Published in final edited form as:

J Exp Psychol Learn Mem Cogn. 2014 September ; 40(5): 1389–1403. doi:10.1037/a0036843.

Unlimited capacity parallel quantity comparison of multiple integers

Daryn R. Blanc-Goldhammer and Dale J. Cohen

Department of Psychology, University of North Carolina Wilmington, 601 South College Road, Wilmington, NC 28403-5612

Abstract

Research has shown that integer comparison is quick and efficient. This efficiency may be a function of the structure of the integer comparison system. The present study tests whether integers are compared with an unlimited capacity system or a limited capacity system. We tested these models using a visual search task with time delimitation. The data from Experiment 1 and 2 indicate that integers are encoded, identified, and compared within an unlimited capacity system. The data from Experiment 3 indicates that second order magnitude comparisons are processed with a highly efficient limited capacity system.

Keywords

Numerical Cognition; Numerical Architecture; Parallel Processing; Capacity Limits; Numerical Distance Effect

There is evidence that integers are represented and compared in an analogue fashion (Banks & Hill, 1974; Dehaene, 1997; Dehaene, Dupoux & Mehler, 1990; Gallistel & Gelman, 2000; Gibbon & Church, 1981; Nieder & Miller, 2003). In general, analogue cognitive processes are theoretically slow (e.g., Shepard & Metzler, 1971; Jolicoeur & Kosslyn 1985; Cohen & Kubovy, 1999). Nevertheless, there is evidence that the system that compares the quantities of two integers is both fast and efficient (Milosavljevic et al., 2011; Pavese & Umiltà, 1998; Reynvoet, Brysbaert, & Fias, 2002; Van Opstal, de Lange, & Dehaene, 2011). This efficiency may be a function of the underlying structure of the system. Here, we present three experiments that (a) assess whether the magnitude comparison system is an unlimited parallel capacity system or a limited capacity system, (b) provide evidence that participants are making semantic comparisons between numbers, (c) assess the influence of post-encoding processing time, and (d) compare processing of first and second order numerical comparisons.

Quantity Comparison

Moyer and Landauer (1967) found that when participants were presented with two Arabic digits and asked to indicate which of the two digits was larger, participants response times

were positively related to numerical distance (termed the *numerical distance effect*). In addition, when numerical distance was held constant, participants were quicker to determine the difference between two small numbers than two large numbers (e.g., 2 and 3 versus 7 and 8; termed the *numerical size effect*). These findings are quite robust and have been replicated extensively (e.g., Buckley & Gillman, 1974; Dehaene, 1989; Dehaene, et al., 1990; Duncan & McFarland, 1980; Hinrichs, Yurko, & Hu, 1981; Schwartz & Stein, 1998; Shepard, Kilpatrick & Cunningham, 1975; Takahashi & Green, 1983).

The dominant explanation for the numerical distance effect supposes that the increase in RT arises from the overlap of the perceptual distribution of quantities associated with the digits. That is, digits denote quantities. Although a digit denotes a distinct quantity, one's perception of the quantity associated with the digit is variable. For example, when we see the number 5, sometimes we may understand it as the quantity five, but sometimes we may understand it as four or six. Because there is variability in quantity representation, there are areas of overlap between the representations. Theoretically, the degree of overlap is positively related to time it takes to determine which digit represents the larger quantity (e.g., Buckley & Gillman, 1974; Dehaene, 1997; Gallistel & Gelman, 1992).

There is no dominant explanation for the size effect. Some researchers hypothesize that the size effect arises because the means of the perceptual distributions are logarithmically spaced (Dehaene, 1997; Dehaene, 2003; Dehaene et al., 1990; Moyer & Landauer, 1967). In this instance, the distributions of numerically small quantities spaced one unit apart have less overlap than those of numerically large quantities spaced one unit apart. Others hypothesize that the size effect arises because the variances of the perceptual distributions increase with the quantity denoted by the number (Gibbon, 1977; Whalen, Galistel & Gelman, 1999). Again, in this instance, the distributions of numerically small quantities spaced one unit apart have less overlap than those of numerically large quantities spaced one unit apart. Still, others hypothesize that the size effect results from the response process associated with the digits (Verguts & Fias, 2004; Verguts, Fias & Stevens, 2005).

There is evidence that the comparison of quantities occurs both quickly and effortlessly. For example, a recent article by Van Opstal, de Lange, and Dehaene (2011) showed that participants were able to make quick semantic comparisons of numbers without awareness. In this task, participants were presented with a subliminal set of digits followed by a visible set of digits and were either asked to determine if the sum or the average of the visible set was greater or less than 5. The data showed an influence of the sum or the average of the prime. The authors interpreted this priming effect to indicate the quick parallel processing of digits, though they did not assess capacity limits directly.

Furthermore, Milosavljevic et al (2011) showed that participants were able to make quantity comparisons in as little as 230 ms. In this task, participants were presented with two numbers on a screen and asked to saccade in the direction of the largest number. Saccading enables the participant to respond quicker than is possible with manual responses. RT of accurate responses was 300 ms on average and varied with numerical distance between the two digits.

The distance effect also appears in tasks that do not require the comparison of quantity, suggesting that quantity comparison can be “automatic” (Tzelgov & Ganor-Stern, 2005; Dehaene & Akhavein, 1995; Ganor-Stern, Tzelgov, & Ellenbogen, 2007). For example, in a numerical Stroop task, Pavese and Umiltà (1998) presented participants with a set of digits of a single value and asked them to identify how many digits were present, regardless of the value the digits represented. Because the irrelevant value of the digit interferes with the relevant number of elements response, the authors inferred that quantity comparison is happening automatically. Furthermore, the closer the quantity value is to the correct response, the stronger the interference. The numerical size effect is also present in the Stroop task (Pavese & Umiltà, 1998).

Furthering the evidence that the quantity comparison does not require intention or effort is the presence of the numerical distance effect in priming tasks. For example, Reynvoet, Brysbaert, and Fias, (2002) presented participants a priming digit before the presentation of the target digit. The participants' task was to name the target digit. Participants' RTs were quickest when the prime and target were numerically similar and slower when the prime and target were further apart. The authors concluded that quantity was activated even though it was not needed to complete the task. Although the numerical distance effect was seen in this priming task, no numerical size effect was found.

The research demonstrating the efficiency of magnitude comparison raises the issue of the structure of the system supporting magnitude comparison. Specifically, are there capacity limits on the magnitude comparison process? Unfortunately, determining whether a process is limited or unlimited capacity is not straightforward.

Capacity Limits

Processing capacity can be represented on a continuum in which unlimited capacity parallel processing would be on one extreme and limited capacity serial processing on the other extreme¹. Serial and parallel processing refer to the way in which multiple processes associated with a given task occur. These individual processes are referred to as operations. Each operation has an average amount of time that it takes to complete and there is variability associated with this time. Depending on the model predictions, different operations may start together or at different times (see Appendix for a detailed description of the models).

In a model of parallel processing all operations start at the same time and occur simultaneously. However, each operation may take different amounts of time to complete. There are two types of parallel processing systems: (a) unlimited capacity and (b) limited capacity. An unlimited capacity system is unaffected by the total number of items processed. However, in a limited capacity system processing suffers when the number of operations increases. That is, the processing of individual operations is slowed when there is a high load from many operations occurring simultaneously. In serial processing, only one

¹Although we will not discuss it here, there is also evidence that a system can be supercapacity, in which performance increases with set size (Miller, 1982).

operation occurs at a time. Subsequent operations only occur after the previous operation has been completed.

Serial and parallel processing systems may follow different stopping rules. A stopping rule describes under what conditions processing will stop (Townsend, 1974; Townsend & Ashby, 1983; Townsend & Nozawa, 1997). There are two main types: (a) self-terminating and (b) exhaustive. Self-terminating processing refers to processing that will end when the first operation that is needed for the participant to come to a conclusion or produce a response is completed. Exhaustive processing refers to processing that ends once all of the operations have finished, regardless of if there is enough information to respond early in the processing.

It is easy for the various processing models to be mistaken for one another. In some cases the models make similar predictions. For example, both serial and limited capacity parallel models predict that RT will increase with set size. A limited capacity parallel model predicts that RT will increase as set size increases because the capacity load slows down processing of individual operations. A serial processing model, on the other hand, predicts that RT will increase with set size because the time to process each operation is added to the total RT. An unlimited capacity model predicts that RT will increase as set size increases because of statistical facilitation and statistical inhibition (see Appendix).

Although various processing models are easy to mistake for each other (Townsend, 1990), there are several methods that can differentiate between them. Some of the most common ways are to differentiate models is to compare how different models predict RTs or accuracy in different experimental paradigms. The time delimitation paradigm provides an elegant method for assessing the capacity limits of a cognitive system. Because this paradigm holds set size constant, statistical inhibition and facilitation are not relevant factors.

Shiffrin and Gardner (1972) developed the time delimitation task that distinguished limited from unlimited capacity processing. In this procedure, participants were presented with an array of four characters, each in a position corresponding to a corner of an invisible square. The characters were presented either simultaneously, or successively. If they were presented simultaneously, the stimuli all appeared at once for 50 ms. If they were presented successively, the stimuli were presented one at a time in a predictable order, each for 50 ms. Participants were asked to determine if a target was presented and if so, its location in the array. In both conditions, the target is presented for 50ms. When processing the target, the unlimited capacity system will be unaffected by the number of stimuli present while the target is visible. Therefore, participants should be equally accurate in both the simultaneous and the successive conditions. In contrast, the ability of the limited capacity system to process the target will be negatively correlated with the number of stimuli present while the target is visible. Therefore, participants should be more accurate in the successive condition than the simultaneous condition.

In a series of experiments that varied the presentation and mask orders, Shiffrin and Gardner (1972) found no advantage for successive condition over the simultaneous condition. The authors concluded that characters were identified with an unlimited capacity system. One variation of the successive presentation condition included presenting two letters at a time,

diagonally to each other. Having the two stimuli presented diagonal to each other encouraged participants focus on a point between two stimuli rather than move their eyes. Duncan (1980) expanded on this work, providing evidence for unlimited capacity processing when only one target was present in the task, but evidence for limited capacity processing when two or more targets were present.

Pashler and Badgio (1987) modified the Shiffrin and Gardner (1972) task to force participants to identify every character rather than rely on simple feature detection. To accomplish this, the authors presented participants a set of integers and asked them to identify and locate the highest number presented. The stimuli were presented on the diagonal, much like Shiffrin and Gardner (1972). The target digits were 3-7 and the distractor digits were 0 through (target-1). There were 10 blocks of 30 trials. Simultaneous and diagonal successive blocks alternated. The exposure durations were calculated in the practice trials so that participants could get 60% accuracy in the simultaneous identification task (66.7ms on average). Participants responded by pressing the number on the keyboard that corresponded to the highest number they saw. Participants also indicated the location of the highest number by pressing one of four keys arranged in a square. The authors found no significant advantage for successive presentation. Three other experiments confirmed this finding. The authors concluded that character identification can be done with unlimited capacity processing.

In summary, models of limited and unlimited capacity processing can be difficult to differentiate. Shiffrin and Gardner (1972) developed the time delimitation paradigm that provided an elegant assessment of capacity limits. Pashler and Badgio (1987) adapted the task to involve an exhaustive search of integers by asking participants to identify the largest number. The evidence from the time delimitation studies suggests that alphanumeric stimuli are identified with an unlimited capacity system.

The current work

Research suggests that numerical comparison is quick and efficient. There is also evidence that integers are identified with an unlimited capacity system. Here, we present three time delimitation experiments that assess the capacity limits of the numerical comparison system. In Experiments 1 and 2, we asked participants to identify the largest number in an array. Experiment 1 replicated Pashler and Badgio (1987) with a two alternative forced choice response condition. The addition of this condition provides us with precise RT data so that, for the first time, we can make interpretations about semantic comparisons in time delimitation tasks. Experiment 2 addressed whether response time limits influenced the pattern of results. Experiment 3 introduced a novel comparison task and assessed whether second order comparisons are processed with a limited capacity or unlimited capacity system.

Experiment 1

In Experiment 1, we assessed whether participants identify the largest number in an array with a limited or unlimited capacity processing system. We used a version of the time delimitation paradigm developed by Shiffrin and Gardner (1972). Specifically, we present

four digits, either simultaneously or successively (two at a time on the diagonals), and ask the participant to identify the largest digit. This task requires that participants identify all of the integers in the display, compare them, and identify the largest. We call this type of comparison a *first order comparison*. First order comparisons occur when the participant determines the difference between numerical distributions directly encoded from the stimuli. We had two response conditions: identification and a two alternative forced choice (2AFC) button press condition. In the identification condition, participants input the largest number on a keyboard display. In the 2AFC condition, the participants pressed a button indicating the diagonal that contained the largest number. This condition, if it provides equivalent results as the identification condition, provides a measure of RT that the identification condition does not. Thus, for the first time, we would be able to explore the specifics of how the system completes multiple integer comparisons.

Methods

Participants—Sixty-nine undergraduates in an introductory psychology class volunteered for class credit.

Apparatus and Stimuli—All stimuli were presented on a 24 inch LED color monitor with a 72-Hz refresh rate controlled by a Macintosh Mini running an OS X. The resolution of the monitor was 1920×1200 pixels.

The target stimuli were randomly chosen from the integers 2-9. The distractor stimuli were randomly chosen with replacement from the integers 1 through $(\text{target} - 1)^2$. The integers were presented as Arabic digits in Arial font subtending about 0.68 degrees visual angle vertically.

The integers were presented on a gray square that subtended 5.8 degrees visual angle. In the center of the square there was a red fixation point 5 pixels in diameter. The remainder of the computer screen was black. The square was divided into four imaginary quadrants, with one integer centered in each quadrant. As a result, the four digits were positioned in the corners of an invisible rectangle that was 3.02 by 3.17 cm (approximately 3° visual angle when participant is 60 cm from screen). These measurements were used by Pashler and Badgio (1987) and allowed the entire square to be viewed within the fovea. The top left and lower right quadrants of the square will be referred to as Diagonal A. The lower left and top right quadrants of the square will be referred to as Diagonal B (see Figure 1).

To ensure that both diagonals were equally salient, the integers in one diagonal were dark gray and the integers in the other diagonal were light gray. These grays were equal luminance contrast from the background gray in opposite directions (one lighter, one darker). The diagonal that was dark gray (A or B) was counterbalanced between participants.

²In some cases all distractors were the same value because they were selected randomly with replacement. It is possible that the target is easier to identify in these instances. If this is the case then the data may be biased towards supporting an unlimited capacity model when in fact a limited capacity system is being used. To address this, we tested the models after removing trials where any distractor was repeated. We did not find a difference between these results and our findings presented in the results section.

A mask consists of randomly positioned lines in the light and dark gray of the integers covering a quadrant of the gray square. The lines were sufficiently dense to fully obscure any afterimage.

Procedure—Each trial consisted of a rapid presentation of several frames: the pre-stimulus mask, the target and distractors, and the post-stimulus mask. Participants responded by indicating the integer that denoted the largest quantity, after which feedback was provided.

There were two within-subject presentation conditions: simultaneous and successive. Both conditions began and ended with a 500 ms mask that appeared in each quarter of the gray square. In the simultaneous condition, the pre-stimulus mask was replaced by a target and three distractors positioned in random quadrants for 56 ms (67 ms in the practice trials). In the successive condition, the pre-stimulus mask was replaced by two frames presented in quick succession. Each frame contained two integers presented on a diagonal with masks in the remaining quadrants. The order of the presentation of Diagonal A and Diagonal B relative to one and other was counterbalanced between participants. Each frame was visible for 56 ms (67 ms in the practice trials). The target was randomly presented in one of the quadrants of Diagonal A or B (randomly chosen). There were no interstimulus intervals. See Figure 2 for a timeline for the simultaneous and successive trials.

After the stimuli had been presented, participants were required to make a response to indicate the largest number presented. There were two between-subject response conditions (two-button or number-pad). In the number-pad response condition, participants were presented an on-screen keyboard and used the mouse to indicate the target number. The keypad consisted of buttons labeled with digits 0 through 9. Once the participant clicked on a digit button, the digit appeared in a box at the top of the keypad. The box only permitted one digit to be entered. The participant could click different digit buttons and the number in the box would be replaced. Participants could do this until they were satisfied with their answer at which point they would click a button labeled “OK.”

In the two-button response condition, each diagonal was assigned to a key on the keyboard. Within this condition, half of the participants were instructed to press “k” when the largest number was in Diagonal A, and “d” then the largest number was in Diagonal B. The other half of the participants in the two-button response condition used the reverse key assignment. This condition was added so that precise RTs could be measured.

After the participants responded, they received feedback stating whether or not their response was correct. Then there was a two second interval followed by the next trial. The start of the next trial was identified by the presentation of the pre-mask and focus point. Accuracy and RTs were recorded in all conditions. The timer to measure RTs started as soon as the last stimuli disappeared from the screen.

The successive and simultaneous conditions were blocked. Each block contained 75 trials. The presentation conditions alternated for a total of six blocks. Twelve practice trials were presented prior to the start of each of the first two blocks. The participants were simply told to identify the largest number and were not told that there were different presentation

conditions. Half of the participants were presented the successive presentation condition first and the other half were presented the simultaneous presentation condition first. Participants were allowed to take self-timed breaks between blocks.

Results

Prior to analyzing the data, we removed participants whose average percent correct was less than or equal to 55% correct (about four standard deviations away from 50%). Permitting participants who performed at chance to remain in the analysis would bias the results toward identifying no difference between the simultaneous and successive conditions. This criterion excluded two participants from the study, both of which were in the two-button condition. We also removed trials where the RT was greater than 5000 ms. (over three standard deviations from the mean). RTs this long suggest that the participant may not have been paying attention to the task. In total, we removed 1.05% of the data from slow RTs (0.62% of trials in the two-button condition and 1.49% of trials in the number-pad condition).

An unlimited capacity model predicts equal accuracy in the simultaneous and successive condition. See Table 1 for the mean and standard deviation of accuracy scores in each condition. Using a 2×2 mixed model ANOVA with participants as a random effect, we found that there was no main effect for presentation type, $F(1,65)=.684$, $p=.411$, or for response type, $F(1,65)=.007$, $p=.934$, on accuracy. There was also no significant interaction between presentation conditions and response conditions on accuracy, $F(1,65)=.283$, $p=.597$.

In addition to null hypothesis testing we also ran a Bayesian analysis using the Bayes Factor package in R (anovaBF). We report the Bayes Factor statistic (BF), which represents the posterior odds for the model being true over the prior odds for the model being true. In this analysis BF refers to the odds of the main effect or interaction relative to the odds of the null. A BF of greater than one indicates the odds of the tested model is greater than the odds of the null, while a BF less than one supports the null. Jeffreys (1961) suggests that a BF of greater than 3 indicates substantial evidence for the tested model while a BF of less than .33 indicates substantial evidence for the null hypothesis.

We came to the same conclusions running a Bayes Factor analysis for presentation type, $BF=.23$, and response type, $BF=.19$. The Bayes Factor analysis shows strong evidence against the interaction, $BF=0.01$.

We also examined differences in RT in each condition. See Table 1 for the mean and standard deviation of RT in each condition. Using a 2×2 mixed model ANOVA with participants as a random effect, we found that there was no significant interaction between presentation conditions and response conditions on RT, $F(1,65)=.353$, $p=.554$. There was no main effect of presentation type on RT, $F(1,65)=.070$, $p=.792$. There was a main effect of response type on RT, $F(1,65)=400.758$, $p<.0001$, but it may be disregarded because the nature of the different response types cause differences in RT. It is much quicker to press one of two keys than it is to key in a response on a number-pad. We came to the same conclusions running a Bayes Factor analysis for presentation type, $BF=.19$, and for response

type, $BF = 6.17e+54$. The Bayes model that includes only response type is preferred over all models including the interaction.

Because both response types showed no difference between accuracy in the simultaneous and successive presentation conditions, we can conclude that response type does not change processing load. As such, we used the RTs from the two-button responses to address the numerical distance effect. The numerical distance effect predicts that participants will take longer to compare numbers that are close together than numbers that are far apart. We were interested in how participants compared the largest number in the array to the second largest number in the array. The possible differences between the two highest numbers are one through seven. Because the target was chosen randomly, there is a higher frequency of trials where the difference between the two highest numbers is small than there is of trials where the difference is large. We excluded differences of 7 because there were only 38 trials.

When we plot the average RT for each difference we see a trend toward a negative linear relation (see Figure 3). That is, RTs are slower when the numbers compared are closer together, and faster when the numbers compared are further apart. This is consistent with the numerical distance effect. Note that for differences of four and five the RTs are higher than would be expected by a linear function. This anomaly may have arisen because the integers 7 and 1 were physically similar, and therefore easily confusable. Therefore, we removed trials where 7 was either the largest or second largest number. We removed 7s, rather than 1s, because 1s were never a target number (this allowed us to retain as much data as possible). Figure 4 reveals a strong linear relationship, $F(1,4)=72.98$, $r^2 = .948$, $p = .001$. Both the slope (slope = -23.80; $t = -8.54$, $p = 0.001$) and the intercept (Intercept=486.33; $t = 44.82$, $p < 0.001$) were significant. Figure 5 shows the relation between percent correct and the differences between the two highest numbers (excluding the digit 7). There is a strong monotone relation that is well captured by a power function, $F(1,4) = 214.30$, $r^2 = 0.982$, $p < .001$. Both the intercept (Intercept=-0.247; $t = -46.56$, $p < 0.001$) and characteristic exponent (exponent=0.062; $t = 4.64$, $p < 0.001$) were significant.

The numerical size effect predicts that when the difference between the largest and second largest number is held constant, RTs will be slower for larger numbers than for smaller numbers. It also predicts that accuracy would be lower for larger numbers than for smaller numbers. We plotted RT and percent correct by the largest number in the array when the difference between the largest and second largest numbers was one (where we had the most trials and the best opportunity to discover a size effect). Just as before, we excluded trials where 7 was one of the two highest numbers. There was a small effect for RT, $F(1,4)=8.66$, $r^2 = .684$, $p = .042$, however, it was in the opposite direction predicted by the size effect. Both the slope (slope=-9.34; $t = -2.94$, $p = 0.042$) and the intercept (Intercept=499.80; $t = 29.49$, $p < 0.001$) were significant. There was no size effect when percent correct was the dependent variable, $F(1,4)=0.003$, $r^2 = .001$, $p = .957$.

Discussion

Consistent with the results from Pashler and Badgio (1987) there was no significant difference between the accuracy of the simultaneous and successive presentation conditions. There was also no difference between the RTs of presentation conditions. These results

indicate that the system that identifies and compares numbers is an unlimited capacity parallel processing system³.

Consistent with much of the literature on quantity comparison tasks, we found the distance effect in the two-button response condition. Participants were less accurate and take longer when numbers are numerically close than when they are numerically different. This suggests that participants are making quantity comparisons between these numbers to determine which of the two is larger. It is difficult to interpret our results with regard to the size effect. The RT data showed a reverse size effect while the accuracy scores do not appear to be dependent on the size of the largest number.

One alternative hypothesis is that the integers were all encoded in an unlimited capacity system, but the comparison process occurred after all stimuli were removed. If the comparison process was both separate and limited capacity, the participants would have needed to process the stimuli longer after the post-stimuli mask in the simultaneous condition to achieve equal accuracy in the simultaneous and successive conditions. Our RT data suggest that this was not the case because there was no effect of presentation condition on RTs. Nevertheless, to explore this possibility further, we conducted Experiment 2.

Experiment 2

Experiment 2 assessed the influence of post-encoding processing on the pattern of results revealed in Experiment 1. Specifically, in Experiment 2 we replicated the two-button response condition of Experiment 1 with the addition of response timing conditions. There were three response timing conditions: (a) RT pressure which encouraged participants to respond quickly, (b) forced delay which prevented responses until a period of time had passed, and (c) a control condition which was identical to the two-button response in Experiment 1. If the comparison system relies on a limited capacity process that the participants must compensate for in the simultaneous condition, then we would expect that participants should exhibit either lower accuracy or longer RTs in the simultaneous condition under RT pressure and the control condition. This effect should be removed in the forced delay condition, because the delay should allow all processing to complete. However, if all conditions show equal performance, then we can be confident that the response processes are not masking a limited capacity system.

Methods

Participants—One hundred and four naïve undergraduates in an introductory psychology class volunteered for class credit.

Apparatus and Stimuli—All stimuli were identical to Experiment 1 with the following exceptions related to the timing condition. In the RT pressure condition, a brief tone was used to encourage faster RTs. The tone was presented after the trial when participants

³Although there may be limited circumstances whereby a unlimited capacity serial system can produce similar patterns of results (e.g., the sum of the times to complete the serial processing of all four characters is less than the time available to process the stimuli in the simultaneous condition), we assume those circumstances are not at play here and refer to the process as an unlimited capacity parallel processing system. Further research is required to fully identify the nature of the system.

responded slower than the 85th percentile in a window of their 20 most recent trials. The tone was not presented if the participant was making more than 60% errors in that trial window. Participants were told that if they heard the tone, it meant that they were responding too slowly and should try to respond faster the next time.

In the forced delay condition, the screen would turn green immediately following the stimuli presentation and would remain green for 1000 ms before turning black. Participants were told not to respond during the green screen. Participants' button presses before the termination of the green screen were not registered.

The control condition was the same as Experiment 1. All participants wore Sennheiser HD 280 Pro headphones, regardless of their assigned condition.

Procedure—The procedure was identical to the two-button response type of Experiment 1 with the inclusion of the response timing conditions (described above). The response timing conditions were manipulated between subjects. The blocking of the trials were modified slightly for Experiment 2 to reduce the total number of blocks. The successive and simultaneous conditions were blocked. Each block contained 120 trials. The presentation conditions alternated for each block for a total of four blocks. Ten practice trials were presented prior to the start of each of the first two blocks.

Results

Before analyzing the data, we trimmed the data similar to Experiment 1. This excluded 5 participants for near chance performance and we removed 0.55% of the data from slow RTs (0.45% of the trials in the delay condition, 0.89% of trials in the control condition, and 0.32% of trials in the RT pressure condition).

To assess models of limited and unlimited capacity, we compared accuracy scores in each condition (see Table 2 for the mean and standard deviation). Using a 2×3 mixed model ANOVA with participants as a random effect, we found that there was no main effect for presentation type, $F(1,96)=.064$, $p=.801$, or for response timing, $F(2,96)= 2.064$, $p=.133$ on accuracy. We came to the same conclusions running a Bayes Factor analysis for presentation type, $BF = .92$, and for response timing, $BF = .66$. There was also no significant interaction between presentation conditions and response timing conditions on accuracy, $F(2,96)=.684$, $p=.507$, $BF = .10$.

We also looked at RTs in each condition (see Table 2 for the mean and standard deviation of RT in each condition). Using a 2×3 mixed model ANOVA with participants as a random effect, we found that there was no significant interaction between presentation conditions and response conditions on RT, $F(2,96)=2.26$, $p=.109$. There was no main effect for presentation type on RT, $F(1,96)=.033$, $p=.857$. There was a main effect for response timing on RT, $F(2,96)= 19.854$, $p<.0001$. A Tukey test revealed that there was no difference between RT pressure and delay conditions. However there was a difference between the control and RT pressure conditions. There was also a difference between the control and the forced delay conditions. This result is notable because it suggests that the RT pressure implementation successfully sped up the participants' responses to equal those of the delay

condition. Presumably, the participants in the delay condition were prepared to respond as soon as the experiment permitted them to do so. This suggests that the RT pressure condition implementation was effective. Aside from this, these differences are not necessarily important for the understanding of processing, as the differences are due to the controlling nature of the conditions. It is also not appropriate to compare the delay condition to the others because a quicker response does not necessarily mean quicker processing. It is likely that participants finish processing the numbers before the response is allowed. Then the participants prepare to respond very quickly after the green screen disappeared. We came to the same conclusions running a Bayes Factor analysis for presentation type, $BF = .77$, and for response timing, $BF = 64675.39$. The Bayes model that includes only response type is preferred over all models including the interaction.

To assess the effect of numerical distance, we looked at the relation between RT and the difference between the two largest numbers in the array for each condition. As in Experiment 1, we removed trials in which 7 was either the largest or second largest number because 7 is easily confused with 1. We also excluded differences of seven and eight because there were very few trials in each condition (between 2-9 trials for differences of eight, and between 31-37 trials for differences of seven). Figure 6 shows the RTs for differences in the RT pressure condition. There is a strong linear relation, $F(1,4)=25.09$, $r^2 = .863$, $p = 0.007$. Both the slope (slope=-13.55; $t = -5.01$, $p = 0.007$) and the intercept (Intercept = 359.17; $t = 34.09$, $p < 0.001$) were significant. Figure 7 shows the accuracy scores for differences in the RT pressure condition. There is a strong linear relationship, $F(1,4)=71.91$, $r^2 = .947$, $p = .001$. Both the slope (slope=0.02; $t = 8.48$, $p = 0.001$) and the intercept (intercept=0.72; $t = 72.16$, $p < 0.001$) were significant. Figure 6 shows the RTs for differences in the control condition. There is a linear relationship, $F(1,4)=7.50$, $r^2 = .652$, $p = 0.052$. Both the slope (slope=-23.67; $t = -2.74$, $p = 0.052$) and the intercept (intercept=619.26; $t = 18.40$, $p < 0.001$) were significant. Figure 7 shows the accuracy scores for differences in the control condition. There is a strong linear relationship, $F(1,4)=21.91$, $r^2 = .846$, $p = .009$. Both the slope (slope=0.02; $t = 4.68$, $p = 0.009$) and the intercept (intercept=0.74; $t = 43.97$, $p < 0.001$) were significant. The RTs in the delay condition were not useful for this analysis because they do not necessarily represent the time needed to compare numbers (The slope was .17, $t = 0.06$, $p = 0.959$, and the intercept was 421.15, $t = 34.11$, $p < 0.001$, $F(1,4)=.003$, $r^2 = .001$, $p = .959$). We can, however, address the numerical distance effect in the forced delay condition by looking at the percent correct. This is because the accuracy scores were not affected by the response timing conditions. Figure 7 shows the accuracy scores for differences in the forced delay condition. There is a linear relationship, $F(1,4)=10.90$, $r^2 = .732$, $p = .030$. Both the slope (slope=0.02; $t = 3.30$, $p = 0.030$) and the intercept (intercept=0.77; $t = 35.89$, $p < 0.001$) were significant.

To address the numerical size effect we examined the RTs and accuracy for each of the largest numbers in the array when the difference between the two largest numbers was held constant. The difference between the two largest numbers was held constant at one to maximize the data available. We excluded trials where 7 was one of the two highest numbers. For the RT analysis, we combined the RT pressure and control conditions. We did not include the RTs from the delay condition because they do not give us precise

information about processing time. Although there was a significant relation of the largest number on RT, $F(1,4) = 15.79$, $r^2 = .798$, $p = .016$ (slope = -9.18; $t = -3.97$, $p = 0.016$; intercept = 513.11; $t = 41.582$, $p < 0.001$), it was in the direction opposite to that predicted by the size effect. For the accuracy analysis, we included all conditions. There was no effect of largest number on accuracy, $F(1,4) = 0.9445$, $r^2 = .191$, $p = .386$. We come to the same conclusions when the analyses are run on individual response conditions.

Discussion

The results indicate that integers are identified with an unlimited capacity parallel system which is consistent with the results from Experiment 1. Our results extend the previous findings to look at the effects of response timing conditions on first order comparisons. Response timing conditions can be used to increase or decrease post-encoding processing time. We found that accuracy does not change when post-encoding processing time is increased or decreased.

We also found the numerical distance effect in the RT data for the RT pressure and control conditions. This suggests that participants are using quantity information to make first order comparisons. Although RT data was uninformative concerning the numerical distance effect in the forced delay condition, the numerical distance effect was present in that condition's accuracy data. We conclude that participants determine the largest number by comparing quantities using an unlimited capacity parallel system. Just as in Experiment 1, we found a reverse size effect in the RT data and no size effect in the accuracy scores. These results are difficult to interpret.

We discuss these findings in the General Discussion.

Experiment 3

In the first two experiments we assessed whether participants make first order comparisons using a limited capacity or an unlimited capacity processing system. We determined they make these comparisons using an unlimited capacity parallel processing system. In the third experiment, we introduced a novel comparison task to assess whether participants make *second order comparisons* using a limited capacity or unlimited capacity processing system. In a second order comparison, participants determine the difference between two derived quantities, rather than quantities that were directly encoded from the stimuli. In this case, we were interested in how participants identify the largest difference between two numbers. We used the same paradigm as in the first two experiments but rather than asking participants to identify which diagonal contained the largest number we asked them to identify which diagonal contained the largest numerical difference.

Methods

Participants—One hundred and forty-seven naïve undergraduates in an introductory psychology class volunteered for class credit.

Apparatus and Stimuli—The apparatus and stimuli were the same as in Experiment 2, with the exception that the numbers in each diagonal were paired such that the numerical

distance between them was between one and three. For example, a diagonal may contain the integers 2 and 5 because the difference between them is three, but the diagonal would not contain the integers 2 and 9 because the difference between them is seven. The purpose of using these pairs was to ensure that the individual integers would not be correlated with target differences.

Procedure

The procedure was nearly identical to Experiment 2 with the exception that participants were asked to identify the diagonal that contained the largest numerical difference between its' two numbers. For example, if diagonal A contained integers 9 and 7 the difference is two. If diagonal B contained the integers 3 and 6 the difference is three. Here, the participant would press the key that corresponded to diagonal B because its difference was the greatest.

Similar to Experiment 2, we implemented a between subjects response timing manipulation. There were two conditions: RT Pressure and Control. We did not include the forced delay response timing condition because we were not interested in the effects of increasing post encoding processing time (i.e. Increasing the use of math strategies). The RT pressure condition was implemented specifically to inhibit participants' use of counting or subtraction strategies rather than comparing quantity representations. Simultaneous and successive conditions were blocked as they were in Experiment 2.

Results

Before analyzing the data, we trimmed the data similar to Experiment 1. We removed 34 participants because their average percent correct was below 55%. This task had a much higher fail rate than the task used in the previous two experiments, likely because it was a much more difficult task. In total, we removed 0.78% of the data due to slow RTs (0.48% of trials in the RT pressure condition and 1.07% of trials in the control condition).

To test models of limited and unlimited capacity, we calculated accuracy scores for each condition. See Table 3 for the mean and standard deviation of accuracy scores in each condition. Using a 2×2 mixed model ANOVA with participants as a random effect, we found that there was no significant interaction between presentation conditions and response timing conditions on accuracy, $F(1, 111)=.289, p=.592$. There was no main effect of response timing condition on accuracy, $F(1, 111)=2.073, p=.153$. However, there was a main effect of presentation condition on accuracy, $F(1, 111)= 7.662, p=.007$. Percent correct was significantly higher in the successive presentation condition than in the simultaneous presentation condition. We came to the same conclusions running a Bayes Factor analysis for presentation type, $BF = 20.48$, and for response timing, $BF = 1.13$. The Bayes model that includes only presentation type is preferred over all models including the interaction.

We also examined differences between RTs in each condition (see Table 3 for the mean and standard deviation of RTs in each condition). Using a 2×2 mixed model ANOVA with participants as a random effect, we found there was a main effect of presentation type on RT, $F(1, 111)=7.838, p=.006$. RTs were significantly faster in the successive presentation condition than in the simultaneous presentation condition. There was also a main effect of

response timing condition on RT, $F(1, 111) = 20.682, p < .0001$. RTs were significantly faster in the RT pressure condition than in the control condition. There was no significant interaction between presentation conditions and response timing conditions on RT, $F(1, 111) = 2.12, p = .149$. We came to the same conclusions running a Bayes Factor analysis for presentation type, $BF = 2.27$, and for response timing, $BF = 468.44$. The Bayes model that includes both presentation type and response timing ($BF = 1358.51$) is preferred over all models including the interaction.

Finally, we examined the numerical distance effect in second order comparisons. Recall, the diagonals could have a difference of one, two, or three. There were three possible combinations of diagonal differences: one and two (1v2), one and three (1v3), or two and three (2v3). To address the numerical distance effect, we compared RT in instances when the distance between the two diagonal differences was one (1v2 and 2v3) to instances when it was two (1v3). See Table 4 for the means and standard deviations of the RT. We analyzed the data with a one-way mixed model ANOVA with subject as a random effect. There was a significant difference between the RT for differences of one and two which indicates the numerical distance effect, $F(1, 112) = 47.00, p < .001$. There is also a significant difference between RTs for differences of one and two when we examine the RT pressure condition separately, $F(1, 52) = 14.13, p < .001$, and the control condition separately, $F(1, 59) = 34.68, p < .001$. To further support evidence for the numerical distance effect we analyzed the accuracy scores with a one-way mixed model ANOVA with subjects as a random effect. See Table 4 for the means and standard deviations of the percent correct. There is a significant difference between the accuracy for differences of one and two, $F(1, 112) = 268.11, p < .001$. There is also a significant difference between the accuracy of differences one and two in the RT pressure condition alone, $F(1, 52) = 154.74, p < .001$, as well as the control condition alone, $F(1, 59) = 124.41, p < .001$.

We will not address the numerical size effect because it is unclear what the predictions in this task would be. Typically, the size effect predicts that it would be more difficult to make comparisons for 2v3 than for 1v2. However, it is likely that the variance of the diagonal differences is a function of the variance of the numbers in the diagonal. There are many possible combinations of numbers in a diagonal to get a given difference. Therefore, it may not be accurate to treat those differences as the same. For example, there is a difference of 2 for both 4 and 6, and 7 and 9. It is possible, but uncertain, that these differences of 2 do not have equal variance. Therefore, it is unclear if RTs would be predicted from the diagonal differences, from the original numbers, or from a combination of the two.

Discussion

The results indicate that the second order comparisons are processed with a limited capacity system under both RT pressure and no RT pressure. There was a significant advantage for speed and accuracy in the successive presentation condition. The consistency between the processing systems in the RT timing conditions is in agreement with the results from Experiment 2. We see the numerical distance effect, which suggests that participants are making quantity comparisons between the two diagonal differences. Overall, the results

from this experiment indicate that participants identify differences by comparing quantity representations with a limited capacity system.

General Discussion

We developed a set of three experiments to explore the capacity limits of the quantity comparison system. In the first two experiments we explored how integers are identified and first order comparisons are processed. In the third experiment we explored how second order comparisons are processed. We found that integers are identified and first order comparisons are made using an unlimited capacity parallel system. Varying post-encoding processing time did not influence the results. We also found that second order comparisons are processed with a limited capacity system and that this does not change when we shorten the post-encoding processing time.

First order comparisons

Experiments 1 and 2 required first order quantity comparisons. In these experiments, integer identification required the encoding of number stimuli and the comparison of quantities to identify the largest. Because participants were able to identify the largest number using an unlimited capacity processing system, we conclude that they were able to encode and compare integers in parallel. Furthermore, the RT and accuracy data reveal the presence of the numerical distance effect but not the predicted numerical size effect.

Capacity Limits—Experiments 1 and 2 reveal that the numerical cognition system can identify Arabic digits and compare the quantity of those digits with an unlimited capacity parallel system. This finding clarifies the structure of the numerical cognition system as well as provides a foundation to explain some of the extant data in the literature. Specifically, there is general agreement in the literature that Arabic digits activate a perceptual distribution of quantities and the numerical distance effect arises from the degree of overlap of these distributions. The question we addressed here is whether the cognitive process that compares these distributions has capacity limits. The data indicates that it does not and that the comparison processes occur in parallel.

The discovery that Arabic digits are identified and compared in an unlimited capacity parallel process provides an architectural foundation for effects in which irrelevant quantity information interferes with a quantity response (e.g., the numerical Stroop effect and size congruency effect). For example, the size congruency effect occurs when two digits are presented, but they are different physical sizes. The participant is required to identify the physically larger stimulus, but the numerical dimension interferes with the response. Here, for interference to occur, two quantity comparisons must be occurring simultaneously: the size comparison and the numerical comparison. This interference becomes most likely if the system that compares quantity is both unlimited capacity and parallel. In such an instance, quantity comparisons will occur simultaneously and efficiently.

An unlimited capacity parallel system that identifies and compares the magnitude of digits also provides an architectural foundation for the processing of proportions. Many formats of proportions express quantity information using two integers. For example, to express the

quantity “half,” one may present a fraction “1/2,” a relative frequency, “50 in 100,” odds “1:1,” etc. To extract the proportion information, one must compare the quantities presented. Take the fraction, for example. Here, to extract the proportion information, one must compare the quantity denoted by 1 to the quantity denoted by 2 and extract the relative difference. Although several researchers have data supporting the conclusion that proportions expressed as two integers are processed featurally (Bonato, Fabbri, Umiltà & Zorzi, 2007; Opfer & DeVries, 2008), others have concluded that proportions expressed as two integers are processed together (Ganor-Stern, Karaik-Rivkin, & Tzelgov, 2011; Ischebeck, Schocke & Delazer, 2009; Meert, Gregoire & Noel, 2009; Meert, Gregoire & Noel, 2010; Schneider & Siegler, 2010; Sprute & Temple, 2011). Although we do not express an opinion on this debate here, we note that an unlimited capacity parallel system would provide the necessary architectural foundation for processing both components of a proportion simultaneously and comparing two proportions.

The finding that RTs were linearly related to the numerical distance between the target and the next largest number in the array provides insight into how the magnitude comparison process deals with multiple comparisons. We start with the assumption that to identify the largest number, all numbers must be compared to each other. Given that there is no difference between the accuracy in the simultaneous and successive presentation conditions, we concluded that all of these comparisons are occurring in parallel in an unlimited capacity system. The numerical distance effect predicts that it will take longer to compare numbers that are closer together than it will to compare numbers that are further apart. Therefore, although all comparisons have the same start time, they will have different processing times depending on the magnitude of the numbers being compared.

We found that RT is a function of the time to find the difference between the two largest numbers, rather than the most difficult comparison (i.e., the comparison where the numerical difference is the smallest). Therefore, it is likely that these long and uninformative processes are being terminated before completion. The termination of these long uninformative comparisons likely arises from the completion of quicker comparisons. The quickest comparison would be between the largest and smallest number. It may be that once the smaller number in that comparison is rejected, all comparisons involving that number are terminated. Similarly, as other comparisons complete, all comparisons involving the smaller number in those pairs are terminated. This cycle would continue until the largest number is determined. This process may be termed a “parallel other-terminating process.”

We found no evidence for the predicted numerical size effect in either Experiment 1 or 2. In fact, we found a reverse size effect in the RT data such that it took longer to compare smaller numbers than larger numbers. It is unclear how these results should be interpreted. The numerical size effect is considered to be one of the main indicators of a compressed mental number line. On a logarithmically compressed number line the means of large numbers are closer together than the means of smaller numbers (Dehaene, 1997; Dehaene, 2003; Dehaene, Dupoux & Mehler, 1990; Moyer & Landauer, 1967). On a linear number line with scalar variance, the variance of large numbers is greater than the variance of small numbers (Gibbon, 1977; Whalen, Galistel & Gelman, 1999). In both of these cases there is more overlap in the perceptual distributions of larger numbers, making them more difficult

to differentiate. This will cause longer RTs and lower accuracy for large number comparisons relative to small number comparisons. Because we do not see this effect, there is no indicator of compressed means or of scalar variance. It is possible that the reverse size effect is evidence that the typical size effect seen in the data has a different origin than the numerical distance effect (Verguts, Fias, and Stevens, 2004; Verguts, & Van Opstal, 2005). Because we are unclear how to interpret these results, we make no strong conclusions regarding the size effect.

The purpose of controlling response timing in Experiment 2 was to manipulate the amount of time that was available for post-encoding processing. Traditional random walk and diffusion models of RT propose that stimuli are encoded first and then samples are taken from perceptual distributions until the accumulated knowledge passes a set threshold. This sampling process can be termed “post-encoding” processing. The traditional interpretation of these RT models is that the amount of post-encoding time should be positively correlated with accuracy. In the RT pressure condition, participants were encouraged to use less post-encoding processing time than participants in the control condition. In the forced delay condition, participants were encouraged to use more post-encoding processing time than participants in the control condition. Therefore, we would expect that participants would be more accurate in the forced delay condition. Although we see a trend toward this pattern, the differences between the response timing conditions are not significant. The lack of a difference in accuracy scores between the response timing conditions suggests that post-encoding processing time has little impact on the comparison process.

Second order comparisons

Our third experiment addressed the issue of second order comparisons, or comparisons of the difference between two previously determined differences. In this experiment, we found that second order comparisons are computed within a limited capacity system.

The literature suggests that an unlimited capacity system is used to process alphanumeric characters unless multiple targets are present (e.g., Duncan, 1980). Duncan (1980) proposes that the basic characteristics and meaning of stimuli are identified using an unlimited capacity system. A second, limited capacity, system is used to process selected stimuli (ie. targets). Duncan's (1980) theory states that the effects of a limited capacity system will only appear if multiple targets must be identified independently.

When Experiment 3 is viewed within the context of Duncan's (1980) theory, the task may require the participant to process two selected stimuli. Specifically, the unlimited capacity system may identify the largest integer in each of the two diagonals. However, the unlimited capacity system may be incapable of completing the second order comparison because the “differences” are not physical stimuli. Therefore, the two differences determined by the unlimited capacity system may be passed to the limited capacity system for further processing. From this perspective, second order comparisons are processed with a limited capacity system that is preceded by an unlimited capacity system.

Alternatively, the task itself may influence the data. In our task, the participant had to use information from the first order comparison to complete the second order comparison.

Therefore, the first order comparison had to begin before the second order comparison could begin. The first order comparisons may or may not need to complete before the second order comparisons begin (e.g., serial processing, or a cascade model). Nevertheless, even a minimal, though consistent, shift in start times may influence the outcome of the task.

The presence of the numerical distance effect in the second order comparisons indicates that the largest difference is determined by comparing the quantity representations of the diagonal differences. The distributions on which the second order comparisons are determined are derived from the comparisons occurring in the first order comparisons. This relation between the first and second order comparisons provided some useful information from which we can make predictions.

In first order comparisons, encoded quantities are compared to each other by comparing samples from the perceptual distributions of quantity. These sampled comparisons provide the values that produce the second order derived distributions (i.e., the difference distributions). The mean of the derived distributions is a function of the difference between the compared distributions. The variance of the derived distributions should be the sum of the variances of the original distributions. This is because for independent random variables, the variance of their sum or difference is the sum of their variances. Therefore, derived distributions should have twice as much variance as encoded distributions. This would increase the amount of overlap between the derived distributions, making second order comparisons more difficult than first order comparisons.

In fact, we can look at the slope of the numerical distance effect between the first and second order comparisons to determine whether there was an increase in difficulty when comparing quantities. However, because we only had the numerical distances of 1 and 2 in the second order comparisons, our analysis was constrained to those two distances. We find, however, that second order comparisons take longer than first order comparisons. That is, in first order comparisons, the RT when the distance is one (460 ms) is only about 25 ms longer than when the distance is two (435 ms). However, in second order comparisons, the RT when the distance is one (644 ms) is about 52 ms longer than when the distance is two (592 ms). This difference is likely accounted for by the fact that derived distributions are noisier than encoded distributions. We see a similar effect with the accuracy data. There is a larger difference between percent correct for distances of one and two in the second order comparisons (.08) than the first order comparisons (.04).

It is important to look at the effects of reducing post-encoding processing time on second order comparisons for two main reasons. First, because second order comparisons are occurring after the initiation of first order comparisons, it is possible that they will be more vulnerable to the effects of limited processing time. Second, it is possible that for more complex operations, like second order comparisons, participants may use strategies such as counting or mental subtraction. Reducing the post-encoding processing time, with the RT pressure condition, will reduce the time participants have to use such strategies. If participants are frequently using mathematical strategies, we would expect their accuracy scores to be lower in the RT pressure condition.

However, we did not see a significant decrease in accuracy when post-encoding processing time was decreased. Therefore, we concluded that the system is efficient enough to make first and second order quantity comparisons with limited processing time. Thus, errors in second order comparisons are likely due to encoding errors that occur during first order comparisons. Because only one stimulus is needed to be encoded incorrectly for a diagonal difference to be incorrect, there is a higher probability that encoding errors will show up in second order comparisons than in first order comparisons. This may explain why overall accuracy is lower for second order comparisons than for first order comparisons.

These results also suggest that second order comparisons do not require mathematical calculations or counting strategies. It is unlikely that participants were performing calculations because they responded in 750 ms on average in the control condition and 500 ms in the RT pressure condition. It would be very difficult to perform a calculation that quickly. Furthermore, if they were performing calculations we would expect a decrease in accuracy when there was less time to process these calculations. When given an RT pressure, participants responded on average 250 ms quicker than the control group, yet there was no significant difference in accuracy. Therefore, we conclude that participants are not making comparisons by using mathematical calculations, but rather comparing quantities pre-attentively.

Conclusion

In summary, we developed three experiments to explore the capacity limits of the system supporting quantity comparison. We found that (a) integers are encoded and identified using an unlimited capacity parallel processing system, (b) first order comparisons are processed with an unlimited capacity parallel processing system, (c) this comparison process appears to be analog, and (c) second order comparisons are processed with a limited capacity system. The limited capacity system is still highly efficient.

Acknowledgments

This work supported by NIH grant RO1HD047796 awarded to Dale J. Cohen.

Appendix

For the purposes of this task, we define an unlimited capacity system as one that has simultaneous, independently occurring operations (i.e. unlimited capacity independent processing; Townsend & Nozawa, 1995; Townsend & Wenger, 2004). The total number of operations does not affect the performance of individual operations. We define a limited capacity system as any system that performs worse than the unlimited capacity system. By definition, if we find evidence for an unlimited capacity system we can conclude that it is parallel processing, but if we found evidence for a limited capacity system we would not be able to immediately identify if it was a serial or parallel processing system. More subtle testing is needed to determine the details of a limited capacity system.

Let us say that the time an operation will end at (T) is a function of the time it started at (S), plus the average ms it takes to complete (x) with some error variance (e); here we assume the error is normally distributed:

$$T=S+x+e, e \sim \eta(0, \sigma) \quad (1)$$

Parallel processing refers to instances in which multiple operations occur simultaneously. In this case, all operations start at the same time, but may take different amounts of time to complete. It is possible that there is variability in the start time but we make the assumption that they all start at the same time without variability. If we assume that multiple occurrences of the same operation proceed in parallel, then we assume that each operation has the same average time and error variance.

$$\begin{aligned} T_1 &= S+x+e \\ T_2 &= S+x+e \\ T_3 &= S+x+e \end{aligned} \quad (2)$$

Parallel processing can have either limited or unlimited capacity. An unlimited capacity system is unaffected by the total number of items processed. However, in a limited capacity system processing suffers when the number of operations increases. That is, the processing of individual operations is slowed when there is a high load from many operations occurring simultaneously. Let us specify f_n as the functional relation between number of operations and time:

$$T_1=S+f_n(x)+e \quad (3)$$

In an unlimited capacity system, $f_n = 1$, and the number of operations would have no effect on the time of processing. In a limited capacity system, $f_n > 1$. Depending on the function, this can have a monotone effect, or can have no effect until a threshold (A) is reached.

$$\text{if } n < A, f_n = 1; \text{if } n \geq A, f_n > 1 \quad (4)$$

That is, if the number of operations is below the threshold there will be no effect on processing time but if the number of operations is above the threshold there will be an effect on processing time.

In serial processing, only one operation occurs at a time. Subsequent operations only occur after the previous operation has been completed.

$$\begin{aligned} T_1 &= S+x+e_1 \\ T_2 &= T_1+x+e_2 \\ T_3 &= T_2+x+e_3 \end{aligned} \quad (5)$$

It is relatively straightforward to show that RT will increase with set size in both serial and limited capacity parallel systems. Limited capacity parallel processing takes longer as set

size increases because f_n is greater than one, which directly increases the mean RT as a function of x . Serial processing, on the other hand, has increased RT with set size because the time to process each operation is added to the total RT.

Serial and parallel processing systems may follow different stopping rules. A stopping rule describes under what conditions processing will stop (Townsend & Nozawa, 1997). There are two main types: self-terminating and exhaustive. Self-terminating processing refers to processing that will end when the first operation that is needed for the participant to come to a conclusion or produce a response occurs. Exhaustive processing refers to processing that ends once all of the operations have finished, regardless of if there is enough information to respond early in the processing.

In a self-terminating parallel model overall RT is equal to the end time of the operation that identifies the target (T_x):

$$RT = T_x \quad (6)$$

In a self-terminating serial model overall RT is equal to the end time of the operation that identifies the target (T_x). This is a function of the number of operations (t) before T_x that do not identify the target:

$$RT = T_x = S + \sum_{i=0}^{t+1} (x + e) \quad (7)$$

In an exhaustive parallel model overall RT is equal to the end time of the slowest operation:

$$RT = \max\{T_1, T_2, T_3 \dots T_n\} \quad (8)$$

In an exhaustive serial model overall TR is equal to the end time of the last operation:

$$RT = T_n \quad (9)$$

Statistical facilitation and statistical inhibition are factors that may make unlimited capacity models mimic limited capacity models. These terms refer to how varying the number of operations occurring in a parallel model can affect RT and accuracy in detection tasks. Statistical facilitation refers to how increasing the number of operations can decrease RT or increase accuracy in processing. With regard to RT, increasing the number of operations increases the likelihood that there will be a quick process that terminates further processing. Imagine that each operation is a runner in a race. If there are more runners in the race, there is a higher chance that there will be a very fast runner and a very slow runner. Once the fastest runner crosses the finish line, there is a winner and it does not matter when the remaining runners finish. In this case, statistical facilitation can make a self-terminating parallel model look a lot like a self-terminating serial model. As set size increases and there

are more target items in a self-terminating serial model, there is an increase in probability that a process that will terminate subsequent processing occurs first.

Statistical facilitation may also make it difficult to differentiate models using accuracy scores. Having more operations can increase the chance that a response will be correct. This is because the probability of having all operations be performed incorrectly (P_{error}) is lower than the probability of any individual operation being performed incorrectly (P_e).

$$P_{\text{error}} = (P_{e1})(P_{e2})(P_{e3}) \dots (P_{en}) \quad (10)$$

For example, if each operation has a .1 chance of being incorrect then the probability of getting two incorrect is .01, and the probability of getting three incorrect is .001. It may appear on the surface that accuracy should be unaffected by set size in a parallel unlimited capacity system. However, we can see that statistical facilitation can increase accuracy in tasks where all operations must be performed incorrectly to give an incorrect response.

Statistical inhibition refers to how increasing the number of operations can increase RT or decrease accuracy in processing. As mentioned above, increasing the number of operations increases the likelihood that there will be a very slow operation and a very quick operation. In a parallel model with exhaustive processing, processing stops only after all operations have been completed. If there are more operations, it is more likely that there will be a slow process that slows the overall RT. Using the race example above, processing would not stop until all of the runners crossed the finish line. If there are many runners, there is a higher chance that there will be a slow runner that will cause the race to end late. It is often difficult to differentiate this from exhaustive serial processing because the more operations in serial processing, the longer the RT because each process must be completed one at a time.

Statistical inhibition can also affect accuracy scores. In a self-terminating unlimited capacity parallel model, making a mistake on one operation terminates processing. Therefore, the percent chance of being correct on one process ($1 - P_e$) is higher than the chance of being correct on all processes (P_{correct}).

$$P_{\text{correct}} = (1 - P_{e1})(1 - P_{e2})(1 - P_{e3}) \dots (1 - P_{en}) = 1 - P_{\text{error}} \quad (11)$$

For example, if a given operation has a .8 chance of being correct, then the chance of being correct on two operations is .64, and the chance of being correct on three is .512. On the surface an unlimited parallel model would not predict that the accuracy would decrease with set size. However, statistical inhibition can decrease accuracy scores when the task requires that only one operation be performed correctly to come to a correct response.

References

- Banks WP, Hill DK. The apparent magnitude of number scaled by random production. *Journal of Experimental Psychology Monograph*. 1974; 102:353–376.
- Bonato M, Fabbri S, Umiltà C, Zorzi M. The mental representation of numerical fractions: real or integer? *Journal of Experimental Psychology: Human Perception and Performance*. 2007; 33(6): 1410–1419. [PubMed: 18085953]

- Buckley P, Gillman C. Comparisons of digits and dot patterns. *Journal of Experimental Psychology*. 1974; 103(6):1131–1136. [PubMed: 4457588]
- Cohen DJ, Kubovy M. Even feature integration is cognitively impenetrable. *Behavioral and Brain Sciences*. 1999; 22:371–372.
- Dehaene S. The psychophysics of numerical comparison: A reexamination of apparently incompatible data. *Perception & Psychophysics*. 1989; 45:557–566. [PubMed: 2740196]
- Dehaene S, Akhavein R. Attention, Automaticity, and Levels of Representation in Number Processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1995; 21(2):314–326.
- Dehaene, S. The number sense: How the mind creates mathematics. New York: Oxford University Press; 1997.
- Dehaene S. The neural basis of the Weber-Fechner law: A logarithmic mental number line. *Trends in Cognitive Sciences*. 2003; 7(4):145–147. [PubMed: 12691758]
- Dehaene S, Dupoux E, Mehler J. Is numerical comparison digital? Analogical and Symbolic effects in two-digit number comparison. *Journal of Experimental Psychology: Human Perception and Performance*. 1990; 16:626–641. [PubMed: 2144576]
- Duncan EM, McFarland CE. Isolating the effects of symbolic distance and semantic congruity in comparative judgments: An additive-factors analysis. *Memory & Cognition*. 1980; 8:612–622. [PubMed: 6163942]
- Duncan J. The locus of interference in the perception of simultaneous stimuli. *Psychological Review*. 1980; 87(3):272–300. [PubMed: 7384344]
- Gallistel CR, Gelman R. Preverbal and verbal counting and computation. *Cognition*. 1992; 44:43–74. [PubMed: 1511586]
- Gallistel CR, Gelman R. Non-verbal numerical cognition: From reals to integers. *Trends in Cognitive Sciences*. 2000; 4:59–65. [PubMed: 10652523]
- Ganor-Stern D, Karasik-Rivkin I, Tzelgov J. Holistic representation of unit fractions. *Experimental Psychology*. 2011; 58(3):201–206. [PubMed: 21106472]
- Ganor-Stern D, Tzelgov J, Ellenbogen R. The representation of two-digit numbers. *Journal of Experimental Psychology: Human Perception and Performance*. 2007; 33:483–496.
- Gibbon J. Scalar expectancy theory and Weber's law in animal timing. *Psychological Review*. 1977; 84(3):279–325.
- Gibbon J, Church RM. Time left: Linear versus logarithmic subjective time. *Journal of the Experimental Analysis of Behavior*. 1981; 7:87–107.
- Hinrichs JV, Yurko DS, Hu JM. Two-digit number comparison: use of place information. *Journal of Experimental Psychology: Human Perception and Performance*. 1981; 7:890–901.
- Ischebeck A, Schocke M, Delazer M. The processing and representation of fractions within the brain: An fMRI investigation. *NeuroImage*. 2009; 47:403–413. [PubMed: 19328235]
- Jeffreys, H. *Theory of Probability*. Third. Oxford University Press; Oxford, England: 1961.
- Jolicoeur P, Kosslyn SM. Demand characteristics in image scanning experiments. *Journal of Mental Imagery*. 1985; 9(2):41–50.
- Meert G, Grégoire J, Noël MP. Rational numbers: Componential vs. holistic representation of fractions in a magnitude comparison task. *The Quarterly Journal of Experimental Psychology*. 2009; 62(8): 1598–1614. [PubMed: 19123119]
- Meert G, Grégoire J, Noël MP. Comparing 5/7 and 2/9: Adults can do it by accessing the magnitude of the whole fractions. *Acta Psychologica*. 2010; 135:284–292. [PubMed: 20797686]
- Miller JO. Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*. 1982; 14:247–279. [PubMed: 7083803]
- Milosavljevic M, Madsen E, Koch C, Rangel A. Fast saccades toward numbers: Simple number comparisons can be made in as little as 230 ms. *Journal of Vision*. 2011; 11(4):1–12.
- Moyer R, Landauer T. Time required for judgments of numerical inequality. *Nature*. 1967; 215(5109): 1519–1520. [PubMed: 6052760]
- Nieder A, Miller EK. Coding of cognitive magnitude: compressed scaling of numerical information in the primate prefrontal cortex. *Neuron*. 2003; 37:149–157. [PubMed: 12526780]

- Opfer JE, DeVries J. Representational change and magnitude estimation: Why young children can make more accurate salary comparisons than adults. *Cognition*. 2008; 108:843–849. [PubMed: 18573491]
- Pashler H, Badgio P. Visual attention and stimulus identification. *Journal of Experimental Psychology: Human Perception & Performance*. 1985; 11:105–121. [PubMed: 3159831]
- Pashler, H.; Badgio, PC. Attentional issues in the identification of alphanumeric characters. In: Coltheart, M., editor. *Attention and Performance XII: The Psychology of Reading*. Hove, UK: Lawrence Erlbaum Associates Ltd.; 1987. p. 63–82.
- Pavese A, Umiltà C. Symbolic distance between numerosity and identity modulates Stroop interference. *Journal of Experimental Psychology: Human Perception & Performance*. 1998; 24:1535–1545. [PubMed: 9988600]
- Reynvoet B, Brysbaert M, Fias W. Semantic priming in number naming. *Quarterly Journal of Experimental Psychology A*. 2002; 55(4):1127–1139.
- Schneider M, Siegler RS. Representations of the magnitudes of fractions. *Journal of Experimental Psychology: Human Perception and Performance*. 2010; 36(5):1227–1238. [PubMed: 20873937]
- Schwartz W, Stein F. On the temporal dynamics of digit comparison. *Journal of Experimental Psychology: Learning, Memory, & Cognition*. 1998; 24:1275–1293.
- Shepard RN, Kilpatrick DW, Cunningham JP. The internal representation of numbers. *Cognitive Psychology*. 1975; 7:82–138.
- Shepard H, Metzler J. Mental Rotations of three-dimensional objects. *Science*. 1971; 171:701–703. [PubMed: 5540314]
- Shiffrin RM, Gardner GT. Visual processing capacity and attentional control. *Journal of Experimental Psychology: General*. 1972; 93:72–82.
- Sprute L, Temple E. Representations of Fractions: Evidence for Accessing the Whole Magnitude in Adults. *Mind, Brain, and Education*. 2011; 5:42–47.
- Takahashi A, Green D. Numerical judgments with Kanji and Kana. *Neuropsychologia*. 1983; 21:259–263. [PubMed: 6877579]
- Townsend TJ. Serial vs. parallel processing: sometimes they look like Tweedledum and Tweedledee but they can (and should) be distinguished. *Psychological Science*. 1990; 1(1):46–54.
- Townsend JT, Nozawa G. Spatio-temporal properties of elementary perception: An investigation of parallel, serial and coactive theories. *Journal of Mathematical Psychology*. 1995; 39:321–360.
- Townsend JT, Wenger MJ. A theory of interactive parallel processing: New capacity measures and predictions for a response time inequality series. *Psychological Review*. 2004; 111:1003–1035. [PubMed: 15482071]
- Townsend JT, Nozawa G. Serial exhaustive models can violate the race model inequality: Implications for architecture and capacity. *Psychological Review*. 1997; 104:595–602.
- Tzelgov, J.; Ganor-Stern, D. Automaticity in processing ordinal information. In: Campbell, JID., editor. *Handbook of mathematical cognition*. New York: Psychology Press; 2005. p. 55–67.
- Van Opstal, de Lange; Dehaene. Rapid parallel semantic processing of numbers without awareness. *Cognition*. 2011; 120:136–147. [PubMed: 21489415]
- Verguts T, Fias W. Representation of number in animals and humans: a neural model. *Journal of Cognitive Neuroscience*. 2004; 16(9):1493–1504. [PubMed: 15601514]
- Verguts T, Fias W, Stevens M. A model of exact small-number representation. *Psychonomic Bulletin & Review*. 2005; 12(1):66–80. [PubMed: 15945201]
- Verguts, Van Opstal. Dissociation of the size effect and distance effect in one-digit numbers. *Psychonomic Bulletin & Review*. 2005; 12:925–930. [PubMed: 16524012]
- Welford AT. The measurement of sensory-motor performance: Survey and reappraisals of twelve years' progress. *Ergonomics*. 1960; 3:189–230.
- Whalen J, Gallistel CR, Gelman R. Nonverbal counting in humans: the psychophysics of number representation. *Psychological Science*. 1999; 10(2):130–1301.

Figure 1. Presentation stimuli. This figure is a representation of the stimuli that were presented. It is not drawn to scale. The white circle in the center represents the red fixation point. The light and dark gray fonts identify the different diagonals.

Figure 2.

Timeline of simultaneous and successive presentation trials. At the start of each trial, all positions will be covered with a mask (represented here as a pound sign for simplicity). Then integers were presented either (a) simultaneously, or (b) successively for 67 ms in the practice trials, and 56 ms in the experimental trials.

Figure 3.
The numerical distance effect in Experiment 1. It takes longer to compare numbers that are numerically close than it does to compare numbers that are numerically distant.

Figure 4.

The RT numerical distance effect without 7 in Experiment 1. When we remove trials where 7 is either the largest or second largest number there is a strong negative linear relationship between RT and numerical distance.

Figure 5.

The accuracy numerical distance effect without 7 in Experiment 1. When we remove trials where 7 is either the largest or second largest number there is a strong relationship between percent correct and numerical distance. The circles indicate the data points and the crosses indicate the predicted values.

Figure 6. RT distance effect in (A) RT Pressure condition, and (B) control condition in Experiment 2. Participants respond more slowly when the difference between the largest two numbers is small than when the difference is large.

Figure 7.

Accuracy distance effect in (A) RT Pressure condition, (B) control condition, and (C) forced delay condition in Experiment 2. Participants respond less accurately when the difference between the largest two numbers is small than when the difference is large.

Table 1
Experiment 1: Means and SD for accuracy and RT

Response condition	Presentation condition	
	Simultaneous	Successive
Two-Button		
Accuracy	.79 (.41)	.80 (.40)
RT	449 (474)	443 (483)
Number-pad		
Accuracy	.79 (.41)	.81 (.40)
RT	1603 (582)	1611 (584)

Table 2
Experiment 2: Means and SD for accuracy and RT

Response timing	Presentation condition	
	Simultaneous	Successive
Delay		
Accuracy	.79 (.41)	.79 (.41)
RT	423 (329)	426 (336)
Control		
Accuracy	.76 (.43)	.78 (.41)
RT	608 (629)	551 (576)
RT Pressure		
Accuracy	.74 (.44)	.76 (.43)
RT	342 (371)	328 (347)

Table 3
Experiment 3: Means and SD for accuracy and RT

Response timing	Presentation condition	
	Simultaneous	Successive
Control		
Accuracy	.71 (.46)	.74 (.44)
RT	792 (752)	711 (683)
RT Pressure		
Accuracy	.68 (.47)	.70 (.46)
RT	508 (507)	488 (518)

Table 4
RT means and standard deviations for the distance effect in Experiment 3

Difference	RT	Percent Correct
1	644 (650)	.69 (.46)
2	592 (616)	.77 (.42)