participants performed both comparison tasks under zero and high verbal WM loads using single squares that varied in physical size. In the comparison to a standard task, participants were first shown a “standard” square which was smaller than half of the stimuli and larger than half of the stimuli. Participants identified whether each subsequently presented square was smaller than or larger than the standard. In the simultaneous square comparison task, participants saw two simultaneously presented squares and identified the larger of the two squares. The same verbal WM load conditions that were used in Experiment 1 were used here. If the dissociation observed in Experiment 1 arises because of the nature of the comparison task, then we should observe a similar dissociation between the two comparison tasks as a function of verbal WM load.

Method

Participants

Sixty undergraduate students from a Canadian university participated and were granted experimental credit towards a course. Seven participants were excluded because they failed to complete both tasks; 6 participants were removed because their accuracy was below 20% on the WM task. Finally, 4 participants were removed for having too low accuracy on the comparison tasks. These participants each had accuracies below chance in one cell of either comparison task. This left us with 43 participants. Excluding these participants did not change the overall pattern of the results.

Stimuli, Apparatus, and Procedure

The apparatus and program used were identical to that of Experiment 1 with minor exceptions. Rather than comparing Arabic digits, participants compared the physical size of squares. In one task participants judged whether a presented square was larger than a pre-determined standard, and in the other task judged which of two simultaneously presented squares were larger. In the comparison to a standard task the standard square was presented for 1000 ms before each WM load display to remind participants of its size. Participants completed both tasks and tasks were presented in a counterbalanced order.

The size of the squares used here was parametrically varied so as to create differences that paralleled the differences in the numeric tasks. There were 9 squares in total, the area of each ranged from 1 cm^2 to 9 cm^2 with a square mapping onto each of the intervening integers within that range (2 cm^2 to 8 cm^2). This set of squares was then used to create stimulus sets with physical differences in area of 1 (e.g., 3 cm^2 to 4 cm^2) and 4 (e.g., 5 cm^2 to 9 cm^2). Within the comparison to a standard task, a 5 cm^2 square was used as the standard, and the presented stimuli were either 1 cm^2, 4 cm^2, 6 cm^2, or 9 cm^2 and in the 2-item comparison task the full set was used. In both tasks there were two blocks each with 160 trials. Trial composition was identical for both blocks (but was presented in a randomized order).

Results

RTs and errors were analyzed across participants, with Physical Distance, WM load, and Task variant as within-subject factors. Trials on which participants responded incorrectly to the WM load (10.6% in the comparison to a standard task and 9.7% in the simultaneous presentation) or made an incorrect response (9.4% and 1.8%, respectively) were removed prior to RT data analysis. Remaining RTs were submitted to a 2.5 standard deviation outlier removal procedure resulting in the removal of an additional 3.0% and 2.9% of trials, respectively. Table 2 depicts mean RTs and error rates for each distance under each WM load for each task variant. Figure 2 depicts the DE under each WM load for each task variant.

Table 2

Response Times as a Function of Working Memory Load and Distance (Percent Errors in Brackets) for Both Tasks in Experiment 2

	Comparison to a Standard		**Simultaneous Comparison**	
WM Load	**Distance 1**	**Distance 4**	**Distance 1**	**Distance 4**
Low Load	1227 (13.5)	784 (4.5)	1024 (1.4)	739 (0.6)
High Load	1363 (16.7)	1044 (4.2)	1145 (2.5)	906 (2.8)

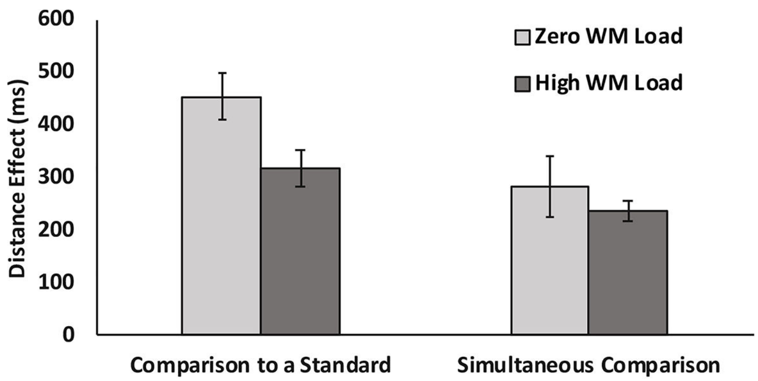

Figure 2. The distance effects under zero and high working memory loads for each task in Experiment 2. Error bars represent +/- 1 standard error of the mean.

A 2 (Task: comparison to a standard vs. 2-item comparison) x 2 (WM load: zero vs. high) x 2 (Physical Distance: 1 vs. 4) ANOVA conducted on the RT data yielded a main effect of Task, $F(1, 42) = 15.8$, $MSE = 123980.2$, $p < .001$, $\eta_p^2 = .27$, a main effect of WM load, $F(1, 42) = 24.9$, $MSE = 100947.8$, $p < 0.001$, $\eta_p^2 = .37$, and a main effect of Distance, $F(1, 42) = 164.8$, $MSE = 53970.1$, $p < .001$, $\eta_p^2 = .78$. There was a significant WM load x Distance interaction, $F(1, 42) = 4.2$, $MSE = 37581.7$, $p = .047$, $\eta_p^2 = .09$, whereby as WM load increased the NDE decreased. There was also a significant Task x Distance interaction, $F(1, 42) = 8.3$, $MSE = 36535.2$, $p < .01$, $\eta_p^2 = .16$, such that the size of the NDE was larger in the comparison to a standard task than in the 2-item comparison task. Most critically, the three-way interaction that was observed in Experiment 1 was not found here, $F(1, 42) = 1.0$, $MSE = 32100.5$, $p = .32$, $\eta_p^2 = .02$.[ii]

A parallel analysis on the error data yielded a main effect of Task, $F(1, 42) = 24.1$, $MSE = 221.1$, $p > .01$, $\eta_p^2 = .37$, a main effect of WM load, $F(1, 42) = 5.7$, $MSE = 37.8$, $p = .02$, $\eta_p^2 = .12$, a main effect of Numerical Distance, $F(1, 42) = 41.8$, $MSE = 62.9$, $p < .01$, $\eta_p^2 = .50$, and a Task x Numerical Distance interaction, $F(1, 42) = 50.0$, $MSE = 57.6$, $p < .01$, $\eta_p^2 = .49$, such that the size of the NDE was larger in the comparison to a standard task than in the 2-item comparison task. There was no Task x WM load x Numerical Distance Interaction, $F(1, 42) = 2.2$, $MSE = 47.8$, $p = .14$, $\eta_p^2 = .05$.

Discussion

In Experiment 2 we assessed whether the dissociation that was observed in Experiment 1 was driven by an inherent difference between the two comparison tasks by removing the numerical components of the tasks. The results of Experiment 2 are inconsistent with this idea. When squares were used as stimuli, the size of the distance effect decreased as a function of increased WM capacity in both of versions of the comparison tasks. If the act of comparing one item to a standard or comparing two simultaneously presented items made differential demands on WM, then we should have observed a similar pattern as that observed in Experiment 1. Specifically, we should have observed an increase in the distance effect in the comparison to a standard task and a decrease in the distance effect in the 2-item comparison task as a function of increased verbal WM load. This was not the case. Thus, it appears that the dissociation observed in Experiment 1 derives from an interaction between the particular comparison task and the to-be-compared stimuli (i.e., Arabic digits).

Experiment 3

While Experiment 2 indicated that the structure of the two tasks cannot account for the differential effect of verbal WM load, Experiment 3 served to test another hypothesis. Specifically, Experiment 3 served to test whether differences in the to-be-compared stimuli could account for the interaction that was seen in Experiment 1. In the comparison to a standard task that was used in Experiment 1, the digits 1, 4, 6, and 9 were compared to 5 (as a standard held in memory). Thus, pairs with numerical distance 1 and 4 were used. In the simultaneous comparison task from Experiment 1, the number pairs were generated from the numbers 1-4 and 6-9, and only those among them and with distances 1 and 4 were included in the stimulus set. It follows that the specific pairs used were different in the two tasks. In Experiment 3 we tested the hypothesis that the differential effect of verbal WM load on the distance effects that arise in the comparison to a standard task and the simultaneous presentation task arises due to the difference in stimuli between the two tasks. To test this hypothesis, we conducted a version of the simultaneous comparison task in which we paired the numbers 1, 4, 6, and 9 with the number 5. That is, the number 5 was present on every trial, making the number pairings identical to those used in the comparison to a standard task. Importantly, if the specific stimulus set that was used in the comparison to a standard task from Experiment 1 was the reason for the overadditive interaction with verbal WM load, then we should also expect to see an increased NDE in the simultaneous comparison task with these stimuli. However, if it was not the specific number pairings but rather something about the nature of comparing two simultaneously presented Arabic digits, then we should expect a decrease in the NDE as a function of increased verbal WM load despite these specific number pairings.

Method

Participants

Forty-five undergraduate students from a Canadian university participated and were granted experimental credit towards a course. One participant was removed for failing to complete the task.

Stimuli, Apparatus, and Procedure

The apparatus and program used were identical to the simultaneous comparison task that was used in Experiment 1. However, unlike in Experiment 1, here the number pairs were composed of the numbers 1, 4, 6, and 9 paired with the number 5. That is, the number 5 was present on every trial. On half of the trials the number 5 appeared to the left of the fixation and on half it appeared to the right.

Results

RTs and errors were analyzed across participants, with Numerical Distance and WM load as within-subject factors. Trials on which participants responded incorrectly to the WM load (13.4%) or made an incorrect response (3.8%) were removed prior to RT data analysis. Remaining RTs were submitted to a 2.5 standard deviation outlier removal procedure resulting in the removal of an additional 3.0% of trials. An additional 2 participants were removed for having NDEs that were greater than 2.5 standard deviations from the mean. Table 3 depicts mean RTs and error rates for each distance under each WM load. Figure 3 depicts the NDE under each WM load.

Table 3

Response Times as a Function of Working Memory Load and Distance (Percent Errors in Brackets) in Experiment 3

WM Load	Distance 1	Distance 4
Low Load	683 (3.0)	634 (2.2)
High Load	903 (3.6)	891 (3.4)

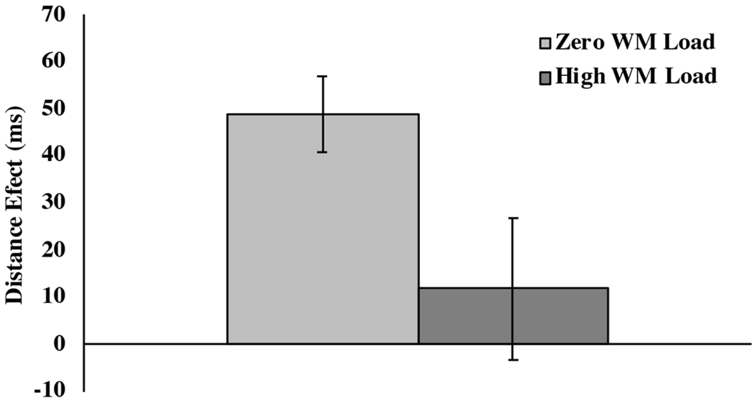

Figure 3. The numerical distance effect under zero and high working memory loads in Experiment 3 calculated as RTs at distance 1 minus RTs at distance 4. Error bars represent +/- 1 standard error of the mean.

A 2 (WM load: zero vs. high) x 2 (Numerical Distance: 1 vs. 4) ANOVA conducted on the RT data yielded a main effect of WM load, $F(1, 41) = 49.71$, $MSE = 47835$, $p < .001$, $\eta_p^2 = .548$, such that response times were quicker in the zero WM load condition (659 ms) than in the high WM load condition (897 ms). There was also a main effect of Numerical Distance, $F(1, 41) = 13.1$, $MSE = 2983$, $p < .001$, $\eta_p^2 = .242$, such that response times were slower at a Numerical Distance of 1 (793 ms) than at a Numerical Distance of 4 (762 ms). Importantly, there was also a WM load x Numerical Distance interaction, $F(1, 41) = 4.62$, $MSE = 3116$, $p = .038$, $\eta_p^2 = .101$. The form of this interaction was such that the effect of Numerical Distance was larger in the zero WM load (49 ms) than in the high WM load (12 ms). In fact, a one-sample T Test conducted on the RT data in the high load condition yielded no effect of Numerical Distance, $t < 1$, while there was a significant effect of numerical distance in the zero load condition, $t(41) = 6.1$, $p < .001$. Parallel analyses on the error data yielded no main effects and no significant WM load x Numerical Distance interaction.

Discussion

Experiment 3 served to test the account that the dissociation that was observed in Experiment 1 resulted from the exact number pairings that were used. Specifically, in Experiment 1, because there were a greater number of possible number pairs in the simultaneous comparison task than in the comparison to a standard task, it is possible that the stimuli themselves could have caused the observed pattern. Thus, in Experiment 3 we ran the simultaneous comparison task using the same number pairings that were used in the comparison to a standard task in Experiment 1. If it was the specific stimuli that were used that was causing the overadditivity in the comparison to a standard task (relative to the underadditivity that is observed in the simultaneous comparison task) then we would expect to see overadditivity in Experiment 3. We did not. Rather, the pattern that was observed in the simultaneous comparison variant used in Experiment 3 replicated that observed for the same variant of the task in Experiment 1. As such, we can conclude that the differential effect of verbal WM load on these two task variants is not dependent upon the specific number pairing used (at least within the 1-9 range of numbers).

Experiment 4

The results from Experiment 2 supported the idea that the structure of the two tasks cannot, in and of itself, account for the differential effect of verbal WM load on the NDE produced under the simultaneous comparison and the comparison to a standard tasks. However, in the comparison to a standard condition in Experiment 2, the standard square, to which all other squares were to be compared, was presented at the beginning of each trial and was thus not required to be maintained in WM. This was done because pilot testing indicated that participants were unable to hold the exact size of the square in WM. In contrast, in the parallel task in Experiment 1, participants were told at the beginning of the task that they were to compare each subsequently presented number to the number 5 and were not reminded of the standard (5) on each trial. To address this confound, in Experiment 4, participants again were told to decide whether the numbers presented were higher than or lower than the number 5. Importantly, the number 5 was presented at the beginning of every trial to exactly parallel what was done in Experiment 2. If the presentation of the standard stimuli on each trial is what caused the NDE to decrease under increased verbal WM load in Experiment 2, then in Experiment 4, we should also expect the NDE to decrease under increased verbal WM load. However, if this confound between Experiments 1 and 2 cannot explain the underadditivity in Experiment 2, then we should not expect underadditivity of the NDE here in Experiment 4.

Method

Participants

Seventy-eight undergraduate students from a Canadian university participated and were granted experimental credit towards a course. One participant was removed as they answered fewer than 20% of the WM trials correctly.

Stimuli, Apparatus, and Procedure

The apparatus and program used were identical to the comparison to a standard task that was used in Experiment 1 with the exception that, unlike in Experiment 1, here the number 5 was presented at the beginning of each trial for 1000ms to parallel the design of Experiment 2.

Results

RTs and errors were analyzed across participants, with Numerical Distance and WM load as within-subject factors. Trials on which participants responded incorrectly to the WM load (11.8%) or made an incorrect response (0.7%) were removed prior to RT data analysis. Remaining RTs were submitted to a 2.5 standard deviation outlier removal procedure resulting in the removal of an additional 2.9% of trials. Table 4 depicts mean RTs and error rates for each distance under each WM load. Figure 4 depicts the NDE under each WM load.

Table 4

Response Times as a Function of Working Memory Load and Distance (Percent Errors in Brackets) in Experiment 4

	Distance			
WM Load	**1**	**2**	**3**	**4**
Low Load	742 (1.4)	724 (0.9)	697 (0.8)	695 (0.9)
High Load	935 (0.7)	938 (0.4)	899 (0.3)	895 (0.3)

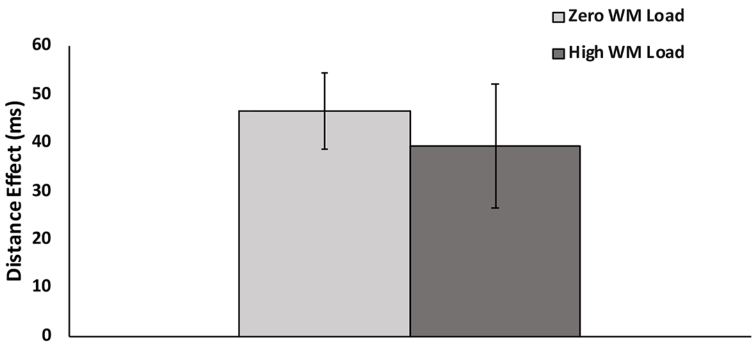

Figure 4. The numerical distance effect under zero and high working memory loads in Experiment 4 calculated as RTs at distance 1 minus RTs at distance 4. Error bars represent +/- 1 standard error of the mean.

A 2 (WM load: zero vs. high) x 2 (Numerical Distance: 1 through 4) ANOVA conducted on the RT data yielded a main effect of WM load, $F(1, 76) = 101.4$, $MSE = 62102.4$, $p < .001$, $\eta_p^2 = .572$, such that response times were quicker in the zero WM load condition (714 ms) than in the high WM load condition (917 ms). There was also a main effect of Numerical Distance, $F(3, 228) = 16.1$, $MSE = 4668$, $p = .001$, $\eta_p^2 = .175$, such that response times increased as numerical distance decreased. Importantly, there was no WM load x Numerical Distance interaction, $F(3, 228) = 0.79$, $MSE = 3427$, $p > .05$, $\eta_p^2 = .010$. Parallel analyses on the error data yielded a main effect of WM load, $F(1, 76) = 7.24$, $MSE = 15.4$, $p = .009$, $\eta_p^2 = .087$, such that more errors were made in the zero load condition (1.0%) than under the high load condition (0.6%). There were no other main effects and no significant WM load x Numerical Distance interaction.

Discussion

Experiment 4 served to test whether the decrease in the DE under increased WM load that was observed in Experiment 2 could be explained by the fact that, in Experiment 2, the standard was presented on each trial. Here, when the standard (5) was presented on each trial, the NDE did not decrease under increased WM load. Therefore, we can conclude that the difference in the patterns observed on the comparison to a standard task in Experiments 1 and 2 (i.e., an overadditive interaction with numerical stimuli and an underadditive interaction with squares) cannot be explained by the presentation of the standard stimulus at the beginning of each trial in Experiment 2 but not in Experiment 1. In Experiment 4, when the standard (5) was presented on each trial, the NDE neither increased nor decreased under increased verbal WM load. This pattern is further unpacked in the General Discussion.

General Discussion

The present investigation of the involvement of verbal WM in numerical comparison has yielded a number of important insights. The results from four experiments converge strongly on the notion that verbal WM is involved in symbolic numerical comparison, one of the most basic numerical abilities (McCloskey, 1992). In addition, the effect of verbal WM load on numerical comparison is dependent upon the nature of the specific comparison task used. In simultaneous presentation, the NDE decreases in magnitude with increases in WM load and in comparison to a standard, the NDE increases with increases in WM load. Critically, this dissociation only held when the comparison tasks involved numbers. When comparing physical size (using squares of different sizes as stimuli) the size of the distance effect decreased with increases in WM load in both the comparison to a standard and simultaneous comparison tasks. Clearly, this pattern is complex; however, we think it provides a unique opportunity for novel insights into numerical comparison in general and the NDE in particular. We explore some of these insights below.

The Elimination of the NDE in Simultaneous Presentation

The interaction between the NDE and WM load provides important new challenges for theories of the distance effect, specifically, and numerical representation in general. Here we have demonstrated in two experiments that in the simultaneous presentation task, an increased WM load leads to the elimination of the NDE. There are at least two theoretically distinct ways that one could explain the elimination of the NDE under a high verbal WM load. The first is a disruption account, which explains the elimination by recourse to the idea that the mechanism that gives rise to the NDE is in some fashion disrupted under a verbal WM load and thus there is no NDE produced (for a similar argument with the SNARC effect see Herrera et al., 2008; van Dijck, Gevers, & Fias, 2009). For example, if access to a noisy representation on a mental number line that is held in long term memory produces the NDE (i.e., the representational overlap account; e.g., Cohen Kadosh et al., 2005; Kaufmann & Nuerk, 2005; Piazza, Izard, Pinel, Le Bihan, & Dehaene, 2004; Turconi, Campbell, & Seron, 2006), then the present results suggest that maintaining a verbal WM load interferes with access to that representation. In other words, the verbal WM load could be preventing either the mapping of the input (Arabic digit) to a representation or the access of said representation from unfolding in the normal manner. Accordingly, the account by van Dijck and Fias (2011) would suggest that the increased verbal WM load prevents participants from actually constructing a mental number line in WM and, as such, no NDE arises under high verbal WM load.

The second account of the elimination of the NDE under high verbal WM load in the simultaneous comparison task is that the NDE is absorbed into slack created by the need to maintain the verbal WM load (Besner & Risko, 2005; Pashler, 1994; Risko & Besner, 2008). In this account, the mechanism that gives rise to the NDE can operate in parallel to the mechanism responsible for the maintenance of the verbal WM load and, in the delay caused by the latter, the interference can be resolved. Critically, this account requires that the mechanism that gives rise to the NDE is one that is capable of reducing with the passage of time (Pashler, 1994). With respect to both the overlapping representations account and Verguts et al.'s (2005) account of numerical comparison, amendments would have to be made that would allow the amelioration of the NDE over short periods of time (e.g., a second). In addition, positing that the elimination of the NDE is due to absorption is actually akin to positing that verbal WM is not involved in numerical comparison in the context of a simultaneous presentation.

One strategy for distinguishing between the disruption and absorption accounts of the elimination of the NDE in the simultaneous presentation condition is to use the time-course of the NDE. According to an absorption account, the NDE should decrease in magnitude as overall RT increases (for the basis of this logic see work on the Simon effect; De Jong et al., 1994; Risko & Besner, 2008; Rubichi et al., 1997). Thus, in the present data, we can look at the time course of the NDE and assess whether it decreases as RTs increase (as predicted by the absorption account). Evidence for this pattern should be present in both the low and high WM load conditions. Alternatively, according to the disruption account outlined above, the NDE should be absent throughout the response time distribution in the high WM load condition (i.e., disruption of access to the representations that give rise to the NDE should be independent of response speed). To look at this, we examined the simultaneous presentation data from Experiment 1 and split the RT data into quartiles (by subject and condition; see Figure 5). A repeated-measures ANOVA was conducted on data from each WM load condition.

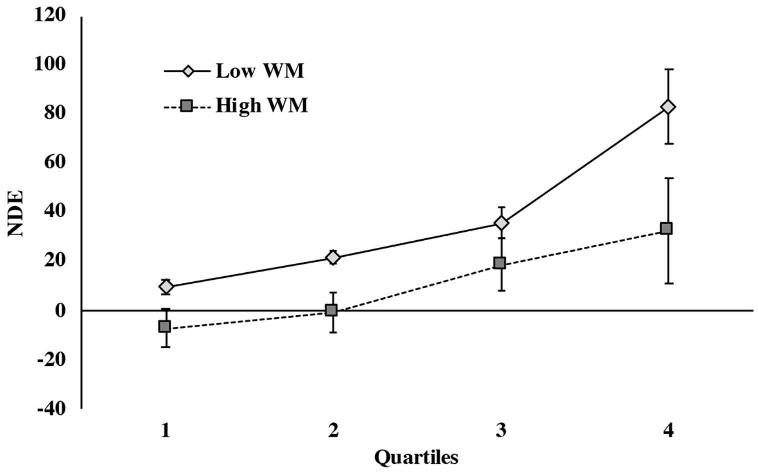

Figure 5. The numerical distance effects under zero and high working memory loads as a function of overall RTs divided into quartiles.

Results of this analysis indicate that in the zero WM load condition there was a main effect of Distance, $F(1, 46) = 11.8$, $MSE = 36220$, $p = .001$, $\eta_p^2 = .21$, and a Distance x Quartile interaction, $F(3, 138) = 4.9$, $MSE = 30929$, $p < .05$, $\eta_p^2 = .10$, such that the effect of Distance increased as RTs increased. Conversely, in the high WM load condition there was no main effect of Distance, $F(1, 46) = 1.9$, $MSE = 13563$, $p > .05$, $\eta_p^2 = .04$, and there was no Distance x Quartile interaction, $F(3, 138) = 1.8$, $MSE = 7083$, $p > .05$, $\eta_p^2 = .04$. These data suggest that the NDE is not being absorbed into the cognitive slack created by the increased WM load. Alternatively, according to the disruption account outlined above, the NDE should be absent throughout the response time distribution in the high WM load condition (i.e., disruption of access to the representations that give rise to the NDE should be independent of response speed). While clarifying the explanation of the elimination of the NDE in the simultaneous presentation condition, the aforementioned analysis leaves open the reason for the dissociation between the two different comparison tasks.

One interesting issue that arises when considering the idea that a high verbal WM load disrupts the processes that produce the NDE is that participants could still accurately compare the two numbers despite this disruption. Specifically, existing accounts of numerical comparison do not offer a mechanism by which numerical comparison can occur without producing an NDE. Thus, if one accepts a disruption account as the explanation for the WM-induced reduction in the NDE, then a complete theory of numerical comparison would require the existence of a separate type of representation or mode of access that could support numerical comparison but not produce a distance effect. This is because participants in the present study are able to perform the numerical comparison task under high verbal load without demonstrating an NDE. One potential explanation could be based on the fact that in Experiments 1, 3, and 4 we used single digits 1-9. Perhaps in addition to storing numerical information on a mental number line we also have a mental store of basic “number facts” for numbers 1 through 9, such as “9 is larger than 5”. Akin to performing basic addition (e.g., 2+2=4) through direct retrieval, accessing these facts would presumably be independent of WM. In addition, these “number facts” need not have stored with them magnitude information per se, but rather semantic knowledge that “9 is larger than 5”, thus, when comparison is completed via recourse to these “number facts” an NDE would not arise.

NDEs in Two Classic Tasks Dissociated

Unlike the NDE in the simultaneous presentation task, the NDE in the comparison to a standard task increases with increased WM load. The dissociation reported here is particularly surprising given that the literature assumes that these two tasks are tapping the same underlying mechanisms (e.g. Van Opstal et al., 2008). This notion is nonetheless consistent with Maloney et al.’s (2010) observation that the NDEs that arise in the simultaneous presentation and the comparison to a standard task do not correlate with one another. Experiment 2 demonstrates that this pattern is not purely a result of the structure of tasks. Specifically, a version of the simultaneous presentation and comparison to a standard that used non-numerical stimuli did not show different patterns as a function of increases in WM load. Experiment 3 suggests that the same pattern emerges when the numerical stimuli used is exactly the same across tasks.

An alternative conceptualization of the dissociation between simultaneous presentation and comparison to a standard with numeric stimuli is that the tasks encourage different comparison strategies. For example, Todkill and Humphreys (1994), outside of the numerical comparison context, suggested that participants could adopt either a long-term comparison or short-term comparison strategy in successive comparison tasks. The critical difference between these strategies concerns where the “comparator” is stored – long-term memory or short-term memory. In the simultaneous presentation condition, where both to-be-compared digits are presented on the screen at the same time, it seems safe to assume that the digits are compared in short-term memory. In a similar vein, the non-numerical simultaneous presentation task used in Experiment 2 would also seem to encourage a short-term comparison strategy. However, the instructions for the comparison to a standard task would seem to encourage a more long-term comparison strategy wherein on each trial participants retrieve the comparison digit (5) from long-term memory in order to carry out the comparison. Interestingly, this seems less true of the comparison to a standard with non-numerical stimuli used in Experiment 2. In that task, the standard was presented on each trial (this was not the case in the numeric version used in Experiment 1). Indeed, pilot testing suggested that it would be difficult to perform the non-numeric comparison to a standard task without providing such support (note that performing the same task with a meaningful digit as the standard stored in memory is relatively easy). Thus, both simultaneous presentation tasks and the non-numeric comparison to a standard could all be viewed as encouraging a short-term comparison strategy whereas the numeric comparison to a standard could be viewed as encouraging a long-term comparison strategy. Critically, this pattern maps qualitatively onto the pattern of the WM by distance effect interactions across tasks. Namely, the three tasks that are putatively relying on a short-term comparison strategy all show WM by distance effect interactions where the distance effect decreases with increasing WM load, whereas the one task that putatively relies on a long-term comparison strategy shows a WM by NDE interaction where the distance effect increases with increasing WM load.

In Experiment 4, participants completed a comparison to a standard task, however, the standard (5) was presented at the beginning of each trial. In this novel version of the task, the NDE did not change as a function of the increase in verbal WM load. While this pattern may, at first, seem counterintuitive, it is, in fact, consistent with the aforementioned theory. Specifically, by presenting the standard (5) at the beginning of every trial, participants are able to either perform the comparison in short-term memory or to adopt a more long-term comparison strategy wherein on each trial they retrieve the comparison digit (5) from long-term memory and carry out the comparison. If participants are sometimes making the comparison in short-term memory (which would yield a decrease in the NDE under high WM load) and sometimes making the comparison on long-term memory (which would yield an increase in the NDE under high verbal WM load), then these two effects would cancel each other out, resulting in an additive interaction. Indeed, this is what we see in Experiment 4.

The present results suggest not only that participants may be using different comparison strategies across tasks, but that those strategies are also differentially reliant on WM. We have already discussed the potential mechanisms mediating the WM by distance effect interactions in the simultaneous presentation tasks. With respect to the long-term comparison strategy, the distance effect may be arising, in part, during the retrieval of the standard from long-term memory. For example, on presentation of the probe digit, it would be harder to recall 5 when the probe is near (e.g., 6) than far (e.g., 9) from that standard. The interaction with WM can be understood from the standpoint of Unsworth and Engle's (2007) dual component theory of WM. In this theory, WM is involved in both active maintenance and controlled search from long-term memory (Unsworth et al., 2013; Unsworth & Engle, 2007). Thus, the increased WM load could impact the efficiency with which the standard is retrieved from long-term memory and because this is the same stage at which the NDE might arise (at least in part), the result is an overadditive interaction between WM load and the NDE. While speculative, this account does capture the complex pattern relating WM load to the NDE across two classic tasks, and, in so doing, provides a starting point for future investigations of this interaction.

Conclusion

We have assessed the role of verbal WM in numerical comparison by examining the effects of verbal WM load on the NDE in two classic numerical cognition tasks. We demonstrated that verbal WM does interact with the NDE and that the nature of the interaction is dependent upon the comparison task used. When comparing two simultaneously presented Arabic digits, the size of the NDE decreases (and here is eliminated) as a function of an increased verbal WM load. On the other hand, when comparing one Arabic digit to a standard, the size of the NDE increases as a function of increased WM load. We further demonstrated that this dissociation is not due to the task structure alone given that when both tasks are performed with squares as stimuli, the distance effect decreases in both tasks. Importantly, we present a novel account of how this dissociation of the NDE can be explained if participants adopt a long-term comparison strategy on the comparison to a standard task and a short-term comparison strategy on the simultaneous comparison task. This complex pattern has straightforward implications: (1) WM is intimately tied up with numerical comparison, (2) the NDE in simultaneous presentation and comparison to a standard are not equivalent, and (3) the NDE can be eliminated. These insights promise to generate further research.

Notes

i) Due to a programming error in the zero WM load condition there were three trials on which the distance was 2 where it should have been 4. These trials were removed from analyses.

ii) Note that we have opted to remove 6 participants because their accuracy was below 20% on the WM task and 4 participants were removed for having too low accuracy on the comparison tasks. If we did not remove those who had too low accuracy on the comparison task (but still remove the 5/6 participants who scored below 20% on the WM task as there were cells in which they had no analyzable data), then the pattern of results does not change. Indeed, even when including these five additional participants, the critical three-way interaction that was observed in Experiment 1 is still not found here, $F(1, 47) = 0.3$, $MSE = 62581.3$, $p > .05$, $\eta_p^2 = .01$.

Funding

This research was supported by Discovery Grants from the Natural Sciences and Engineering Research Council of Canada (NSERC) to EAM, EFR, and JAF.

Competing Interests

The authors have declared that no competing interests exist.

Acknowledgments

The authors have no support to report.

Data Availability

The lead author does not have authorisation from their ethics board to release the data publicly. However, should readers wish to see the data, they should contact Erin Maloney (Erin.Maloney@uottawa.ca). The requester simply needs to complete a brief data sharing agreement before the data can be accessed.

References

Ashcraft, M. H., & Kirk, E. P. (2001). The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology: General, 130*, 224-237. https://doi.org/10.1037/0096-3445.130.2.224

Baddeley, A., & Hitch, G. (1974). Working memory. In G. Bower (Ed.), *The psychology of learning and motivation* (pp. 47–89). New York, NY, USA: Academic Press.

Besner, D., & Risko, E. F. (2005). Stimulus-response compatible orienting and the effect of an action not taken: Perception delayed is automaticity denied. *Psychonomic Bulletin & Review, 12*, 271-275. https://doi.org/10.3758/BF03196371

Brysbaert, M. (1995). Arabic number reading: On the nature of the numerical scale and the origin of phonological recoding. *Journal of Experimental Psychology: General, 124*, 434-452. https://doi.org/10.1037/0096-3445.124.4.434

Cohen Kadosh, R., Henik, A., Rubintsen, O., Mohr, H., Dori, H., Van Deven, V., . . . Linden, D. E. J. (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*(2), 132-147. https://doi.org/10.1016/j.pneurobio.2007.11.001

Dehaene, S., Dupoux, E., & Mehler, J. (1990). Is numerical comparison digital: Analogical and symbolic effects in two-digit number comparison. *Cognition, 16*, 626-641.

Dehaene, S. (2011). *The number sense: How the mind creates mathematics.* New York, NY, USA: Oxford University Press.

De Jong, R., Liang, C.-C., & Lauber, E. (1994). Conditional and unconditional automaticity: A dual-process model of effects of spatial stimulus-response correspondence. *Journal of Experimental Psychology: Human Perception and Performance, 20*(4), 731-750. https://doi.org/10.1037/0096-1523.20.4.731

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*(2), 193-200.

Healy, A. F., & Nairne, J. S. (1985). Short-term memory processes in counting. *Cognitive Psychology, 17*, 417-444. https://doi.org/10.1016/0010-0285(85)90015-5

Herrera, A., Macizo, P., & Semenza, C. (2008). The role of working memory in the association between number magnitude and space. *Acta Psychologica, 128*, 225-237. https://doi.org/10.1016/j.actpsy.2008.01.002

Holloway, I. D., & Ansari, D. (2008). Domain-specific and domain-general changes in children's development of number comparison. *Developmental Science, 11*, 644-649. https://doi.org/10.1111/j.1467-7687.2008.00712.x

Kalaman, D. A., & LeFevre, J.-A. (2007). Working memory demands of exact and approximate addition. *The European Journal of Cognitive Psychology, 19*, 187-212. https://doi.org/10.1080/09541440600713445

Kaufmann, L., & Nuerk, H. C. (2005). Numerical development: Current issues and future perspectives. *Psychological Science, 47*, 142-170.

LeFevre, J., DeStefano, D., Coleman, B., & Shanahan, T. (2005). Mathematical cognition and working memory. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 361-377). New York, NY, USA: Psychology Press.

Lyons, I. M., & Beilock, S. L. (2011). Numerical ordering ability mediates the relation between number-sense and arithmetic competence. *Cognition, 121*(2), 256-261. https://doi.org/10.1016/j.cognition.2011.07.009

Maloney, E. A., Ansari, D., & Fugelsang, J. (2011). The effect of mathematics anxiety on the processing of numerical magnitude. *Quarterly Journal of Experimental Psychology, 64*(1), 10-16. https://doi.org/10.1080/17470218.2010.533278

Maloney, E. A., Risko, E. F., Preston, F., Ansari, D., & Fugelsang, J. (2010). Challenging the reliability and validity of cognitive measures: The case of the numerical distance effect. *Acta Psychologica, 134*, 154-161. https://doi.org/10.1016/j.actpsy.2010.01.006

McCloskey, M. (1992). Cognitive mechanisms in numerical processing: Evidence from acquired dyscalculia. *Cognition, 44*, 107-157. https://doi.org/10.1016/0010-0277(92)90052-J

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgments of numerical inequality. *Nature, 215*, 1519-1520. https://doi.org/10.1038/2151519a0

Nairne, J. S., & Healy, A. F. (1983). Counting backwards produces systematic errors. *Journal of Experimental Psychology: General, 112*, 37-40. https://doi.org/10.1037/0096-3445.112.1.37

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin, 116*, 220-244. https://doi.org/10.1037/0033-2909.116.2.220

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. https://doi.org/10.1016/j.neuron.2004.10.014

Pinel, P., Dehaene, S., Riviere, D., & LeBihan, D. (2001). Modulation of parietal activation by semantic distance in a number comparison task. *NeuroImage, 14*, 1013-1026. https://doi.org/10.1006/nimg.2001.0913

Raghubar, K. P., Barnes, M., & Hecht, S. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences, 20*, 110-122. https://doi.org/10.1016/j.lindif.2009.10.005

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

Risko, E. F., & Besner, D. (2008). A role for set in the control of automatic spatial response activation. *Experimental Psychology, 55*, 38-46. https://doi.org/10.1027/1618-3169.55.1.38

Rubichi, S., Nicoletti, R., Iani, C., & Umiltà, C. (1997). The Simon effect occurs relative to the direction of an attention shift. *Journal of Experimental Psychology: Human Perception and Performance, 23*(5), 1353-1364. https://doi.org/10.1037/0096-1523.23.5.1353

Schneider, W., Eschman, A., & Zuccolotto, A. (2001). *E-Prime user's guide*. Pittsburgh, PA, USA: Psychology Software Tools Inc.

Seyler, D. J., Kirk, E. P., & Ashcraft, M. H. (2003). Elementary subtraction. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(6), 1339-1352. https://doi.org/10.1037/0278-7393.29.6.1339

Todkill, A. L., & Humphreys, M. S. (1994). Stimulus comparison strategies and task demands in successive discrimination. *The Quarterly Journal of Experimental Psychology, 47*(3), 761-779. https://doi.org/10.1080/14640749408401136

Turconi, E., Campbell, J. I. D., & Seron, X. (2006). Numerical order and quantity processing in number comparison. *Cognition, 98*, 273-285. https://doi.org/10.1016/j.cognition.2004.12.002

Unsworth, N., Brewer, G. A., & Spillers, G. J. (2013). Working memory capacity and retrieval from long-term memory: The role of controlled search. *Memory & Cognition, 41*(2), 242-254. https://doi.org/10.3758/s13421-012-0261-x

Unsworth, N., & Engle, R. W. (2007). On the division of short-term and working memory: An examination of simple and complex spans and their relation to higher order abilities. *Psychological Bulletin, 133*, 1038-1066. https://doi.org/10.1037/0033-2909.133.6.1038

Unsworth, N., Redick, T. S., Heitz, R. P., Broadway, J., & Engle, R. W. (2009). Complex working memory span tasks and higher-order cognition: A latent variable analysis of the relationship between processing and storage. *Memory, 17*, 635-654. https://doi.org/10.1080/09658210902998047

van Dijck, J. P., & Fias, W. (2011). A working memory account for spatial–numerical associations. *Cognition, 119*(1), 114-119. https://doi.org/10.1016/j.cognition.2010.12.013

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

Van Opstal, F., Gevers, W., De Moor, W., & Verguts, T. (2008). Dissecting the symbolic distance effect: Comparison and priming effects in numerical and nonnumerical orders. *Psychonomic Bulletin & Review, 15*, 419-425. https://doi.org/10.3758/PBR.15.2.419

Verguts, T., Fias, W., & Stevens, M. (2005). A model of exact small-number representation. *Psychonomic Bulletin & Review, 12*, 66-80. https://doi.org/10.3758/BF03196349

Walsh, V. (2003). A theory of magnitude: Common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences, 7*, 483-488. https://doi.org/10.1016/j.tics.2003.09.002

PsychOpen GOLD is a publishing service by
Leibniz Institute for Psychology Information (ZPID),
Trier, Germany. www.leibniz-psychology.org