Title
Representing Number Sets: Encoding Statistical Properties

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 35(35)

ISSN
1069-7977

Authors
Cravalho, Patrick
Morris, Bradley
Was, Christopher
et al.

Publication Date
2013

Peer reviewed

Representing Number Sets: Encoding Statistical Properties

Patrick F. Cravalho (pcravalh@kent.edu), Bradley J. Morris (bmorri20@kent.edu), Christopher A. Was (cwas@kent.edu)
Department of Educational Psychology, 405 White Hall
Kent, OH 44242 USA

Amy M. Masnick (Amy.M.Masnick@hofstra.edu)
Department of Psychology, Hauser Hall
Hempstead, NY 11549 USA

Abstract

Previous research suggests that people often recall individual items when sets are smaller than four and aggregate set features for sets larger than four. One intriguing possibility is that the process of aggregating sets creates summary representations that maintain the statistical properties of the set itself. For sets of numbers, this process might implicitly create approximate means. We report the results of two experiments investigating memory for number sets and its relation to working memory and metacognitive monitoring. In both experiments, participants were shown a series of data sets that varied in size (4, 6, or 8 numbers) and variance (10% or 20% of the mean) and were then presented with an actual value from the set and the set mean. In Experiment 1, participants were asked to select the actual value, and in Experiment 2, participants were asked to select the set mean. Results indicated that the proportion of correct selections and metacognitive confidence decreased with set size. Working memory was related to performance only when the set size was 6. The findings suggest that participants often erroneously reported the set mean as being a member of the set and that this effect increased for sets larger than four. The findings suggest that the process of aggregating number sets results in approximate means.

Keywords: Number sets; Representation; Metacognition.

Introduction

Data interpretation (i.e., interpreting numbers in context) is critical to everyday reasoning. Numbers are represented in two ways, as an approximate magnitude and as a symbolic value (Dehaene, 2009). Previous research suggests that people compare single numbers as analog magnitudes, meaning that numbers are represented as approximate rather than exact values. Key evidence for this argument is that comparisons are faster and more accurate as the ratio of difference between values increases (Feigenson, Dehaene, & Spelke, 2004; Hyde, 2011). Yet little is known about how people represent numerical data sets. Previous research involving the cognitive processing of number sets suggests that people's comparisons are highly related to the statistical properties of the sets. Specifically, as the ratio of mean differences between number sets increase and variance within each set decreases, comparisons to assess which set has the higher mean become faster and more accurate (Masnick & Morris, 2008). These findings suggest that when asked to make number comparisons, people implicitly (i.e., without deliberate effort) represent the properties of number sets, such as means and variance, rather than information about individual numbers.

Sets vs. Individuals

A well-replicated finding is that people individuate items (i.e., represent individual values within a set) for sets smaller than four (Scholl, 2001). For example, given a set of three dots, participants are likely to remember the individual dots rather than the properties shared by the dots (Airely, 2001). Given sets larger than four, people aggregate items, or average over sets retaining information about set features (Scholl, 2001; Masnick & Morris, 2008). For example, given a set of a dozen dots of various sizes, participants were likely to erroneously recall a dot that represents the average size, rather than correctly recall an actual dot from the set they viewed (Airely, 2001). Earlier research suggested two separate memory stores for individual versus aggregated sets. For example, Feigenson, Dehaene, and Spelke (2004) suggested one system for individuating small values (< 4) and another system for aggregating larger sets. However, recent research (Hyde, 2011) suggests that the difference in representation may be functional (i.e., due to differences in strategy) rather than structural (i.e., separate stores).

The process of aggregating appears to maintain information about the statistical properties of sets. Previous research on category learning provides evidence that aggregating over individual objects results in a prototype, or a most representative set member. Researchers have suggested that the process of aggregating yields "a measure of central tendency" (e.g., an average or prototypical category member; Medin, Altom, & Murphy, 1984, p. 334). Early research into prototypes suggested that when people aggregate across multiple exemplars, they extract a prototypical category member that represents an average of category features (Medin et al., 1984; Nosofsky, Denton, Zaki, Murphy-Knudsen, & Unverzagt, 2012). For example, a robin is often seen as a prototypical bird because it shares most of the features common among birds.

Numbers provide an intriguing extension of this work because aggregation of sets may produce approximate measures of central tendency. Specifically, encoding a set of numbers may implicitly provide information about the set mean. The reviewed research suggests two related effects. One, for small set sizes (< 4), participants will encode individual set members. For larger set sizes (> 4), participants will aggregate across values, losing information about individual values. Two, the process of aggregating values creates summary representations of the number sets, including approximate mean and variance values.

Set Encoding and Working Memory

Working memory (WM) is often described as the processes and mechanisms involved with the maintenance of task relevant information necessary for performance of cognitive tasks (Miyake & Shah, 1999). In the current study, we predict that set representation is more likely to occur as the number sets increase in size because maintenance of all items from larger sets requires greater WM resources than maintenance of items from smaller sets. Therefore, we also predict that participants with larger WM capacity should be able to maintain a greater number of items in WM and correctly identify presented items.

Set Encoding and Metacognitive Monitoring

A secondary focus of this experiment was to investigate the influence of metacognitive monitoring on the representation of number sets. Monitoring refers to one's judgment of their current cognitive processing, specifically, the degree to which a process is deliberate and a rough estimate of one's performance (Dunlosky & Metcalfe, 2009). We assessed monitoring through participant confidence judgments.

Aims and Predictions

This experiment extends the work of Masnick and Morris (2008) by examining how people represent and remember number sets, and how accurately they judge their memory for number sets. Independent variables were number set size (4, 6, and 8) and coefficient of variation (20% and 10% of mean). Dependent variables were number selection accuracy, reaction time, confidence judgments, strategy use, and working memory capacity. We predicted that people would erroneously recall seeing set means or medians rather than actual numbers from the sets as set sizes increased. As suggested by Hyde (2011), different strategies are predicted to be associated with encoding goals. Specifically, encoding individual numbers is expected to be associated with longer reaction times than encoding set aggregates, because the latter is a relatively implicit process. Finally, we predicted that number recall and mean identification would be related to confidence judgments in terms of accuracy, or that they would increase or decrease together.

Experiment 1

Method

Participants. Participants (N = 51) were undergraduate students at a Midwestern state university. The average age was 22.04 (*SD* = 6.84), 84% of the participants were female and most were of Caucasian descent.

Procedure. Experiment 1 consisted of two counterbalanced tasks: the number set task and a working memory task. All participant tasks were presented using E-Prime software. The number set task consisted of 126 trials. Each trial had four parts, described here in order of presentation (see Figure 1a).

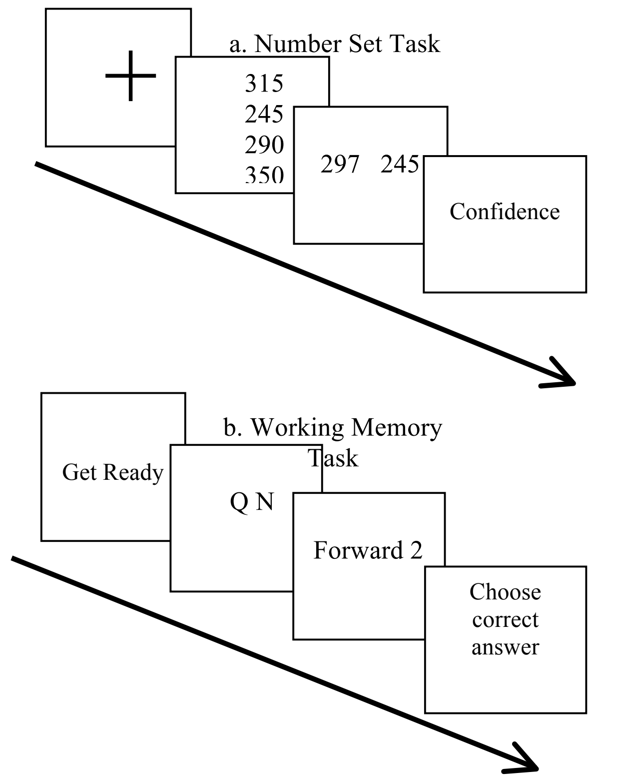

Figure 1. Number set and working memory task procedures.

First, a fixation cross was presented to focus the participant's attention. Then, a number set was shown (details of the sets and presentation durations below). Next, two numbers were presented, an actual value from the set and the mean or median value from the set. The participant was asked to indicate which of the two numbers was in the set presented in the previous slide by pressing a computer key. The two numbers and indication prompt remained on the screen until the participant chose an answer. After responding, participants were asked how sure they were of their answer. Confidence judgment response options were presented as follows: 1) 0-25%, 2) 26-50%, 3) 51-75%, & 4) 76-100%. The response options and prompt remained on the screen until the participant chose an answer.

There were four sets of stimuli, each preceded by a set of instructions explaining the process outlined above to the participant. The first six number sets were practice trials and were not analyzed. Experimental trials included 40 sets of four numbers (each set presented for 2 s), 40 sets of six numbers (each set presented for 2.5 s), and 40 sets of eight numbers (each set presented for 3 s). Each set consisted of three digit numbers. Within each set size, 10 sets were drawn from one of four variance types. Set variance was either 10% (low variance) or 20% (high variance) of the set mean. For half of the choice trials, participants were shown the set mean. For the remaining half, participants were shown the set median; although for sets of 4 the mean and median were identical.

The three blocks of experimental trials were presented sequentially, randomized within-participants. The presentation location of the actual value and mean or median value was also randomized so that the actual value and mean were presented on the left or right side of the screen in 50% of trials. After completing all number set trials, participants were surveyed about their strategy use during the task. Participants were presented nine strategy descriptions (see Table 1) and were asked to estimate how often they used each strategy during the experiment using the following scale: 1) never, 2) some trials, 3) most trials, or 4) always. The response options and prompt remained on the screen until the participant chose an answer.

Table 1: Strategy Descriptions and Examples.

Strategy	Example
Look at the first digit.	The “1” in 125.
Look at the second digit.	The “2 “in 125.
Look at the third digit.	The “5” in 125.
Try to figure out the average.	Calculate mean value.
Find the biggest number.	Scan set for highest value.
Find the smallest number.	Scan set for lowest value.
Just get a sense of the numbers.	Scan set values.
Look for a number that is not like other numbers.	Find any value unlike other values.
Try to memorize specific numbers.	Memorize a few numbers.

Participants were also given the alphabet mathematics working memory task (cf. Was, Rawson, Bailey, & Dunlosky, 2011). The general format of this task was as follows (see Figure 1b): the words “Get Ready” preceded a letter or set of nonadjacent letters (presented for 2.5 s), followed by a prompt containing a transformation direction and number (ex. “Forward 2”). The participant then counted forward or backward from the given letter or set of letters by the number given. The transformation prompt remained on the screen until the participant was ready to choose an answer from a list of eight alternatives. Participants were instructed to solve the problem before advancing to the response alternative screen. Once a participant had advanced to the response alternative screen, a time limit of 6 s was imposed to prevent one from solving the problem while examining the alternatives in the response window. This pattern spanned four practice trials and a set of 24 stimuli, which was preceded by a set of instructions explaining the process outlined above to the participant. Feedback of “Correct!” or “Incorrect” was given for each response. More detailed feedback was given during the practice trials. Specially, if a correct answer was given it was accompanied by a brief explanation of why that answer was correct, but if an incorrect answer was given it was accompanied by a more detailed explanation of how the correct answer would have been attained. The block of 24 stimuli was randomized within-participants.

Results and Discussion

Participants were most accurate (*M* = .79, *SD* = .12) in their responses for sets of four numbers. Accuracy (see Figure 2) then declined as set size increased to six (*M* = .68, *SD* = .10) and again when it increased to eight (*M* = .61, *SD* = .08). Participants were also most confident (*M* = 3.05, *SD* = .52) in their responses for sets of four numbers. Confidence also declined as set size increased to six (*M* = 2.63, *SD* = .54) and again when it increased to eight (*M* = 2.43, *SD* = .63).

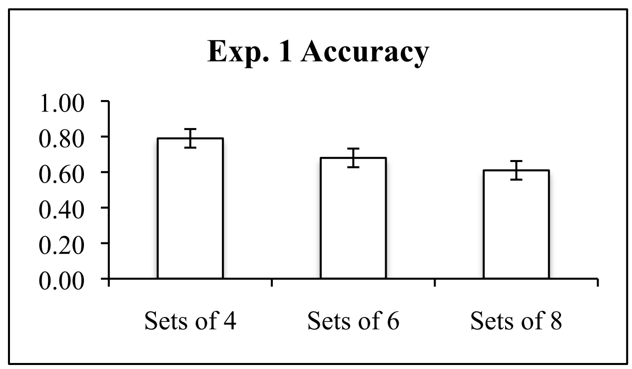

Figure 2. Mean accuracy scores for Experiment 1.

Reaction times (see Figure 3) for sets of four (*M* = 2269.37, *SD* = 585.34) were faster than for sets of six (*M* = 2338.76, *SD* = 620.24), but not faster than for sets of eight (*M* = 2197.69, *SD* = 618.27).

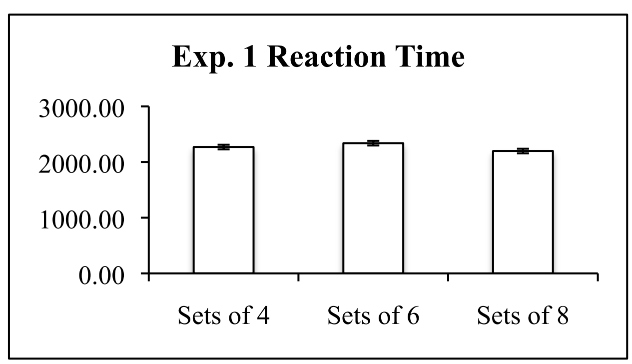

Figure 3. Mean reaction times for Experiment 1.

Repeated measures ANOVA were used to analyze all main variables. There was a significant decline in accuracy as set size increased, $F(2, 104) = 59.94$, $p = .000$, $\eta^2 = .535$. More specifically, accuracy decreased significantly as set size increased from 4 to 6, $F(1, 52) = 39.62$, $p = .000$, $\eta^2 = .432$, and as set size increased from 6 to 8, $F(1, 52) = 20.98$, $p = .000$, $\eta^2 = .288$.

There was a significant decrease in reaction times as set size increased, $F(2, 104) = 5.15$, $p = .007$, $\eta^2 = .090$. While there was no significant difference in reaction time between sets of 4 and sets of 6, $F(1, 52) = 2.48$, $p = .121$, $\eta^2 = .046$, there was a significant difference in reaction time between sets of 6 and sets of 8, $F(1, 52) = 11.38$, $p = .001$, $\eta^2 = .180$.

There was a significant decrease in confidence in one's answer as set size increased, $F(2, 104) = 52.07$, $p = .000$, $\eta^2 = .500$. Contrasts showed significant decreases in confidence as set size increased from 4 to 6, $F(1, 52) = 55.80$, $p = .000$, $\eta^2 = .518$, and as set size increased from 6 to 8, $F(1, 52) = 12.19$, $p = .001$, $\eta^2 = .190$. Zero-order correlations indicated a positive relation between selection accuracy and confidence judgment for set size of 4, $r(52) = .58$, $p = .000$, and set size of 6, $r(52) = .45$, $p = .001$, but not set size of 8, $r(52) = .19$, $p = .173$.

We compared the effect of high and low variance number sets, aggregated across set size. A paired samples t-test indicated that there was greater accuracy when variance in the number set was higher ($M = .72$) than when it was lower ($M = .67$), $t(52) = 4.30$, $p = .000$, $d = 1.181$. However, there was no significant difference in reaction times between high ($M = 2301.25$) and low ($M = 2260.82$) variance, $t(50) = 1.39$, $p = .171$, $d = .393$.

No correlations were found between working memory or any of the variables of interest, so we performed a mean split of the participants based on working memory task performance and conducted independent samples t-tests to compare the groups in terms of accuracy. There was no significant differences between groups based on sets of 4, $t(51) = .46$, $d = .02$, $p =.65$, or sets of 8, $t(51) = .95$, $d = .02$, $p = 35$. However, the group scoring above the mean on the WM task performed better than the group scoring below the mean when set size was 6, $t(51) = 2.08$, $d = .06$, $p = .04$, 95% CI [.002, .114].

The most frequently used strategies were to "just get a sense of the numbers" ($M = 3.19$, $SD = .90$) and to "try to memorize specific numbers" ($M = 2.87$, $SD = 1.01$). These data reflect the general pattern in recall accuracy, as participants likely tried to memorize numbers from the sets of 4, and then were forced to abandon that strategy in favor of "just getting a sense of the numbers" as set sizes exceeded four. This pattern supports the finding of Scholl (2001) that people individuate items for sets smaller than four. One of the least frequently used strategies was to "try and figure out the average" ($M = 1.26$, $SD = .62$). This low rating provides some evidence that the participants were not consciously calculating the mean for the number sets. This pattern is interesting because their response patterns indicate that they chose the mean or median value of the set more often than the actual value from the set as set size increased. This pattern supports findings that people aggregate items for sets larger than four (Scholl, 2001; Masnick & Morris, 2008).

The results demonstrate a clear effect of set size in that participants were more likely to recall the actual value given smaller set sizes (e.g., 4) than larger set sizes (e.g., 8). This finding is consistent with findings using different types of stimuli (e.g., dots) suggesting that encoding individual items is the default operation until the number of items exceeds working memory limits, demonstrated as four elements in previous research (Ariely, 2001; Feigenson et al., 2004). Once working memory limits are reached, the default operation appears to be encoding the set as an aggregate. The data also suggest that the process of aggregating numbers yields as approximate mean. As set size increased, participants were more likely to select the set mean rather than the actual value as the number they had seen in the experiment. Further, when the numbers varied less and were within a smaller range, accuracy also decreased. It became more difficult to distinguish set members from their mean when the set members were more similar.

Working memory capacity also played a role in the ability to identify presented values. When the sets presented contained six items, individuals scoring above the mean on the WM task out performed those scoring below the mean. When the set size was four, WM capacity was not taxed for either group, and when set size was eight, WM capacity was likely exceeded for both groups.

In general, the confidence level of the participants fell along with their performance, although the influence of metacognitive monitoring on number recall was greater for sets of 4 and sets of 6 than for sets of 8. The participants showed accurate recognition that their performance was declining as set size, and task difficulty, was increasing. The lack of correlation between recall and confidence accuracy for sets of 8 may be related to the pattern of strategy use, as this set size would have been too large for one to memorize the numbers presented, as well as large enough to pose difficulty in trying to get a sense of the whole set.

Experiment 2

One possible limitation of Experiment 1 is that the task explicitly asked participants to attend to and encode individual values. Experiment 2 provided the same experimental conditions but changed instructions to ask participants to identify the set mean. This task parallels the Deese-Roediger-McDermott (DRM) paradigm (Deese, 1959; Roediger & McDermott, 1995), as participants will study lists that are comprised of numbers (rather than words) related to a non-presented response target (Sugrue & Hayne, 2006). The change in instructions changes the task demands and likely changes the strategy used for the task. Specifically, if participants are asked to recall set means, they will be less likely to encode individual values. Previous research demonstrates that although people are relatively accurate when estimating set means, estimation accuracy decreases with set size (Peterson & Beach, 1967).

Method

Participants. Participants (N = 27) were undergraduate students at a Midwestern state university. The average age was 20.67 (*SD* = 5.88), 70% of the participants were female and most were of Caucasian descent.

Procedure. The number sets and working memory tasks for experiment two followed the same procedure as for experiment one, except that for number sets participants were asked to indicate which of the two numbers represented the average of the set of numbers (rather than which of the values was actually in the set).

Results and Discussion

Participants were most accurate (*M* = .71, *SD* = .19) in their responses for sets of four numbers. Accuracy (see Figure 4) then declined as set size increased to six (*M* = .62, *SD* = .14) and again when it increased to eight (*M* = .58, *SD* = .13). Participants were also most confident (*M* = 2.76, *SD* = .64) in their responses for sets of four numbers. Confidence also declined as set size increased to six (*M* = 2.37, *SD* = .66) and again when it increased to eight (*M* = 2.15, *SD* = .56).

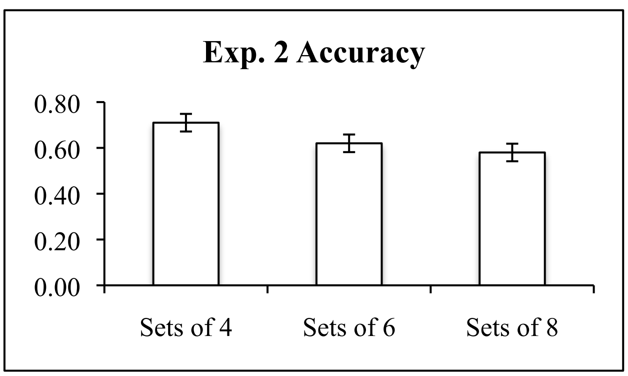

Figure 4. Mean accuracy scores for Experiment 2.

Reaction times (see Figure 5) for sets of four (*M* = 2743.02, *SD* = 673.50) were slower than for sets of six (*M* = 2512.29, *SD* = 951.66). Reaction times for sets of eight (*M* = 2238.00, *SD* = 922.49) were faster than for sets of six.

As in Experiment 1, repeated measures ANOVA were used for all main analyses. There was a significant decline in accuracy as set size increased, $F(2, 52) = 9.00$, $p = .000$, $\eta^2 = .257$. Contrasts showed that accuracy decreased significantly as set size increased from 4 to 6, $F(1, 26) = 7.26$, $p = .012$, $\eta^2 = .218$, but not as set size increased from 6 to 8, $F(1, 26) = 2.32$, $p = .140$, $\eta^2 = .082$.

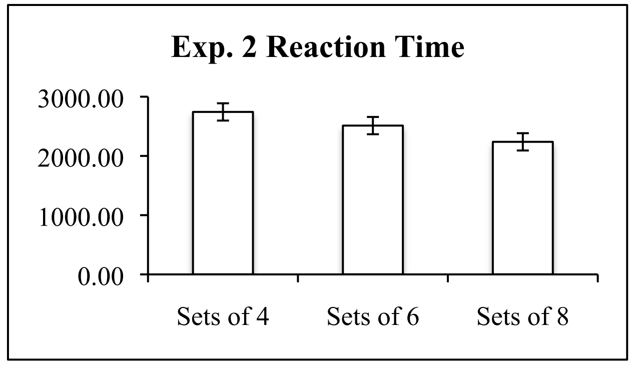

Figure 5. Mean reaction times for Experiment 2.

There was a significant decrease in reaction times as set size increased, $F(2, 52) = 11.01$, $p = .000$, $\eta^2 = .298$. There was no significant difference in reaction time between sets of 4 and sets of 6, $F(1, 26) = 3.76$, $p = .063$, $\eta^2 = .127$, but there was a significant decrease in reaction time between sets of 6 and sets of 8, $F(1, 26) = 13.98$, $p = .001$, $\eta^2 = .350$.

There was a significant decrease in confidence ratings across set size, $F(2, 52) = 22.44$, $p = .000$, $\eta^2 = .463$. Contrasts showed significant decreases in confidence as set size increased from 4 to 6, $F(1, 26) = 18.52$, $p = .000$, $\eta^2 = .416$, and as set size increased from 6 to 8, $F(1, 26) = 8.55$, $p = .007$, $\eta^2 = .248$. Zero-order correlations indicated no relation between identification accuracy and confidence judgment for set size of 4, $r(27) = .07$, $p = .709$, set size of 6, $r(27) = -.07$, $p = .708$, or set size of 8, $r(27) = .33$, $p = .087$.

We compared high and low variance, aggregated across set size. A paired samples t-test indicated that there was no difference in accuracy when variance in the number set was higher (*M* = .64) than when the variance was lower (*M* = .64), $t(26) = .184$, $p = .855$, $d = .072$. However, there was a significant difference in reaction times between high (*M* = 2543.05) and low (*M* = 2439.62) variance, $t(26) = 2.24$, $p = .034$, $d = .879$.

No correlations were found between working memory or any of the variables of interest, so we performed a mean split of the participants based on working memory task performance and conducted independent samples t-tests to compare the groups in terms of accuracy. There was no significant differences between groups based on sets of 4, $t(25) = -1.07$, $d = -.09$, $p = .29$, on sets of 6, $t(25) = -1.50$, $d = .15$, $p = .146$, or sets of 8, $t(25) = -1.03$, $d = .31$, $p = .311$.

The most frequently used strategies were to "just get a sense of the numbers" (*M* = 2.85, *SD* = 1.02) and to "look at the first digit" (*M* = 2.59, *SD* = .93). As with experiment one, most participants tried to get a sense of each number set. Predictably, fewer participants cited memorization (*M* = 2.33, *SD* = 1.07) as a frequently used strategy than those from experiment one, as this would not have been necessary due to the change in answer format. These data reflect the general pattern in identification accuracy, as participant performance stabilized once the sets grew larger than 4 numbers, the point at which one would expect participants to have a better chance at inferring set characteristics rather than trying to focus on memorizing individual numbers. This pattern supports the finding that people individuate items for sets smaller than four, but aggregate items for sets larger than four (Scholl, 2001; Masnick & Morris, 2008).

These results reveal a similar pattern of behavior as in Experiment 1, with a few exceptions. One exception was that variance in set size did not affect accuracy in identifying the mean in the same way it did identifying members of the original set. Another exception is that working memory did not appear to play a role in accuracy, even in the largest sets, when identifying the mean. It is possible that differences between a recognition task and a slightly more inferential task are playing a role, but more data with direct comparisons will be necessary to explore this issue in more detail.

General Discussion

Our results provide new insights into how people represent number sets. One, it appears that number sets and non-numerical sets are aggregated similarly. Number sets of four appear to be accurately encoded as individual values, likely represented as symbolic values. Larger sets appear to be encoded as aggregates and represented as analog magnitudes. It appears that the process of aggregating over sets results in an approximate mean. In Experiment 1, as set size increased, participants were increasingly likely to erroneously select the set mean, rather than the actual value from the set. It appears that even with larger sets, participants quickly approximated means, though less accurately, than for smaller sets.

The results suggested that strategies are important in creating different representations. One intriguing piece of evidence from Experiment 1 is that reaction times for sets of four and eight were shorter than for sets of six. This suggests that participants might have been attempting to use the same individuation strategy with sets of four and six. In correspondence, we found a relation between working memory capacity and performance only for those with high capacities and only for sets of six items. This outcome was not surprising, because the individuation strategy would be more difficult for sets of six than for sets of four. This suggestion is supported by a decline in accuracy and confidence from sets of four to sets of six.

In conclusion, our results suggest that memory for number sets is similar to memory for sets of objects in that sets smaller that four are likely to be individuated while sets larger than four are likely to be aggregated. The process of aggregation appears to maintain approximate representations of the statistical properties of the sets. One possible explanation for the lack of relation between working memory and recall accuracy found in this study is that the tasks themselves load on spatial working memory, rather than verbal memory (Morris & Masnick, 2008).

References

Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science, 12(2)*, 157-162.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58*, 17-22.

Dehaene, S. (2009). Origins of mathematical Intuitions: The case of arithmetic. *The Year in Cognitive Neuroscience 2009: Ann. N.Y. Acad. Sci., 1156*, 232-259.

Dunlosky, J., & Metcalfe, J. (2009). *Metacognition*. London: Sage Publications, Inc.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences, 8(7)*, 307-314.

Hyde, D. C. (2011). Two systems of non-symbolic numerical cognition. *Frontiers in Human Neuroscience, 5*, 1-8.

Masnick, A. M., & Morris, B. J. (2008). Investigating the development of data evaluation: The role of data characteristics. *Child Development, 79(4)*, 1032-1048.

Medin, D. L., Altom, M. W., & Murphy, T. D. (1984). Given versus induced category representations: Use of prototype and exemplar information in classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(3), 333-352.

Miyake, A. & Shah, P. (1999). *Models of Working Memory.* Cambridge, UK: Cambridge University Press.

Morris, B. J., & Masnick, A. M. (2008). Making numbers out of magnitudes. *Behavioral and Brain Sciences, 31(6)*, 662-663.

Nosofsky, R. M., Denton, S. E., Zaki, S. R., Murphy-Knudsen, A. F., & Unverzagt, F. W. (2012). Studies of implicit prototype extraction in patients with mild cognitive impairment and early Alzheimer's disease. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15(46)*, 1-21.

Peterson, C. R., & Beach, L. R. (1967). Man as an intuitive statistician. *Psychological Bulletin, 68(1)*, 29-46.

Roediger, H. L., III, & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 803-814.

Scholl, B. J. (2001). Objects and attention: The state of the art. *Cognition, 80(1/2)*, 1-46.

Sugrue, K., & Hayne, H. (2006). False memories produced by children and adults in the DRM paradigm. *Applied Cognitive Psychology, 20*, 625-631.

Was, C. A., Rawson, K. A., Bailey, H., & Dunlosky, J. (2011). Content-embedded tasks beat complex span for predicting comprehension. *Behavioral Research Methods.* doi: 10.3758/s13428-011-0112-x