Journal of Experimental Psychology:
Human Perception and Performance

0096-1523/13/$12.00 DOI: 10.1037/a0034131

Sets Within Sets: The Influence of Set Membership on Numerical Estimates

Sara Cordes, Alison Goldstein, and Elizabeth Heller
Boston College

Navigating the world requires attention to number; however, sets in real-world contexts are rarely homogeneous or presented in isolation, thus the task of determining what constitutes a relevant set for enumeration can be a difficult one. This contextual ambiguity increases the likelihood that irrelevant sets may bias our ability to accurately track number. In the current study, we investigated whether numerical estimates are influenced by irrelevant set information, such as the size of other present subsets or the number of subsets composing the set. Adult observers were shown brief arrays of dots containing 1 or 2 intermixed subsets, differentiated by color, and were asked to estimate the number of either: (1) 1 of the subsets or (2) the superset (total number of dots). When estimating the size of a subset, numerical estimates were greatly influenced by the size of the other, irrelevant subset, suggesting that the presence of extraneous sets may hinder accurate number judgments. Furthermore, when asked to judge the total number of items (the superset), observers judged supersets comprising 2 subsets as more numerous compared to those comprising only a single subset. Importantly, both trends were apparent even when observers had prior information about the identity of the target set, suggesting adults preattentively parse the world into sets, causing numerical biases that are not under conscious control. Potential explanations for this pattern of results, including simultaneous numerical contrast effects and superset summation, are discussed.

Keywords: numerical estimation, set membership, numerical perception, subset, superset, enumeration

The ability to keep track of number has been shown to be critical for basic functioning (e.g., decision-making, foraging; Gallistel, 1990; Gallistel, Gelman, & Cordes, 2006) as well as mathematical development (e.g., Geary, 2013; Halberda, Mazzocco, & Feigenson, 2008; Siegler & Booth, 2004). However, the task of knowing what constitutes a relevant set for enumeration is rarely simple. Stepping outside, you can count the number of people on the street or the number of trees you pass, among other things. Likewise, you can count the number of living things you encounter. Despite the variety and hierarchical nature of sets one may encounter in the real world (e.g., trees and people are subsets of the superset, living things), numerical estimation research has almost exclusively focused on observer's abilities to estimate the size of a single homogeneous set presented in isolation. We know nothing, however, of how the presence of multiple sets may affect our abilities to make accurate numerical estimations. Does the size of irrelevant sets influence numerical judgments of a target set such that, for example, the number of boys in a classroom affects estimates of the number of girls (and vice versa)? In that same vein, does the makeup of the superset play a role in set size estimation such that, for example, coed and single-sex classes are perceived as different in size simply because of their make-up?

In theory, the presence of other irrelevant sets and/or the make-up of the set in question should not play a role in numerical judgments—number is an abstract quantity, independent of set attributes (Cantlon, Cordes, Libertus, & Brannon, 2009; Cordes, Williams, & Meck, 2007; Jordan & Brannon, 2006). But in practice, our perception of number may behave otherwise. Evidence reveals perceptual, affective, and social factors regularly bias our perception of number (e.g., Baker, Rodzon, & Jordan, in press; Durgin, 1995, 2008; Redden & Hoch, 2009; Young & Cordes, 2013a, 2013b). For example, adults estimate highly dense arrays (i.e., with small inter-item distances) as being more numerous (Durgin, 1995, 2008; but see Allik & Tuulmets, 1991), and both children and adults perceive homogeneous arrays (e.g., all red circles) as being more numerous than heterogeneous arrays (e.g., a mix of red circles, blue squares, green triangles, etc.; Posid, Huguenel, & Cordes, 2013; Redden & Hoch, 2009). Further, the identity of the items to be enumerated also matters; adults report seeing fewer items when arrays are composed of threatening stimuli (spiders) compared to neutral items (flowers; Young & Cordes, 2013b). While the underlying mechanisms driving these numerical biases necessarily differ, these findings do suggest that, at least in some cases, perceptual information may be processed prior to enumeration, leading to subconscious influences on our subsequent perception of number. The current study investigates: Does set information operate at this preattentive level, leading to systematic numerical estimation biases?

Sara Cordes, Alison Goldstein, and Elizabeth Heller, Department of Psychology, Boston College.

Funding was provided by an Alfred P. Sloan Foundation fellowship to Sara Cordes.

Correspondence concerning this article should be addressed to Sara Cordes, Boston College, Department of Psychology, 300 McGuinn Hall, 140 Commonwealth Avenue, Chestnut Hill, MA 02467. E-mail: cordess@bc.edu

Work by Halberda, Sires, and Feigenson (2006) suggests set membership is likely processed prior to enumeration. In their task, adult observers were presented with arrays of dots composed of 1–6 colors and asked either to estimate the total number of dots in the display (the superset) or the size of one of the subsets (e.g., only the red dots). On some trials (probe-before trials), observers knew which set to enumerate prior to seeing the display, and on other trials (probe-after trials), observers did not learn this information until after the display disappeared. Results revealed that when judging the size of an individual subset, error rates in the probe-before and probe-after trials did not differ when arrays contained one or two subsets for the majority of observers, suggesting observers were able to simultaneously enumerate both subsets and the superset. Once the display contained more than two subsets, however, the adults' abilities to enumerate individual subsets deteriorated in the probe-after trials, as they were unable to track set size information for more than two subsets simultaneously. In contrast, when estimating the size of the superset, the number of subsets presented did not matter; estimates in the probe-before and probe-after conditions were similar regardless of the number of colors in the display. The authors suggested that together their results reveal adults are limited to being able to track the numerosity of a total of three sets in parallel[1]—that is, two independent subsets and the superset containing those sets (although there was some variability in this limit across observers). Further, the similar error rates observed across the probe-after and probe-before conditions led Halberda et al. to posit "the notion of a set may operate prior to enumeration" (p. 576). If true, then set information may lead to observable and predictable biases in numerical estimation data.

In the current study, adult observers were presented arrays of dots composed of either a single set or two intermixed sets (sets defined by color). Similar to Halberda et al. (2006), observers were asked across trials to estimate the number of dots belonging either to a particular subset (e.g., number of red dots) or to the superset (e.g., total number of dots, regardless of color). To explore the role of attention, on some trials participants were informed beforehand which set to attend to (probe-before trials) and on other trials, the identity of the target set was only available after the display presentation (probe-after trials). Importantly, we were interested in the influence of two specific types of irrelevant set information: (1) Subset Effects: When enumerating a subset, does the size of other subsets impact judgments (e.g., does the number of blue dots influence estimates of the number of red dots), and if so, how? And, (2) Superset Effects: When enumerating a superset, how does the composition of the set affect numeric estimates (e.g., are supersets composed of two subsets perceived as more or less numerous than those composed of a single subset)?

Subset Effect Hypotheses

Objectively, the size of one set is completely independent of the size of another. However, it is possible that, subjectively, these two variables interact. If sets are attended to automatically and prior to enumeration, then it is reasonable to assume that information about other nearby sets may impact our perception of number. Given that adults can store the size of two subsets simultaneously (Halberda et al., 2006), then one possibility is that estimates of the size of one subset may positively correlate with size of other present, yet irrelevant, subsets. Similar positive correlations between subjective estimates of quantity have been observed in the interval timing literature, such that when more than one duration is stored in memory, estimates are inadvertently combined in memory ("memory mixing") resulting in a temporal distortions (Penney, Allan, Meck, & Gibbon, 1998). While memory mixing has not been documented in the domain of number, work has shown a slight positive correlation between an individual's tendency to under- or overestimate the numerical size of two independent sequential sets (Cordes, Gallistel, Gelman, & Latham, 2007). Thus, it is not unlikely that memory mixing may cause estimates of the size of one subset to be skewed in the direction of the size of another subset, such that observers may perceive more red dots when there are many blue dots, and perceive fewer red dots when there are not as many blue dots.

Alternatively, evidence of simultaneous contrast effects in subjective magnitude estimates suggests that the presence of larger subsets may lead to smaller subset estimates. Contrast effects, demonstrated for many non-numerical quantities (e.g., weights, durations, loudness; Preston, 1936; Melamed & Thurlow, 1971; Mo, 1971), result in the underestimation (or overestimation) of a stimulus magnitude as a result of exposure to a greater (or lesser) magnitude in the same dimension (e.g., a weight is perceived as lighter when held at the same time as another very heavy weight). Although contrast effects have yet to be documented in the domain of number, there is no reason to assume numerical contrast effects are any less likely or prevalent than other magnitude contrast effects. If so, then in contrast to predictions of memory mixing, a simultaneous numerical contrast should result in a negative correlation between subset estimates (e.g., estimates of the number of red dots should be lower as the number of blue dots increases). The current study provides a direct test of these opposing Subset Effect hypotheses.

Superset Effect Hypotheses

Is there any reason to suspect that the make-up of the superset affects numerical estimates? Perhaps there is. As mentioned earlier, evidence suggests that set membership appears to be an important attribute adults attend to even prior to basic enumeration (Halberda et al., 2006). Thus, when set membership is made salient (such as when color identifies subsets), adults may be unable to focus on the superset without first preattentively parsing the set into its component subsets. This preattentive parsing may lead to the employment of alternative strategies for enumeration, such as a "group and add" process (e.g., Trick & Pylyshyn, 1994), in which subsets are enumerated individually and added together to form a superset size estimate. Studies of approximate numerical arithmetic indicate that both infants and adults overestimate sums ("operational momentum"; Knops, Viarouge, & Dehaene, 2009; McCrink, Dehaene, & Dehaene-Lambertz, 2007; McCrink & Wynn, 2009). If so, then this implicit summing should result in higher set size estimates for supersets comprising two subsets compared to those comprising a single subset.

[1] Recent work by Levinthal and Franconeri (2011), however, suggests that this tracking may not be a parallel process—such that each set is instead enumerated in sequence.

On the other hand, subset composition may affect superset estimates, but not because set membership is preattentively processed, but because in order to delineate multiple subsets, supersets must be perceptually heterogeneous (i.e., composed of dots of two colors, e.g., red and blue dots, as opposed to only a single color for a single subset, e.g., all red dots). Both children and adults perceive perceptually heterogeneous sets as being less numerous than homogeneous ones (possibly due to facilitated item individuation; Posid et al., 2013; Redden & Hoch, 2009). Thus, if perceptual variability, as opposed to set membership, is the critical factor, then supersets composed of multiple subsets may be perceived as being less numerous than those composed of a single subset.

In five experiments, we perform the first evaluation of the effects of irrelevant sets and/or set information on adults' numerical estimates. In our first experiment, we provide the first evidence that (1) the presence of other subsets impact estimates of a single subset (Subset Effect) and (2) that set composition of a superset affects superset size estimates (Superset Effect). In further experiments, we explore the mechanisms responsible for these effects. Together, results provide robust evidence to suggest that our subjective perceptions of number are inherently biased by irrelevant set information.

Experiment 1: Subset and Superset Effect

In Experiment 1, we provide the first test of the Subset and Superset Effects. In this study (modeled after Halberda et al., 2006), adult observers were presented with arrays of dots composed of a single subset (all the same color) or of two subsets (each dot was either e.g., red or blue). On some trials, observers were asked to estimate the number of dots contained within a single subset (e.g., the number of red dots) and on other trials, they were asked to estimate the size of the superset (i.e., the total number of dots on the screen, regardless of their color). As in Halberda et al. (2006), on 50% of trials, observers were told beforehand which set they would be asked to enumerate (Probe-Before trials); on the other 50% of trials, they were not told until after the display had disappeared (Probe-After trials).

Method

Participants. Thirty undergraduate students participated in this a single 1-hr session ($M = 19.9$ years, range 18–22 years; 19 females) for course credit.[2]

Procedure. The experiment was run in a quiet room in the laboratory. Participants sat approximately 41.5 cm from a 22" computer screen (resolution 1440 × 900) and indicated their responses on a keyboard. Each participant wore JVC noise cancelling headphones (HA-NC80).

Participants were instructed not to attempt to count. Each Probe-After trial began with the participant viewing a fixation cross (20 pts in diameter) on a black background (500 ms), followed by a visual mask (500 ms), the stimulus display (500 ms), and another visual mask (500 ms—see Figure 1). Stimuli consisted of a dot array of either a single color (e.g., all red dots), or of two interleaved dot arrays (e.g., some red and some blue).

After the second visual mask, participants were prompted with "How many [target set] dots were there?" in the center of the screen. On subset trials, the target set referred to the color of one of the subset arrays from the stimulus display (e.g., "How many red dots were there?"). On superset trials, the target set included all dots on the screen (i.e., "How many total dots were there?"). The question stayed on the screen until the participants typed in a response and selected "next" to proceed to the next trial. If the participant missed seeing the array presentation, they were instructed to type "xxx." These trials were excluded from analyses.

Probe-Before trials (50% of trials) were identical to Probe-After trials except the identity of the target set was displayed during the initial fixation cross presentation at the beginning of the trial. This led to four trial types (Probe-Before Superset, Probe-After Superset, Probe-Before Subset, and Probe-After Subset), randomly intermixed throughout the session. Subjects participated in four practice trials followed by a total of 252 trials (144 subset and 108 superset trials).

Stimuli. All displays were arranged on a black background. Stimulus array colors were randomly chosen from the following: red, yellow, blue, brown, purple, green, orange, and white, with the exception that certain color combinations (orange and red; orange and brown; orange and yellow; yellow and green) were not presented together to ensure the two subsets were maximally discernible. Each color was equally likely to be chosen for the Target array as it was to be chosen for the Distractor array, and all colors had an equal probability of being chosen on any trial.

All dots were 60 pixels in diameter, approximately 2.1 degrees in visual angle. The computer program generated dot arrays on each trial by randomly placing dots within an invisible 13 × 8 grid centered on the screen (each box on the grid was 100 × 100 pixel^2 (3.65 × 3.65 degrees) taking up a majority of the 1440 × 900 screen). To maximize the appearance of randomness, columns in adjacent rows were offset slightly by adding random jitter to the horizontal placement of the columns in each row and the location of the 60-pixel diameter dot within the 100 × 100 pixel^2 box on the grid was also randomly determined. Following the array presentation, the visual mask consisted of a single array of multicolored overlapping dots of varying sizes. The same visual mask was used on all trials.

Subset trials. The number of dots in the target subset array was chosen from one of six logarithmically spaced set sizes (12, 15, 18, 22, 28, or 34). The size of the other subset (the nontarget array or "distractor" array) was randomly chosen from one of four set sizes: (1) 0 (No Distractor trials), (2) 50% fewer than target subset (Small Distractor trials), (3) Same number as target (Equal Distractor trials), or (4) 50% larger than target subset (Large Distractor trials). This led to a total of 6 (target size) × 4 (distractor size) × 2 (probe order) types of Subset trials, each presented three times each for a total of 144 subset trials.

Superset trials. The number of dots on Superset trials was chosen from one of six logarithmically spaced set sizes (24, 32, 42, 54, 72, or 94 dots). There were three types of Superset trials, based upon the division of the superset, in which participants saw either: (1) a single array of the same color (100/0 trials), (2) two subsets of equal size (50/50 trials), or (3) two subsets of unequal size, with the superset divided up as 75%/25% between the two subsets

[2] Two additional subjects were excluded due to computer error.

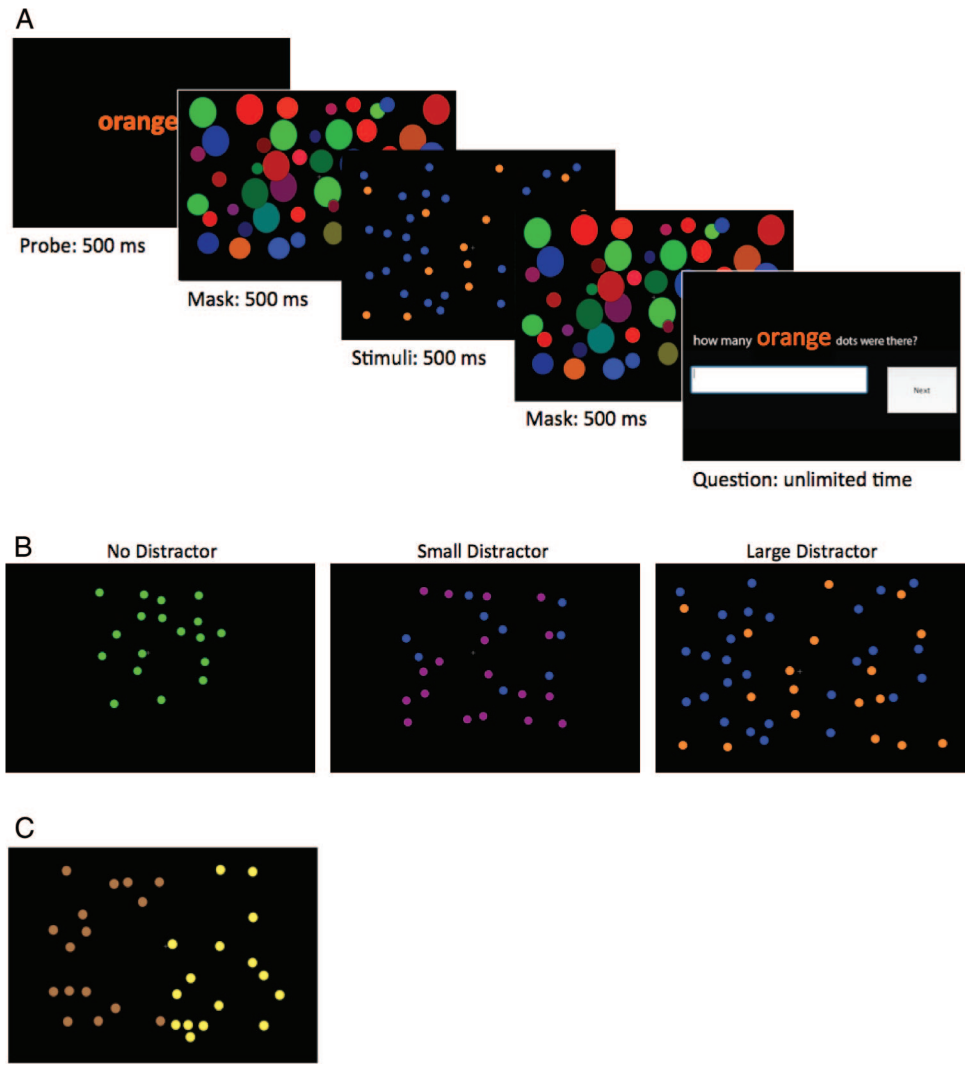

Figure 1. (A) Depiction of events in a Probe-Before Subset trial. In Probe-After trials, the initial screen was left blank. (B) Example arrays from Experiment 3 (target set size 18), in which the size of the background was changed in order to keep dot density constant across the Large Distractor (left, target = orange dots), Small Distractor (middle, target = purple dots), and No Distractor (right, target = green dots) arrays. (C) Example of a spatially separated array from Experiment 5.

(75/25 trials; e.g., 8 red dots and 24 yellow dots).[3] Each of the 6 (set size) × 3 (division) trials × 2 (probe order) trials was presented three times each for a total of 108 Superset trials.

Data analyses. Across all experiments reported in this paper, responses greater than three standard deviations from the mean, "xxx" responses, and inordinately small response (five or fewer for Superset trials; three or fewer for Subset trials[4]) were excluded from analyses (<2%). Subset and Superset trials were analyzed independently.

Results[5]

Subset trials. Mean estimates from Subset trials were subjected to a 6 (target set size) × 4 (distractor size) × 2 (probe order) repeated measures ANOVA and follow-up LSD post hoc tests. Not surprisingly, a significant effect of target, $F(2.4, 70.2) = 308.9$, $p < .00$ revealed mean estimates increased with the size of the target subset, indicating that subjects were engaged in the numerical estimation task. Interestingly, a significant main effect of distractor size was found, $F(3, 87) = 41.0$, $p < .00$, such that target size estimates were highest when it was the only set presented (No Distractor trials, $M = 18.03$, $SE = .40$), and decreased systematically as the distractor set size increased (Small Distractor $M =$

[3] The two different divisions were included so as to ensure subjects did not invoke strategies such as attending to only a single subset and doubling their estimate, and so forth.

[4] We used different allowable minimum responses for Subset compared to Superset trials based on the minimum set sizes presented in each of these types of trials.

[5] For all data analyses reported, when assumptions of sphericity were violated, Greenhouse-Geisser adjustments are reported.

16.59, SE = .46; Equal Distractor M = 15.86, SE = .37; Large Distractor M = 14.81, SE = .35; all differences $p < .02$). Furthermore, a target × distractor size interaction, $F(6.7, 193.4) = 6.4$, $p < .001$ revealed that as target set sizes increased, so did the magnitude of this distractor effect (see Figure 2). No other significant effects or interactions were obtained ($p > .05$). Importantly, probe order did not interact with distractor ($p > .25$), indicating that the size of the distractor set had the same effect regardless of whether the participants knew which set to attend to beforehand (Probe Before: No Distractor M = 17.7, SE = .41; Small Distractor M = 16.4, SE = .41; Equal Distractor M = 15.8, SE = .33; Large Distractor M = 14.4, SE = .33; Probe After: No Distractor M = 18.2, SE = .43; Small Distractor M = 16.4, SE = .43; Equal Distractor M = 15.9, SE = .51; Large Distractor M = 15.0, SE = .49; see Figure 3).

Superset trials. Superset trial mean responses were subjected to a 6 (superset size) × 3 (division) × 2 (probe order) repeated measures ANOVA and follow-up t tests. A significant superset size effect, $F(1.2, 36.0) = 175.7$, $p < .001$ revealed mean estimates increased as a function of set size as expected. Additionally, a main effect of division, $F(1.6, 46.9) = 7.2$, $p < .01$ indicated estimates of supersets comprised of a single subset (100/0 trials; M = 35.5; SE = 1.1) were significantly lower than those of two subsets (50/50 trials: M = 36.8, SE = 1.2 $t(29) = 2.8$, $p < .01$; 75/25 trials: M = 36.9, SE = 1.3, $t(29) = 3.0$, $p < .01$; Figure 4). There was no difference in estimates between the 50/50 and 75/25 division trials, $t(29) = .12$, $p > .9$. Lastly, a superset size × division interaction, $F(6.5, 188.7) = 2.11$, $p = .049$ indicated that these numerical biases decreased as a function of the size of the superset. No other significant main effects or interactions were found. Again, the probe order × division interaction was not significant ($p > .4$), revealing that prior knowledge that participants should attend to the superset and ignore the subsets did not diminish the impact of set composition on estimates (Probe Before: 100/0 M = 35.6, SE = 1.2; 50/50 M = 36.7, SE = 1.3; 75/25 M = 37.1, SE = 1.3; Probe After: 100/0 M = 35.3, SE = 1.1;

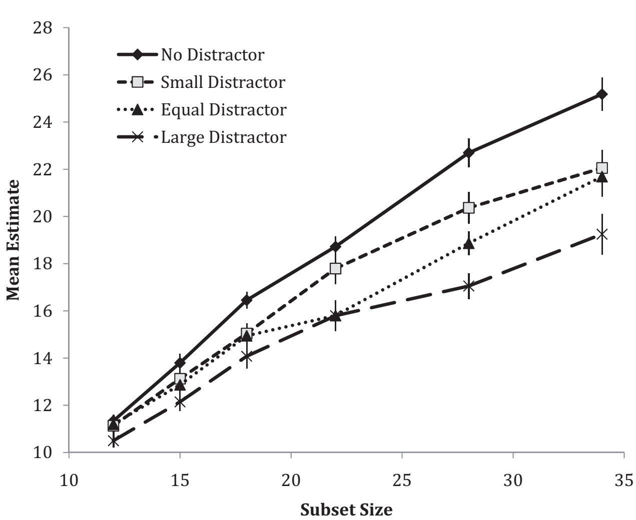

Figure 2. Mean estimates of subset sizes as a function of both set size and the size of the distractor subset from Experiment 1. Error bars depict standard errors.

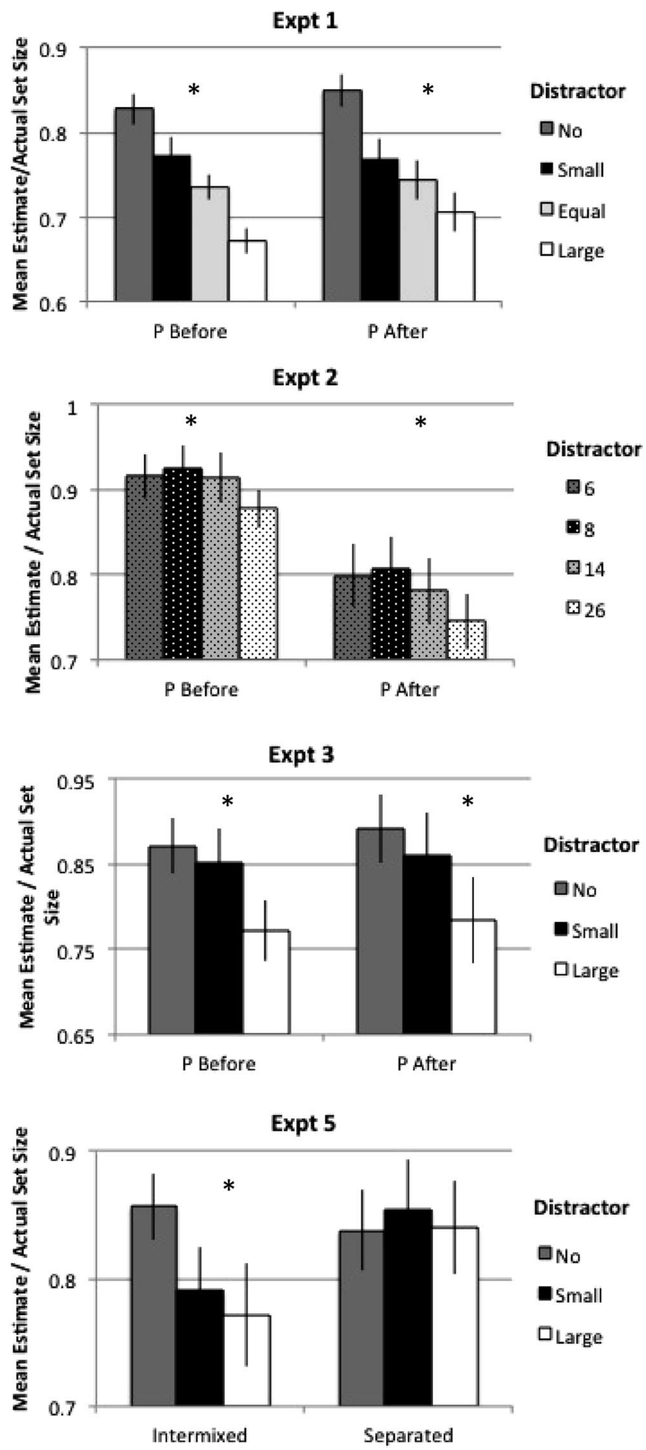

Figure 3. The ratio of mean subset size estimates to the actual size of the subset as a function of the size of the Distractor set for Experiments 1, 2, 3, and 5. Although none of the experiments revealed a Distractor size × Probe-Order interaction, graphs depict data from the Probe-Before and Probe-After trials separately for Experiments 1–3 to illustrate the similar patterns were obtained in both condition. The bottom graph (Experiment 5) presents the data from the spatially intermixed and separated trials separately (all trials are Probe-After). Asterisks indicate a significant Subset Effect and error bars depict standard errors.

50/50 $M = 36.9$, $SE = 1.4$; 75/25 $M = 36.6$, $SE = 1.2$; see Figure 5).

Discussion

Results of Experiment 1 reveal that irrelevant set membership had a significant impact on set size estimates. In line with predictions of a simultaneous numerical contrast effect, when estimating the size of a subset, participants produced lower estimates as the size of the other set in the display increased, even though the other set was completely irrelevant (Subset Effect). Furthermore, when estimating the total number of dots present, set make-up influenced set size estimates, with sets composed of two subsets being estimated as larger than sets composed of a single subset (Superset Effect), suggestive of a "group and add" enumeration strategy. Interestingly, these effects did not appear to be under the subject's conscious control, as they were equally likely to occur whether or not the subject knew the identity of the target set prior to viewing the stimuli.

In the next four experiments, we examine the robustness of these effects to determine whether they persist despite reduced trial-to-trial variability in task demands (Subset Effect−Experiment 2; Superset Effect−Experiment 4), controls for item density (Subset Effect−Experiment 3), reduced salience of the subset composition (Superset Effect−Experiment 4), and decreased visual search demands (Subset and Superset Effects−Experiment 5).

Experiment 2: Subset Effect With Reduced Trial-to-Trial Variability

In Experiment 2, we replicate the Subset Effect with a different sample of individuals. Importantly, to ensure that this effect was not driven by trial-to-trial variability in task demands, subjects were randomly assigned to either a Probe-Before or a Probe-After condition; thus, probe order was manipulated between subjects as opposed to within subjects. In addition, to decrease perceptual variability of the displays across trials, all stimulus displays contained two subsets.[6]

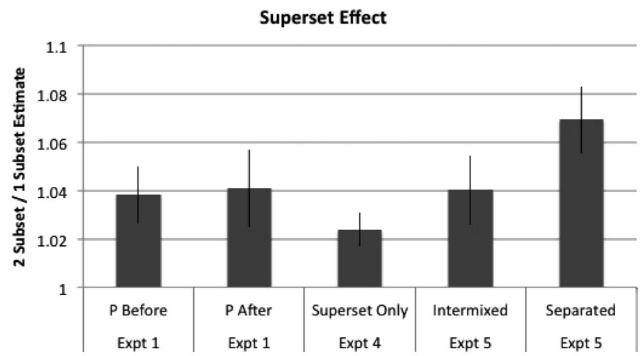

Figure 5. The ratio of mean estimates for supersets composed of two subsets to estimates for supersets only containing a single subset for Experiments 1, 4, and 5. Again, although no effect of probe order was obtained, data from Probe-Before and Probe-After trials are depicted separately for Experiment 1. Error bars depict standard errors.

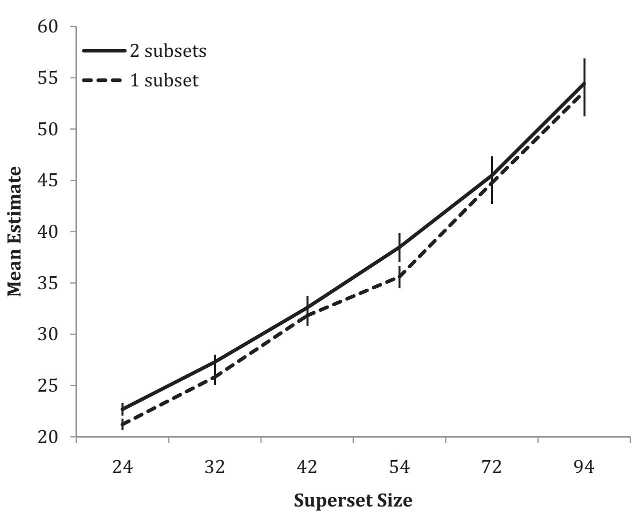

Figure 4. Mean superset estimates as a function of superset size and the number of subsets from Experiment 1. The dashed line depicts estimates of supersets consisting of only a single subset (100/0 trials) and the solid line depicts estimates of supersets consisting of two subsets (average of 50/50 and 75/25 trials). Error bars depict standard errors.

Although participants were asked to estimate the total number of dots in the display on some Superset trials in order to minimize the chances that they attended to only a single subset on every trial, data from these Superset trials were not analyzed because all trials involved two subsets, and thus were unable to speak to the existence of a Superset Effect. Instead, we address the Superset Effect in Experiments 4 and 5.

Method

Participants. Forty undergraduate students participated in this a single 1-hr session ($M = 19.1$ year, range 18–22 years; 30 females) for course credit.[7] None had participated in Experiment 1.

Procedure. The procedure of Experiment 2 was identical to that of Experiment 1 with the following exceptions. In Experiment 2, all stimulus displays consisted of two interleaved dot arrays (e.g., red and blue dots intermixed) so as to reduce trial-to-trial variability in dot arrays. In contrast to Experiment 1 in which probe order was manipulated within-subjects, in Experiment 2, probe order was manipulated between subjects so as to reduce trial-to-trial variability in task demands. Subjects were randomly assigned to one of two conditions—Probe-Before or Probe-After. In the Probe-Before condition, the identity of the Target set (the set to be enumerated) was presented to the participants prior to stimulus presentation on every trial. In the Probe-After condition, however, participants were only made aware of the identity of the target set after the stimulus display on every trial. Thus, target set identity information provided to participants was consistent across trials for all participants; participants in the Probe-Before condi-

[6] With this design we were able to explore the impact of distractor set size on target set size estimates. However, since target sets were never presented in the absence of a distractor set, in Experiment 2b we did not assess whether merely the presence of a distractor set affected estimates (as in Experiment 1). This was again explored in Experiments 3 and 5.

[7] One additional subject was excluded due to failure to complete the task.

tion always knew the identity of the target set prior to stimulus presentation, whereas those in the Probe-After condition did not have this information until after the stimulus display had disappeared.

Again, Subset and Superset trials were randomly intermixed throughout the session. Subjects participated in three practice trials followed by a total of 288 trials (192 Subset and 96 Superset trials). However, because all Superset trials contained two subtests and thus were not specifically designed to explore the Superset Effect, data from these trials were not analyzed and will not be discussed.

Stimuli. The number of dots in the Target subset array was chosen from one of six logarithmically spaced set sizes (6, 8, 11, 14, 19, 26). The size of the other subset (the Nontarget array or "Distractor" array) was randomly chosen from one of four possible set sizes (6, 8, 14, 26).[8] This led to a total of 6 (target size) × 4 (distractor size) types of Subset trials, each presented eight times each for a total of 192 subset trials.

Results

Mean estimates from Subset trials were subjected to a 6 (Target set size) × 4 (Distractor size) × 2 (Probe Condition) mixed measures ANOVA and follow-up LSD post hoc tests. Not surprisingly, a significant effect of Target, $F(1.8, 69.2) = 457.0$, $p < .00$ revealed mean estimates increased with the size of the Target subset, indicating that subjects were engaged in the numerical estimation task. Importantly, a significant main effect of Distractor size was found, $F(2.4, 92) = 17.5$, $p < .00$, such that Target size estimates generally decreased as the Distractor set size increased (Distractor Size 6: $M = 12.0$, $SE = .32$; Size 8: $M = 12.12$, $SE = .32$; Size 14: $M = 11.87$, $SE = .34$; Size 26: $M = 11.36$, $SE = .28$; Figure 3). In addition, a main effect of Probe Condition, $F(1, 38) = 8.1$, $p = .007$ revealed that participants in the Probe-Before condition ($M = 12.7$, $SE = .44$) produced higher estimates overall compared to those in the Probe-After condition ($M = 11.0$, $SE = .44$), and a Probe-Condition × Target set size interaction, $F(5, 190) = 10.1$, $p < .000$ revealed that the difference in estimates across the two conditions was most dramatic for the largest Target set sizes. Importantly, however, Probe Order did not interact with Distractor ($p > .6$), indicating that the size of the Distractor set had the same effect regardless of whether the participants knew which set to attend to beforehand. No other significant effects or interactions were obtained ($p > .1$).

Discussion

Results from Experiment 2 mimic those of Experiment 1. In particular, despite reduced variability in task demands, set size estimates decreased systematically as a function of the size of other present subsets such that, for example, the number of red dots influenced the perception of the number of blue dots. Furthermore, this Subset Effect was found when participants knew beforehand which subset to attend to, and data patterns were identical to those found when participants did not have prior knowledge of the identity of the target set. The lack of a significant Probe-Order × Distractor interaction in either Experiment 1 or 2 provide strong evidence to suggest that the size of other sets impacts target set estimates at a preattentive level. That is, the Subset Effect does not appear to be under the subject's conscious control.

Experiment 3: Subset Effect With Density Control

Results of Experiments 1 and 2 reveal that larger extraneous subsets lead to lower estimates of the target subset. These results are consistent with predictions of a simultaneous numerical contrast, suggesting that the size of the distractor set contrasted with estimates of the target subset. There is an alternative explanation, however, that has yet to be addressed. Because the previous experiments did not control for dot density, then as the number of distractor dots increased, so did the item density of the display. Studies suggest that numerical estimates are greatly influenced by density; although numerical estimates have typically been shown to positively correlate with density (that is, as density increases, estimates generally increase too; Durgin, 1995, 2008), in some cases, estimates have been shown to decrease as dot occupancy increases (Allik & Tuulmets, 1991). If our Subset Effect is best accounted for by a perceptual bias brought on by increased dot occupancy, then keeping dot density constant across displays should make this effect disappear. Alternatively, if a simultaneous numerical contrast effect more accurately described the observed pattern of results, then a Subset Effect should still be observed despite density controls. To test this hypothesis, in Experiment 3, the background size of each display varied as a function of the total number of dots in the display. In doing so, item density was held constant across varying Distractor subset sizes, eliminating dot density as a potential cue for estimation.

Method

Participants. Twenty-one undergraduate students participated in this a single 1-hr session ($M = 20.1$ year, range 18–22 years; 14 females) for course credit or $5 compensation.

Procedure. The procedure of Experiment 3 was identical to that of Experiment 1 with the following exceptions. To reduce task demands, no Superset trials were presented—all trials required participants to estimate the size of one of the subsets in the array. As in Experiment 1, half of the trials were Probe-Before, while the other half were Probe-After (randomly intermixed). Subjects participated in four practice trials followed by a total of 150 trials.

Stimuli. The number of dots in the target subset array was chosen from one of six logarithmically spaced set sizes (12, 15, 18, 22, or 28). The size of the distractor subset was randomly chosen from one of three set sizes: (1) 0 (No Distractor trials), (2) 50% fewer than target subset (Small Distractor trials), or (4) 50% larger than target subset (Large Distractor trials). This led to a total of 5 (target size) × 3 (distractor size) × 2 (probe order) types of Subset trials, each presented 5 times each for a total of 150 subset trials.

To control for density, the size of the background on which the dots were placed varied on every trial based on the total number of dots in the display (total of target + distractor). Large Distractor trials were randomly placed within the same 13 × 8 invisible background grid as in Experiment 1. However, because the total

[8] Note, in contrast to Experiment 1, distractor set sizes were chosen as one of four absolute set sizes, and did not vary as a function of a proportion of the target set size.

number of dots in Small Distractor trials was 60% that of Large Distractor trials, the size of the grid was reduced by 60% on Small Distractor trials to a 9 × 7 grid (thus keeping the ratio of the number of dots to background size constant). Similarly, the grid on No Distractor trials was reduced by 40% (compared to Large Distractor) to 7 × 6. All grids were centered on the screen. Thus, although density continued to covary with target set size (because the number of dots in the display increased with target set size yet the size of the grids did not), within each target size, density was held constant across No Distractor, Small Distractor, and Large Distractor trials (see Figure 3).

Results

Mean responses were subjected to a Target set size (5) × Distractor Size (3) × Probe-Order (2) repeated measures ANOVA. In addition to the main effect of Target Set Size, $F(1.4, 30.0) = 153.6$, $p = .00$, an effect of the size of the Distractor set, $F(1.5, 30.1) = 18.14$, $p = .00$ and a Target x Distractor interaction, $F(4.6, 92.5) = 7.2$, $p = .00$ were again found (see Figure 3). Despite controls for dot density, just as in Experiments 1 and 2, target subset size estimates decreased as the size of the Distractor subset increased (No Distractor: $M = 16.75$, $SE = .69$; Small Distractor: $M = 16.27$, $SE = .84$; Large Distractor $M = 14.78$, $SE = .79$) and the interaction revealed that this Subset Effect was most exaggerated for the largest set sizes. As in Experiment 1, Probe Order did not affect estimates ($p > .3$), nor did it interact with any other variables (p's$>.1$).

Discussion

Again, data reveal that subset estimates are affected by the presence and size of other irrelevant subsets. Findings held despite strict controls for dot density, suggesting that a simple perceptual explanation based upon increased dot occupancy cannot accurately account for the observed Subset Effect. Instead, evidence of a negative correlation between Distractor set sizes and mean estimates are consistent with claims of a simultaneous numerical contrast effect, such that numerical information subjectively contrasts with numerical percepts. Lastly, in all three experiments, the timing of when observers learned the identity of the target set did not affect data patterns, providing strong support that these simultaneous numerical contrasts are not under the observer's conscious awareness.

Experiment 4: Superset Effect With Reduced Set Membership Saliency

In Experiment 4, we return to the Superset Effect to explore the robustness of the finding that supersets comprising two subsets are perceived as larger than those comprising a single subset. Notably, although participants in Experiment 1 could have ignored the composition of the Superset on Probe-Before Superset trials, data suggest that they did not do so. Again, this finding may have been driven by an inability to attend to or process the information provided in the first probe due to exorbitant trial-to-trial variability in task demands (i.e., ignoring the Probe-Before information). Thus, in Experiment 4, we simplified task demands by making subset membership (as defined by color) completely irrelevant to the task, thereby reducing trial-to-trial variability in task demands as well as eliminating dot color as an informative or relevant cue in the task. A second group of naïve adult observers were asked to estimate the total number of dots presented on the screen on every trial (identical to Probe-Before Superset trials of main experiment). Critically, participants were never asked to enumerate sets based upon color, and were always asked to estimate the total number of dots presented. Thus observers knew to ignore the composition of the set (the colors of the dots) and only attend to the total number.

Method

Participants. Thirty undergraduate students ($M = 19.6$ years; range 18–22 years; 15 females) participated in a single 1-hour session for course credit.[9] None of the observers from the other experiments participated in this experiment.

Materials and procedure. The materials and procedure were identical to Experiment 1 with the following exceptions. All trials consisted of Probe-Before Superset Trials. That is, on every trial, participants were asked to indicate the total number of dots presented, regardless of the number of subsets. Thus, the color of the dots (and thus subset membership) was never relevant to task demands.

Superset sizes were chosen from eight possible values (18, 24, 30, 38, 46, 60, 76, or 96). Again, there were three types of set divisions: 100/0 (single subset), 50/50 (dots evenly divided among 2 colors), and 75/25 (25% of dots were a different color). After three practice trials, each superset size × division combination was presented six times each for a total of 144 trials.

Data analyses. Data analyses focused on the comparison of estimates for supersets containing a single subset versus two subsets. Thus, for both Experiments 4 and 5, estimates on 75/25 division and 50/50 division trials were averaged to create a Two Subset mean response for each superset size, for each subject.

Results

Mean responses were subjected to a Superset size (8) × Number of Subsets (2) repeated measures ANOVA. Again, a main effect of set size was found, with mean estimates increasing in direct proportion to set size, $F(1.3, 37.4) = 181.9$, $p = .00$ indicating adults were engaged in the task. Importantly, just as in Experiment 1, the number of subsets affected estimates, $F(1, 29) = 7.3$, $p < .02$ indicating that even when color was completely irrelevant to the task and participants knew only to track the total number of dots in the display, participants estimated sets containing two subsets ($M = 40.0$, $SE = 1.6$) as having more elements than sets containing only a single set

[9] Two additional subjects were excluded for failing to follow task directions, as their mean estimates were inordinately large (estimating ~50 for the smallest superset size of 18) and did not increase monotonically with set size.

($M = 39.1$, $SE = 1.6$; Figure 5). No other significant main effects or interactions were found ($p > .05$).[10]

Discussion

Experiment 4 results replicate the Superset Effect of Experiment 1, revealing that set membership has robust impacts on numerical estimates. Participants systematically estimated supersets containing two subsets as greater in size than supersets containing only a single subset. Importantly, set composition affected estimates even when subset membership (i.e., dot color) was completely irrelevant to the task at hand, providing strong evidence to suggest that identification of set membership occurs automatically, prior to basic numerical processing.

Experiment 5: Reduced Visual Search Demands

In our last experiment, we explored what role, if any, the task of set delineation played in our observed Subset and Superset Effects by varying the spatial arrangement of the two subsets. In our previous experiments, when displays contained more than one subset, the subsets were always intermixed throughout the displays, requiring observers to first visually search through the display in order to identify all members of a given subset (e.g., search through all the red dots to find the blue dots and vice versa). It is unclear whether this visual search contributed to the pattern of results. In Experiment 5, we presented some trials in which subsets were spatially separated, such that one subset was presented entirely on the left side of the screen and the other subset was presented entirely on the right side (Figure 1C). In doing so, spatial separation of the subsets facilitated set individuation, highlighting subset membership, thereby reducing visual search demands. It was expected that highlighting subset membership in this way would enhance the impact of set membership on superset estimates—that is, we expected a magnified Superset Effect. On the other hand, highlighting set membership could make it easier for observers to keep subset size estimates independent in memory, thereby reducing the likelihood of numerical contrast effects on subset trials. If so, then this would predict a reduction or elimination of the Subset Effect. We explored these distinct hypotheses in Experiment 5.

Method

Participants. Twenty-two undergraduate students ($M = 19.8$ years; range 18–22 years; 13 females) participated in a single-half hour session for course credit.[11] None of the observers from the prior experiments participated in this experiment.

Materials, stimuli, and procedure. The materials, stimuli, and procedure were identical to Experiment 1 with the following exceptions. All trials consisted of Probe-After Trials. On half of the trials, dot arrays were randomly intermixed across the computer screen, as in previous experiments. On the other half of trials, however, arrays were spatially separated on the left and right side of the computer screen such that, for example, all red dots appeared on the left side and all blue dots appeared on the right side. Spatially intermixed and spatially separated trials were randomly intermixed throughout the session. To accommodate the different spacing on spatially separated trials, the size of all individual dots was reduced slightly to approximately 1.7 degrees of visual angle. After six practice trials, subjects participated in a total of 264 trials (144 subset and 120 superset).

Subset trials. The number of dots in the target subset array was chosen from one of four logarithmically spaced set sizes (16, 18, 22, 28). The size of the distractor array was randomly chosen from one of three set sizes: (1) 0 (No Distractor trials), (2) 50% fewer than target subset (Small Distractor trials), or (3) 50% larger than target subset (Large Distractor trials). This led to a total of 4 (Target size) × 3 (Distractor size) × 2 (Spatial arrangement—together or separate) types of Subset trials, each presented six times each for a total of 144 subset trials.

Superset trials. Superset sizes were chosen from four possible values (24, 32, 42, 54). Again, there were three types of set divisions: 100/0, 50/50, and 75/25. On half of all trials involving two subsets (50/50 and 75/25 conditions), subsets were spatially separated, and on the other half, subsets were randomly intermixed across the screen. Because the design of spatially intermixed 100/0 trials (involving a single subset) was identical to that of spatially separated 100/0 trials (making these trials indistinguishable), participants were only presented half as many 100/0 trials and data were used as a baseline measure in both the spatially separate and spatially intermixed analyses. Thus, subjects experienced 24 100/0 trials (four superset sizes presented six times each), 48 50/50 trials and 48 75/25 trials (four superset sizes × two spatial arrangements presented six times each) randomly intermixed, for a total of 120 trials.

Results

Subset trials.

Spatially intermixed trials. Mean responses to all spatially intermixed subset trials were subjected to a 4 (Target set size) × 3 (Distractor set size) repeated measures ANOVA. A main effect of Target set size, $F(3, 63) = 93.4$, $p < .000$, a main effect of Distractor set size, $F(2, 42) = 8.1$, $p = .001$, and a significant interaction, $F(6, 126) = 7.9$, $p < .000$ were found, again revealing that the size of the Distractor had a greater impact as the size of the Target set increased. Importantly, the pattern of results from the spatially intermixed subset trials replicated those of Experiments 1–3, with larger Distractor set sizes resulting in lower Target set size estimates (No Distractor: $M = 18.0$, $SE = .55$; Small Distractor: $M = 16.6$, $SE = .72$; Large Distractor: $M = 16.2$, $SE = .85$; Figure 3).

Spatially separated trials. An identical ANOVA on data from the spatially separated trials revealed a main effect of Target set size, $F(1.6, 34.2) = 91.9$, $p < .000$ and no other significant main effects or interactions ($p > .4$). It is important to note that the

[10] Although not significant, the target set size × division interaction approached significance, $F(5.0, 143.7) = 2.0$, $p < .08$ but this interaction was in the opposite direction as that found in Experiment 1, such that the effect of set division was most evident for the largest set sizes. Although it is not clear why results differed across the two experiments, it is likely that the lack of inclusion of subset trials in the design of Experiment 4 may have contributed to the direction of this effect.

[11] Two additional subjects were excluded for failing to follow task directions ($n = 1$), and for providing a large number (9) of mean estimates that were higher than the mean of all responses by more than three standard deviations ($n = 1$).

Subset Effect disappeared, such that the main effect of Distractor set size was not significant ($p > .6$), with estimates from the No Distractor ($M = 17.6$, $SE = .66$), Small Distractor ($M = 18.0$, $SE = .80$), and Large Distractor ($M = 17.6$, $SE = .76$) not differing from each other, contrasting with the pattern of results found in the spatially intermixed conditions (see Figure 3).

Combined analyses. Mean estimates (averaged across target values) were subjected to a Spatial Arrangement (2) × Distractor (3) repeated measures ANOVA. Results revealed a main effect of Spatial Arrangement, $F(1, 21) = 9.9$, $p = .005$, revealing that mean estimates on spatially separated trials ($M = 17.7$, $SE = .71$) were generally higher than those from spatially intermixed trials ($M = 16.9$, $SE = .67$). Importantly, a significant interaction, $F(2, 42) = 14.0$, $p < .000$ reflected the fact that the distractor set size only impacted estimates in the spatially intermixed trials (see Figure 4). The main effect of Distractor approached significance, $F(2, 42) = 2.5$, $p < .1$.

Superset trials.

Spatially intermixed trials. Mean responses to all spatially intermixed superset trials were subjected to a 4 (Superset size) × 2 (Number of Subsets) repeated measures ANOVA. Analyses revealed a main effect of Superset size, $F(1.5, 31.5) = 144.6$, $p < .000$ and a main effect of Number of Subsets, $F(1, 21) = 8.0$, $p = .01$ and no interaction ($p > .2$). Again, results replicated that of Experiments 1 and 4, with estimates of supersets containing one subset ($M = 28.0$, $SE = 1.1$) lower than estimates of supersets containing two subsets ($M = 29.1$, $SE = 1.2$; Figure 5).

Spatially separated trials. An identical ANOVA on Spatially Separated Superset trials revealed an almost identical pattern of results. In particular, analyses revealed significant main effects of Superset size, $F(1.4, 30.1) = 138.5$, $p < .000$ and Number of Subsets, $F(1, 21) = 24.5$, $p < .000$ and no interaction ($p > .13$). Again, Supersets containing a single subset produced lower estimates ($M = 28.0$, $SE = 1.1$) than those containing two ($M = 29.8$, $SE = 1.2$; Figure 5).

Combined analyses. To determine whether the Superset Effect observed in our spatially separated trials was comparable or greater in magnitude to that observed in our spatially intermixed trials, a Superset Size (4) × Spatial Arrangement (2) repeated measures ANOVA was performed on estimates of supersets containing two subsets (the only trials which differed between the two conditions). Results revealed a main effect of Spatial Arrangement, $F(1, 21) = 5.7$, $p < .03$, revealing that 2-subset superset estimates in the spatially separated trials ($M = 29.8$, $SE = 1.2$) were significantly greater than those in the spatially intermixed trials ($M = 29.1$, $SE = 1.2$). This magnified Superset Effect observed for the spatially separated trials was constant across all set sizes tested (no interaction, $p > .8$).

Discussion

Again, results of both subset and superset intermixed trials replicated those of previous experiments, providing further evidence that the Subset and Superset Effects are robust. When presented with two intermixed subsets, observers estimated fewer members of one subset as the size of the other, irrelevant subset increased (Subset Effect). And, when estimating the total number of items present, observers provided higher estimates when the superset could be divided into two distinct subsets (based upon color) than when it was composed of only a single subset (all one color; Superset Effect). In sum, results of the intermixed trials corroborate findings from previous experiments to suggest that irrelevant set information, such as the size of other subsets and/or subset composition, make a significant impact on numerical estimates.

Results of the spatially separated trials, however, yielded a disparate pattern of results. Whereas the Subset Effect disappeared on spatially separated trials, in contrast, the Superset Effect was magnified. That spatial separation led to a greater discrepancy between superset estimates for sets containing a single as opposed to two subsets (compared to intermixed trials) is not surprising; separation highlighted subset membership (although irrelevant to the task), making it harder for participants to ignore subset membership on these trials. Given that observers were already unable to ignore subset membership when making superset estimates when sets were intermixed, the heightened awareness of subset membership provided by spatial separation could only have made it more difficult to ignore the subset composition, thereby exacerbating the Superset Effect.

On the other hand, this heightened subset awareness from spatial separation made the task of identifying members of individual subsets easier, leading to reduced visual search demands on subset trials. Although evidence from intermixed trials in this experiment (as well as those in Experiments 1–3) suggests that set estimates are partially dependent upon the size of other irrelevant sets, data from spatially separated trials reveal that this is only the case when the sets are intermixed, forcing the observer to search through both sets in order to individuate members of each set prior to the enumeration process. That is, when visual search demands are eliminated or reduced, numerical contrast effects disappear. Returning to a real-world example, results of Experiment 5 suggest that the number of girls in a coed classroom only affects estimates of the number of boys in the class so long as the girls and boys are randomly dispersed throughout the classroom. As soon as the children are separated by gender, then estimates should be truly independent (of course, this is assuming gender is the only salient subset grouping).

General Discussion

Results from five experiments reveal that irrelevant subsets have a robust impact on numerical estimation. When estimating the number of items contained within a subset, the existence and size of other, irrelevant sets mattered such that observers perceived fewer items in the target set as the number of items within the distractor subset increased. Furthermore, when estimating the superset, adults perceived a greater number of items when the superset could be broken down into two distinct subsets compared to when it was comprised of only a single subset. Remarkably, both effects were obtained even when participants knew beforehand which set they should attend to and with reduced task demands, suggesting they were not under the subject's conscious control and cannot be attributed to cognitive overload of simultaneously estimating the size of more than one set. In fact, observers could not have reasonably known which set to attend to based solely upon the display, as neither the total number of items in the display, size of the smallest or largest subset, nor the ratio of the subset sizes were indicative of what the target set would be. Instead, we posit that these numerical biases are the result of a preattentive perceptual process, such that adults could not help but enumerate both subsets within a display, even when not necessary. Thus, not only can adults attend to two subsets simultaneously (Halberda et al., 2006), our findings indicate that they do so automatically and unintentionally.

Subset Enumeration

There is still the open question of why this preattentive enumeration yielded the particular pattern of results observed. Why did larger distractor sets result in smaller target subset estimates? Results of Experiment 3 rule out the simple perceptual explanation of density accounting for these results. Alternatively, we suggest findings are indicative of a simultaneous contrast effect for number, similar to those found for other magnitudes such as weight, duration, and loudness (e.g., Preston, 1936; Melamed & Thurlow, 1971; Mo, 1971). Across four experiments, participants estimated fewer items in the target set as the number of distractor items increased, consistent with a simultaneous contrast effect for number. Thus, these data provide the first evidence that estimates of number subjectively contrast with our percepts of number.

It is interesting that the Subset Effect disappeared on spatially separated trials of Experiment 5 when visual search demands were eliminated. Why might this be? One possibility is that by spatially separating the sets and reducing visual search demands, observers were able to enumerate each set sequentially. In contrast, when sets are intermixed, both sets may be enumerated in parallel as observers visually search through the array. This sequential enumeration may have allowed observers to keep counts of the two arrays independent in memory, eliminating the impact of contrast effects. Future research should explore this hypothesis, possibly using eye-tracking procedures to determine whether, when presented with intermixed arrays, adults attempt to fixate on dots of all one color first followed by the other (in sequence) or if instead, as hypothesized, fixation patterns are independent of dot color.

Superset Enumeration

Findings of a Superset Effect reveal that, despite seeing the same number of dots, observers estimated the size of sets composed of two subsets as larger than those composed of a single subset. Halberda et al. (2006) reported that, on average, adults were capable of simultaneously enumerating three different sets—two subsets and the superset—however, their analyses did not explore whether the composition of the superset resulted in different estimates. Although our data align with Halberda et al.'s overall findings—that subset and superset estimates were unaffected by probe order (so long as the number of subsets was two or fewer)—our analyses suggest that the conclusion that observers were simultaneously enumerating subsets and the superset may not be accurate. In contrast, our data are consistent with the idea that observers did not, in fact, enumerate the superset, but instead enumerated each individual subset and performed an approximate summation of the two estimates. Previous investigations into the responses to approximate addition and subtraction of numerical estimates have reported infants and adults overestimate sums and underestimate differences (McCrink et al., 2007; McCrink & Wynn, 2009). Our data are consistent with this "operational momentum," such that observers likely estimated the sizes of the two individual subsets and added the estimates together, with the resulting sum larger than estimates of a single set of the same size.

It should be noted that Halberda et al. (2006) did address the possibility of a summation account but ruled it out based upon two arguments. First, they suggested that flat RTs, regardless of the number of subsets in the array, were inconsistent with a summation account that would predict higher RTs for more summations. But given that observers typed in their responses using a computer keyboard (as they did in our task), RTs on their task would not be a sensitive enough measure to reflect the minute differences predicted by mental summations. Second, they pointed to the fact that subjects in their study were able to estimate the superset even when it contained as many as six subsets, despite being unable to enumerate six sets simultaneously. While true, this finding alone does not necessarily preclude a summation account. It is plausible that when presented with more subsets than can be simultaneously enumerated, adults shift from tracking the size of individual subsets to tracking the size of the superset as a whole. As posited by Halberda et al. (2006), set membership may be processed prior to numerical tracking. When presented with a small number of subsets (e.g., two), set membership is easily identified, and thus subsets are enumerated separately. When presented with a large number of subsets (e.g., six) that are less easily parsed, on the other hand, observers may be unable to keep track of the sizes of all subsets, and thus set membership may be ignored.

Regardless, results of our study strongly suggest that despite being asked to attend to the superset, adults ignored the superset and instead simultaneously enumerated the individual subsets and summed these estimates, resulting in slightly larger estimates for supersets composed of two subsets than for those composed with only a single subset (the superset). This clearly was a subconscious process, as the Superset Effect was observed even when observers knew beforehand to attend to the superset (Experiments 1, 4, and 5) and when subset membership was not relevant in any way to the task (Experiment 4), and was magnified when subset membership was highlighted (Experiment 5). Further, this finding contradicts Halberda et al.'s conclusions, and instead suggests that, at least under these circumstances, adult observers will only enumerate two sets, not three, at a time.

Conclusion

In sum, our numerical perception abilities are far from perfect. When presented with multiple sets of items, we monitor the number of items within each of these individual subsets, but our estimates of the size of one set depend heavily upon the size of the other set. Furthermore, even when we only need to attend to the set as a whole and subset membership is irrelevant, we still deal with the set in a piecemeal-wise manner, attending to each individual subset, judging supersets composed of two subsets as larger than those with only a single subset. Combined, our findings reveal that set composition and extraneous numerical information reliably affect our perceptions of number, a finding with serious implications for our ability to accurately estimate number in real-world contexts.

References

Allik, J., & Tuulmets, T. (1991). Occupancy model of perceived numerosity. *Perception & Psychophysics, 49,* 303–314. doi:10.3758/BF03205986

Baker, J. M., Rodzon, K., & Jordan, K. (in press). The impact of emotion on numerosity estimation. *Frontiers in Psychology*.

Cantlon, J. F., Cordes, S., Libertus, M. E., & Brannon, E. M. (2009). Numerical abstraction: It ain't broke. *Behavioral and Brain Sciences, 32,* 331–332. doi:10.1017/S0140525X09990513

Cordes, S., Gallistel, C. R., Gelman, R., & Latham, P. (2007). Nonverbal arithmetic in humans: Light from noise. *Perception & Psychophysics, 69,* 1185–1203. doi:10.3758/BF03193955

Cordes, S., Williams, C. L., & Meck, W. H. (2007). Common representations of abstract quantities. *Current Directions in Psychological Science, 16,* 156–161. doi:10.1111/j.1467-8721.2007.00495.x

Durgin, F. H. (1995). Texture density adaptation and the perceived numerosity and distribution of texture. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 149–169. doi:10.1037/0096-1523.21.1.149

Durgin, F. H. (2008). Texture density adaptation and visual number revisited. *Current Biology, 18,* R855–R856. doi:10.1016/j.cub.2008.07.053

Gallistel, C. R. (1990). *The organization of learning*. Cambridge, MA: MIT Press.

Gallistel, C. R., Gelman, R., & Cordes, S. (2006). The cultural and evolutionary history of the real numbers. In S. Levinson & P. Jaisson (Eds.), *Evolution and culture: A Fyssen Foundation symposium* (pp. 247–274). Cambridge, MA: MIT Press.

Geary, D. C. (2013). Early foundations for mathematics learning and their relations to learning disabilities. *Current Directions in Psychological Science, 22,* 23–27. doi:10.1177/0963721412469398

Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature, 455,* 665–668. doi:10.1038/nature07246

Halberda, J., Sires, S. F., & Feigenson, L. (2006). Multiple spatially-overlapping sets can be enumerated in parallel. *Psychological Science, 17,* 572–576. doi:10.1111/j.1467-9280.2006.01746.x

Jordan, K. E., & Brannon, E. M. (2006). The multisensory representation of number in infancy. *PNAS: Proceedings of the National Academy of Sciences of the United States of America, 103,* 3486–3489. doi:10.1073/pnas.0508107103

Knops, A., Viarouge, A., & Dehaene, S. (2009). Dynamic representations underlying symbolic and nonsymbolic calculation: Evidence from the operational momentum effect. *Attention, Perception, & Psychophysics, 71,* 803–821. doi:10.3758/APP.71.4.803

Levinthal, B. R., & Franconeri, S. L. (2011). Common-fate grouping as feature selection. *Psychological Science, 22,* 1132–1137. doi:10.1177/0956797611418346

McCrink, K., Dehaene, S., & Dehaene-Lambertz, G. (2007). Moving along the number line: Operational momentum in nonsymbolic arithmetic. *Perception & Psychophysics, 69,* 1324–1333. doi:10.3758/BF03192949

McCrink, K., & Wynn, K. (2009). Operational momentum in large-number addition and subtraction by 9-month-olds. *Journal of Experimental Child Psychology, 103,* 400–408. doi:10.1016/j.jecp.2009.01.013

Melamed, L. E., & Thurlow, W. R. (1971). Analysis of contrast effects in loudness judgments. *Journal of Experimental Psychology, 90,* 268–274. doi:10.1037/h0031548

Mo, S. S. (1971). Judgment of temporal duration as a function of numerosity. *Psychonomic Science, 24,* 71–72.

Penney, T. B., Allan, L. G., Meck, W. H., & Gibbon, J. (1998). Memory mixing in duration bisection. In D. A. Rosenbaum & C. E. Collyer (Eds.), *Timing of behavior: Neural, psychological, and computational perspectives* (pp. 165–193). Cambridge, MA: MIT Press.

Posid, T., Huguenel, B., & Cordes, S. (2013). *Stimulus heterogeneity facilitates difficult number judgments in preschoolers*. Manuscript in preparation.

Preston, M. G. (1936). Contrast effects and the psychometric judgments. *The American Journal of Psychology, 48,* 389–402. doi:10.2307/1415884

Redden, J. P., & Hoch, S. J. (2009). The presence of variety reduces perceived quantity. *Journal of Consumer Research, 36,* 406–417. doi:10.1086/598971

Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development, 75,* 428–444. doi:10.1111/j.1467-8624.2004.00684.x

Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review, 101,* 80–102. doi:10.1037/0033-295X.101.1.80

Young, L. N., & Cordes, S. (2013a). Fewer things, lasting longer: The effects of emotional stimuli on quantity judgments. *Psychological Science, 24,* 1057–1059. doi:10.1177/0956797612465294

Young, L. N., & Cordes, S. (2013b). *Quantifying a threat: Numeric processing bias in panic-prone individuals*. Manuscript submitted for publication.

Received August 23, 2012
Revision received June 11, 2013
Accepted July 18, 2013 ■