Summary Statistics of Size: Fixed Processing Capacity for Multiple Ensembles but Unlimited Processing Capacity for Single Ensembles

Mouna Attarha, Cathleen M. Moore, and Shaun P. Vecera
University of Iowa

We assessed the processing capacity of establishing statistical summary representations (SSRs) of mean size in visual displays using the simultaneous-sequential method. Four clusters of stimuli, each composed of several circles with various diameters, were presented around fixation. Observers searched for the cluster with the largest or smallest mean size. In the simultaneous condition, all four clusters were presented concurrently; in the sequential condition, the clusters appeared two at a time. We found that the processing capacity of SSRs for multiple ensembles was as extreme as a fixed-rate bottleneck process (Experiment 1). A control experiment confirmed that this was not caused by having to compare the results of multiple averaging processes (Experiment 2). In contrast to computing SSRs across ensembles, computing SSRs for a single ensemble using the same stimuli was consistent with unlimited-capacity processing (Experiment 3). Contrary to existing claims, summary representations appear to be extracted independently for items within single ensembles but not multiple ensembles. A developing understanding of capacity limitations in perceptual processing is discussed.

Keywords: summary statistics, ensemble representations, processing capacity, simultaneous-sequential method

Imagine walking into a lecture hall late. After a quick glance through the crowded room, you search for an available seat in areas with the fewest number of people. How is a representation of average density within each portion of the room achieved? One possibility is that visual information is summarized by processes that analyze the scene in a manner analogous to how descriptive statistics summarize a dataset (e.g., mean density).

A coarse representation of statistical properties, such as the mean, variance, and range of stimulus characteristics (statistical summary representations; SSRs) has been proposed to guide behavior more efficiently than would processing the features and objects of a scene individually (e.g., Ariely, 2001; Chong & Treisman, 2003, 2005). For example, forming a surface-area ratio of brown-to-yellow for bananas may facilitate choosing the bunch with the fewest bruises more quickly than inspecting each banana in isolation, forming a summary of mean berry size across multiple bushes may lead to choosing the most fruitful bush, and forming a representation of the average emotion of students in a classroom may inform an instructor if he should speak more slowly. The proposed function of SSRs in this guidance process is to decrease processing load by summarizing redundant information, especially for unattended regions of the visual field and for areas of the periphery (e.g., Alvarez & Oliva, 2008; Alvarez, 2011; Joo, Shin, Chong, & Blake 2009). Under this view, the detailed impression of the world that we experience depends on the integration of details that are sampled at fixation with a low-resolution representation of the general properties of groups of stimuli across the visual field. Understanding how SSRs are computed then is important for understanding visual perception more generally (Chong & Treisman, 2003).

Early work on this question demonstrated that SSRs can be computed with surprising accuracy. In one study, observers saw displays that included a set of different-sized circles, and were asked to compare the perceived mean size of the circles to the diameter of a subsequently presented probe circle (Ariely, 2001). The set never included a circle that matched the mean size of the set exactly, yet observers could report whether the size of the probe was smaller or larger than the mean size of the group for diameter differences as small as 4–6%. Critically, when observers were asked to report which of the two probe circles had been a member of the set, performance was at chance levels. These results indicate that observers were able to compute the mean size of a set of stimuli quite accurately, even when they failed to either identify or remember the sizes of individual stimuli from the set. Poor performance for member identity is consistent with the view that there exists a tradeoff between SSRs and the representation of individual stimuli; precise information is lost in favor of a more holistic, low-resolution summary of the environment (Alvarez, 2011). This might occur in a variety of different specific ways. The identity of individual members might be discarded after a global percept of the group is formed (Ariely, 2001). Alternatively, individual representations might be so noisy that they cannot be used reliably for...
later identification (Alvarez & Oliva, 2008). Regardless, the results suggest that averaging is distinct from individual object processing (e.g., Alvarez, 2011; Corbett & Oriet, 2011; Im & Halberda, 2013; Jacoby, Kamke, & Mattingley, 2013; Joo et al., 2009).

The ability to establish SSRs is not restricted to calculating the mean size of stimuli (e.g., Ariely, 2001; Brady & Alvarez, 2011; Chong & Treisman, 2003, 2005; Corbett & Oriet, 2011). The visual system can accurately average over many different feature dimensions such as orientation (e.g., Dakin & Watt, 1997; Robitaille & Harris, 2011), length (e.g., Weiss & Anderson, 1969), brightness (e.g., Bauer, 2009), spatial position (e.g., Alvarez & Oliva, 2008), speed (e.g., Emmanouil & Treisman, 2008; Watamaniuk & Dachon, 1992), and direction of motion (e.g., Watamaniuk, Sekuler, & Williams, 1989; Dakin & Watt, 1997). SSRs also act on higher-order information such as the average emotion, gender, and identity of a group of faces in a crowd (de Fockert & Wolfenstein, 2009; Haberman & Whitney, 2007, 2009; Haberman, Hurt, & Whitney, 2009). Further, averaging is not limited to the spatial dimension. It can operate over stimuli that change size over time (Albrecht & Scholl, 2010), and across sensory modalities (Albrecht, Scholl, & Chun, 2012). Although less studied, other statistics, such as the variance, proportion, range, skew, and kurtosis of stimulus values, can also be estimated surprisingly well (Morgan, Chubb, & Solomon, 2008; Peterson & Beach, 1967; Pollard, 1984). The breadth and scope of stimuli over which statistical summaries occur is consistent with the possibility that SSRs reflect a general mechanism that pools redundant information over all available stimulus dimensions (Alvarez & Oliva, 2008; Alvarez, 2011).

In addition to establishing that observers can compute SSRs, a focus in the literature has been that SSRs are established quickly and in parallel across the visual field. Chong and Treisman (2003), for example, showed that mean-size estimates could be extracted in as little as 50 ms and were unaffected by further increases in exposure duration. The authors argued that given this short timeframe, it is unlikely that the mean was calculated through a serial process akin to adding the sizes of individual items and dividing by the sum of the number of circles present. Consistent with these claims, Ariely (2001) found that discrimination thresholds of mean size were unaffected by the number of items (4 or 16) in the to-be-averaged set, and Chong and Treisman found that mean-size estimates for a group of heterogeneously sized circles were as accurate as those for single circles. Finally, in a later study, Chong and Treisman (2005) showed that observers were able to report the mean of one of two interpersed sets of circles that were defined by color, and that performance was unaffected by whether the color of the to-be-reported subset was preceded or postcedue relative to the display. Observers in this study were also able to report the mean size of one of two colored subsets of stimuli as well as they could for individually presented sets, but Brand, Oriet, and Tottenham (2012) were unable to replicate this aspect of the results. Together these findings have lead researchers to conclude that SSRs for multiple sets are established through processes that “...precede the limited capacity bottleneck” (Chong & Treisman, 2005), and by implication are established through unlimited-capacity processes (e.g., Oriet & Brand, 2013; Robitaille & Harris, 2011).

The goal of the current study was to test the hypothesis that SSRs are established through unlimited-capacity processes. Unlimited-capacity models state that processing occurs independently (i.e., without interference) across stimuli. These models therefore predict that performance will not vary with the number of stimuli that must be processed simultaneously. In contrast, limited-capacity models state that the processing of one stimulus is compromised by having to process other stimuli simultaneously. These models therefore predict that performance will decline with increasing numbers of simultaneous stimuli.

Although the absence of set size effects in the studies reviewed above is consistent with an unlimited-capacity model of SSRs (e.g., Chong & Treisman, 2003), the evidence is equivocal with regard to the issue of interference because of the way in which set size was manipulated. Specifically, in order to maintain given average sizes, Ariely (2001) varied set size between 4 and 16 items by varying the frequency of only four distinct circle sizes. Observers therefore did not have to sample all of the stimuli in a set to do the task. They could instead sample from only a portion (e.g., an average of 4 items), effectively nullifying the set-size manipulation. When size regularity across items was minimized, forcing observers to sample from the whole set, significant set size effects were observed (Marchant, Simons, & de Fockert, 2013; Myczek & Simons, 2008; but cf. Ariely, 2008; Robitaille & Harris, 2011).

Set size manipulations, in general, are not ideal for testing between unlimited- and limited-capacity processing models because varying set size affects aspects of the task other than the number of to-be-processed stimuli (e.g., Palmer, 1994; Shaw, 1984; Townsend, 1990). Specifically, set size confounds the number stimulus that must be processed with the number of perceptual representations that contribute to the task decision. Because every representation is associated with a certain amount of noise, a greater number of representations implies a greater amount of noise that is fed into decision processes. Poorer performance with larger set sizes could reflect this difference alone. One strategy for handling this confound has been to develop specific models of the task in question and use them to make quantitative predictions regarding how large an effect the increased noise should have on performance. These predictions can then be compared to the observed effect of set size on performance, which will either be more or less the same as that predicted by increased decision noise alone or not (e.g., Palmer, 1994; Shaw, 1984). Though effective within specific contexts, this strategy is limited in that it is dependent on the development of specific processing models that require, and these models require specific, and often ancillary to the question of interest, assumptions about how processing unloads.

The simultaneous-sequential paradigm is an alternative approach to avoiding sensory confounds associated with set size manipulations (e.g., Eriksen & Spencer, 1969; Scharff, Palmer, & Moore, 2011a; Scharff, Palmer, & Moore, 2011b; Scharff, Palmer, & Moore, 2013; Shiffrin & Gardener, 1972). This paradigm holds constant the number of stimuli that contribute to the response, and therefore the amount of noise at decision, across conditions. Instead, what varies is how many stimuli must be processed at any given time. Specifically, in the simultaneous condition, all stimuli appear at the same time, and therefore must be processed at the same time. In contrast, in the sequential condition, the stimuli appear in smaller subsets at different times, and therefore fewer need be processed at any given time. Critically, the total amount of exposure time to any given stimulus is the same in the two
conditions; the only variable that differs is how many stimuli must be processed at a given time. Unlimited-capacity models predict no difference between the simultaneous and sequential conditions. This follows because if processing unfolds independently across multiple stimuli, then barring physical or sensory interference, it should make no difference how many stimuli require processing. In contrast, limited-capacity processing predicts an advantage for sequential over simultaneous presentation because the sequential condition allows fewer stimuli to engage the process at any given time.

In the current study, observers viewed four clusters of circles and reported whether the mean size of one of the clusters was larger or smaller than the others (Figure 1). The four clusters were presented all at once in the simultaneous condition (Figure 1A) or in subsets of two in the sequential condition (Figure 1B). If SSRs unfold independently across groups of stimuli (i.e., they are established through unlimited-capacity processes), then performance should be just as good in the simultaneous condition as in the sequential condition. In contrast, if computing SSRs interfere with each other across clusters (i.e., they require limited-capacity processes), then performance should be better in the sequential condition than in the simultaneous condition because fewer items require processing at any given time.

Finally following Scharff et al. (2011a), we included a repeated condition (Figure 1C). This was just like the simultaneous condition, except that the display was presented twice. This provided two advantages over the basic design. First, it allowed us to confirm that if SSRs do involve limited-capacity processes, our conditions were such that observers could have taken advantage of the sequential condition. If, for example, the stimulus duration that we used was too long, SSRs for all clusters could be established by shifting processing within that one display period, then performance might be equal across the simultaneous and sequential conditions, despite SSRs depending on limited-capacity processes. If performance is equal across the simultaneous and sequential conditions but better in the repeated condition, however, then we can be assured that this was not the case. Another advantage of including the repeated condition is that in the event that processing is limited capacity, it allows one to assess a particular limited-capacity model—fixed capacity—which states that processing is limited to a fixed amount of information per unit time. A serial model (i.e., one cluster at a time) is a specific example of a fixed-capacity model. A fixed-capacity model predicts not only that performance will be higher in the sequential condition than in the simultaneous condition, but also that it will be as good as performance in the repeated condition. Formal details of these predictions are given in Scharff et al., 2011a.

As a preview of our results, we found that computing SSRs for multiple ensembles of stimuli was inconsistent with unlimited-capacity processing, and consistent with fixed-capacity processing (Experiment 1). A control experiment confirmed that this was not caused by having to compare the results of multiple averaging processes (Experiment 2). In contrast to computing SSRs across ensembles, computing SSRs for a single ensemble was consistent with unlimited-capacity processing (Experiment 3). The striking contrast in results for computing SSRs across multiple ensembles (fixed capacity) versus computing an SSR for a single ensemble (unlimited capacity) provides an explanation for apparently conflicting results and conclusions regarding the processing limitations of SSRs within the literature.

## Experiment 1

### Method

**Observers.** An N² power analysis, which calculates the number of subjects necessary to have at least 80% power for every factor (Cohen, 1988), determined the minimum number of observers needed in our experiments. Effect size estimates for this analysis were based on a pilot run of the experiment with three subjects. This analysis indicated that at least 7 subjects were necessary to detect effects in this design if they were there. For good measure, we increased this number by five prior to running any study. Twelve undergraduate volunteers from the University of Iowa participated in exchange for course credit (six men, six women, age range: 18–26 years, all right-handed). All observers were naïve as to the purpose of the study and all reported normal visual acuity and color vision.

**Equipment.** Stimuli were displayed on a flat-screen CRT monitor (19-inch ViewSonic G900B) controlled by a Macintosh Pro (Mac OS X) with a 512-MB NVIDIA GeForce 8800 GT graphics card (1024 by 768 pixels, viewing distance of 61.5 cm, horizontal refresh rate of 100 Hz). Stimuli were generated using the Psychophysics Toolbox Version 3.0.8 (Brainard, 1997; Pelli, 1997) for MATLAB (Version 7.5, Mathworks, MA). Observers sat in a height-adjustable chair and used an adjustable chin rest to maintain a constant viewing distance from the monitor.

**Stimuli.** Displays consisted of 16 filled circles of various sizes (Figure 1), which were presented as luminance increments (43.03 cd/m²) on an achromatic background (39.45 cd/m²). The circles...
were configured to give rise to the perception of four clusters centered 4.19° from fixation. Each cluster was composed of four circles whose sizes were chosen from a target or distractor distribution. The center of the circle closest to fixation was 3.26° away, while the circle furthest from fixation was 5.59° away. Clusters were separated horizontally and vertically by 6.05° center-to-center.

On every trial, the sizes of circles within three of the four clusters were randomly chosen from a uniform distractor distribution (Range: 1.09°–1.96°), while the sizes of circles within the fourth cluster were equally chosen from either a uniform small-target distribution (Range: 0.34°–1.38°) or a uniform large-target distribution (Range: 1.40°–2.21°). Each distribution contained 122 possible diameter sizes. The sizes were equally spaced on a power function with an exponent of 0.76, identified by Teghtsoonian (1965) as the psychological scale for size (see also Chong & Treisman, 2003). The lower bound of the distractor distribution was the median of the small-target distribution while the upper bound of the distractor distribution was the median of the large-target distribution. The heavy overlap between target and distractor distributions minimized the degree to which observers could bypass the averaging process by using size information of individual circles to perform the task. While this potential strategy was not eliminated in the current experiment (i.e., the distributions did not fully overlap), it is only a concern if evidence of unlimited capacity is obtained. Stated another way, simply using feature information to determine target identity predicts equal performance between the simultaneous and sequential conditions because extreme sizes exclusive to the target distributions would “pop-out” and would be processed with parallel, unlimited capacity (Huang & Pashler, 2005).

Procedure. Observers completed one 45-min session that consisted of a practice block of 30 trials, followed by six experimental blocks of 48 trials each (96 observations per condition, 288 experimental observations per subject). Practice trials were excluded from all analyses.

Trials began with a centrally located black fixation cross (0.25° × 0.25°) for 500 ms. In the simultaneous condition, this was followed by the four clusters for 50 ms, and then a blank screen until response (Figure 1A). In the sequential condition, fixation was followed by two clusters for 50 ms presented along either the positive or negative diagonal, a blank ISI of 1,100 ms, the other two clusters for 50 ms presented along the opposite diagonal, and a blank screen until response (Figure 1B). The repeated condition was the same as the sequential condition except that all four clusters appeared in both of the two 50 ms displays (Figure 1C). Written feedback was given for 1,000 ms in the form of words “correct” or “incorrect” at fixation following each response. The next trial automatically began 1,000 ms after the presentation of feedback.

Display type (simultaneous, sequential, repeated), target type (small, large), and target position (upper-left, upper-right, lower-left, lower-right) were randomly mixed within blocks of trials and appeared equally often. Which of the two diagonally opposite positions were presented first in the sequential display was constant for a given observer but varied across observers. The purpose of this was to eliminate uncertainty of presentation positions.

Task. The task was to find the cluster of circles that had a different mean size than the other three, and to report whether it was smaller or larger than the others by pressing the “F” or “J” key, respectively. Observers were instructed to maintain central fixation and respond as accurately as possible.

Method of analysis. All theoretical models assume a reliable advantage in the repeated condition relative to the simultaneous condition. Subjects who did not meet this criterion were omitted from all analyses and replaced until a total of 12 subjects in each experiment were collected. One, one, and seven subjects failed to show a repeated advantage in Experiments 1, 2, and 3. In the experiments that follow, accuracy data were transformed to arcsin values to normalize their distributions. The underlying assumptions of all statistical tests were confirmed and corrections were made if needed. Violations of normality and sphericity were confirmed using a one-sample Kolmogorov–Smirnov test and Mauchly’s test. Violations of sphericity were corrected using the Greenhouse-Geisser epsilon. Follow-up t tests were used after significance of the final model was verified.

Results and Discussion

Figure 2 shows the mean percent correct as a function of condition, collapsed across observers. Performance was higher in the sequential condition than the simultaneous condition, which is inconsistent with an unlimited-capacity model of SSRs but consistent with a limited-capacity model. Moreover, because performance was as high in the sequential condition as in the repeated condition, the results are consistent with the fixed capacity version of the limited-capacity model. Inferential statistics confirmed these descriptive patterns.
Because the sizes of circles were chosen randomly from partially overlapping distributions, a small percentage of trials would by chance include a distractor cluster whose mean size was either greater than (or less than) the mean size of the large or small target cluster, respectively. As a result, the cluster that appeared to be the target might in fact be a distractor cluster. We omitted these trials from the reported analyses. Out of 3,456 experimental trials across all observers, a total of 248 (7%) were omitted for this reason. Elimination of these trials did not alter the results qualitatively.

Arcsin transformed values were submitted to a one-way repeated-measures ANOVA with condition as the within-subjects factor (all Kolmogorov–Smirnov ps > .018; Mauchly’s p = .027, Greenhouse–Geisser ε = .661). The final model was significant, F(1,32,14.54) = 16.91, p < .001, pη² = .606, MSE = .010. As predicted by fixed-capacity processing, accuracy was not reliably greater in the sequential condition (83.4%) than in the repeated condition (84.5%), t(11) = 1.06, p = .313. However, performance in the sequential condition was significantly greater than in the simultaneous condition (69.9%), t(11) = 4.22, p = .001.

An assumption of the simultaneous-sequential method is that the conditions differ only with respect to how many stimuli must be processed simultaneously. They did necessarily differ, however, in when the target appeared within the trial sequence. In the simultaneous condition the target always appeared in the “first” frame because that was the only frame, whereas in the sequential condition, the target could appear in either the first frame or in the second frame. This difference might provide an advantage to the simultaneous condition if there are any memory differences across the two conditions. To assess this possibility, we compared performance in the sequential condition for trials on which the target appeared in the first and second frames. No reliable differences were observed: 82.3% (first frame) versus 83.7% (second frame), F(1, 11) = 0.25, p = .625, pη² = .023, MSE = .005 (all Kolmogorov–Smirnov, ps > .027).

The results of this experiment indicate that establishing SSRs of size engage limited-capacity processes, and, in particular, that only a fixed amount of information can be processed per unit time. Furthermore, the reliable difference between the simultaneous and sequential conditions shows this experiment had the power to detect an unlimited-capacity result. In summary, the results of Experiment 1 are consistent with a limited-capacity model of SSRs for multiple ensembles, and not with an unlimited-capacity model.

Experiment 2

We have interpreted the results of Experiment 1 as evidence that SSRs of mean size involve limited-capacity processes. Successful performance, however, required that observers not only compute the mean size of each cluster, but also compare those means and determine whether the mean furthest away, in numerical terms, was relatively smaller or larger. To rule out the possibility that it was some other aspect of the task that caused performance to be limited capacity, we conducted a control experiment in which the task required all of the same processes except computing mean size. Subjects were shown clusters of homogeneous circles, the size of each determined by the mean of their respective cluster from Experiment 1. Averaging was no longer required since the mean of each cluster was directly provided and since all circles within a cluster were of equal size (Figure 3). The task was the same otherwise. If the limited capacity results of Experiment 1 were caused by limited capacity SSR formation and nothing else, then we should find evidence of unlimited-capacity processing in this second experiment.

Method

All aspects of the stimuli and procedure were identical to Experiment 1, with the following exceptions.

Observers. Twelve undergraduate volunteers from the University of Iowa participated in exchange for course credit (four men, eight women, age range: 17–20 years, 11 right-handed).

Stimuli. As in Experiment 1, the sizes of 16 circles were randomly chosen from the appropriate target or distractor distribution. But before the stimuli were presented, the mean size for each cluster was computed. The size of all circles within a given cluster was adjusted according to that cluster’s mean prior to presentation (Figure 3). As a result, subjects could circumvent the averaging process by directly comparing individual circles to determine whether the oddball cluster was relatively larger or smaller than the others.

Results and Discussion

Figure 4 shows the mean percent correct as a function of condition collapsed across observers. Performance was no different in the sequential condition than in the simultaneous condition, but it was higher in the repeated condition than the other two. This pattern is consistent with an unlimited-capacity model and inconsistent with a limited-capacity model. The inferential statistics confirmed this pattern.
As in Experiment 1, we filtered trials in which the perceptually correct response may have led to an “incorrect” feedback message (241 trials of 3,456 total trials across observers, for 7%). Arcsin transformed values were submitted to a one-way repeated-measures ANOVA with condition as the within-subjects factor (all Kolmogorov–Smirnov ps > .895; Mauchly’s $p = .003$, Greenhouse–Geisser $\varepsilon = .591$). The final model was significant, $F(1,18,13.01) = 10.89, p = .004$, $R^2 = .498$, $MSE = .004$. As predicted by unlimited-capacity processing, accuracy was not reliably greater in the sequential condition (77.8%) than in the simultaneous condition (77.7%), $t(11) = 0.37, p = .722$. However, performance in the sequential condition was significantly lower than performance in the repeated condition (84.3%), $t(11) = 2.88, p = .015$.

We again compared performance within sequential trials when the target was presented in the first frame versus the second frame. Again, we found that performance across both frames were statistically equal, 78.5% (first frame) versus 76.4% (second frame), $F(1, 11) = 0.42, p = .532$, $R^2 = .036$, $MSE = .010$ (all Kolmogorov–Smirnov ps > .877), suggesting that targets presented first did not suffer from more memory loss than targets presented closer in time to response.

In summary, the results of Experiment 2 indicate that when the task no longer requires the computation of averages, processing becomes unlimited capacity. The fact that the results of Experiment 1 indicated fixed capacity can be confidently interpreted as evidence that SSRs depend on limited averaging processes, and not on limited comparison or decision processes. The crowding of items within each cluster also cannot explain the reported limitation since the stimulus spacing in Experiment 1 was preserved in Experiment 2 (Banno & Saiki, 2012; Bouma, 1970).

### Experiment 3

Experiment 1 showed that computing SSRs of size for multiple ensembles engages fixed-capacity processes. But what about computing an SSR for a single ensemble of stimuli? Previous studies asking about SSRs have rarely made this distinction. Some have used tasks in which SSRs are computed across a single ensemble (e.g., Ariely, 2001; Robitaille & Harris, 2011), whereas others have used tasks in which SSRs are computed across multiple ensembles (e.g., Banno & Saiki, 2012; Oriet & Brand, 2013). It is possible that computing SSRs is limited by the number of ensembles for which an SSR is extracted, but that computing a single SSR is not limited by the number of stimuli across which the summation is made. If that were the case, then not distinguishing between tasks that depend on SSRs of single versus multiple ensembles could lead to apparently conflicting conclusions about whether computing SSRs is limited capacity. Indeed such conflicting conclusions exist. For example, the fixed-capacity conclusion drawn from Experiment 1 of this study, which involved multiple ensembles, contrasts with the unlimited-capacity conclusion drawn from a study reported by Robitaille and Harris (2011), which focused on single ensembles. Experiment 3 used the simultaneous-sequential method to test the capacity limitations for a task that depended on only a single-ensemble SSR.

### Method

All aspects of the stimuli and procedure were identical to Experiment 1, with the following exceptions.

**Observers.** Twelve undergraduate volunteers from the University of Iowa participated in exchange for course credit (eight men, four women, age range: 18–30 years, all right-handed).

**Stimuli.** Sixteen filled circles of various sizes were placed on a square grid spaced horizontally and vertically by 2.21° and centered at fixation (Figure 5).

**Procedure.** On each practice trial, a black probe circle (1.39°) appeared on the response screen after the simultaneous, sequential, and repeated displays. The size of the probe circle was fixed and subjects were instructed to report whether the average of all 16 circles was smaller or larger than the size of the probe circle appearing afterward. After the practice block, subjects were told that while the probe circle would not appear on experimental trials, their task remained the same because the probe remained a fixed size. This kept the trial events consistent across all three experiments. The size of the probe circle was a unique value in the distractor distribution; it did not match any of the sizes falling in either the small or large target distributions. On trials in which the target was “small,” the average of the entire cluster was shifted lower than the size of the probe circle; conversely, on trials in which the target was “large,” the average of the entire cluster was shifted higher than the size of the probe circle.

**Task.** The task was to report whether the average of the single set was smaller (“F” key) or larger (“J” key) than the probe circle that had been presented throughout the practice block.

### Results and Discussion

Figure 6 shows the mean percent correct as a function of condition collapsed across observers. Performance was no differ-
in Experiment 1, the results were consistent with the opposite processes. When the same 16 items were grouped into four clusters mean size for single ensembles engage only unlimited-capacity processing extreme. Computing summary representations for multiple ensembles introduces interference unlike single ensembles.

**General Discussion**

Applying the simultaneous-sequential method to test the capacity limitations of SSRs, we found evidence that was consistent with a limited-capacity model for the formation of SSRs for multiple ensembles and an unlimited-capacity model for the formation of an SSR for a single ensemble. Specifically, observers were poorer at responding on the basis of mean size of four clusters of circles when the clusters were all presented simultaneously compared to when they were presented two at a time sequentially. In fact, because performance was equally good in the sequential condition as in the repeated condition, the results suggest that computing mean size involve fixed-capacity processing, an extreme version of the limited-capacity model (Scharff et al., 2011a). When the same items were presented as a single perceptual unit, mean size was computed through unlimited-capacity processes and performance was equally good across the simultaneous and sequential conditions. One large set can be averaged more efficiently than multiple smaller sets.

Before discussing these findings more broadly, we note that Chong and Treisman (2003, Experiment 1) did compare performance in a task that depended on SSRs across multiple ensembles under simultaneous versus sequential presentation conditions. They found equal performance across these two conditions, which seems contrary to the results and conclusion of Experiment 1 of the present study. In their experiment, however, the simultaneous display was presented for 200 ms, whereas each frame of the

![Figure 5](image5.png)

**Figure 5.** Trial events for the (A) simultaneous, (B) sequential, and (C) repeated conditions in Experiment 3. The four clusters presented in Experiment 1 were presented on an equally spaced grid to produce the perception of a single cluster with 16 items. During practice trials (not pictured), a probe circle appeared on the response screen and subjects reported whether the mean size of the single cluster was larger or smaller than the size of the probe circle. In the real experiment (pictured), presentation of the probe circle was removed because it remained the same size on every trial. In this example, the correct response is “smaller.”

![Figure 6](image6.png)

**Figure 6.** Mean correct responses (%) as a function of display collapsed across observers in Experiment 3. Evidence consistent with unlimited capacity was obtained when summary statistics were computed for a single cluster. Error bars are within-subject 95% confidence intervals (Cousineau, 2005; Morey, 2008).
sequential display was only 100 ms each. Therefore, their simultaneous condition was similar to the repeated condition of Experiment 1 (i.e., twice the duration of the other condition), and indeed performance in this double-duration condition achieved that of the sequential condition. We suggest that rather than conflicting with our results, the results from the Chong and Treisman experiment are, like ours, consistent with a fixed-capacity model of SSRs across multiple ensembles.

The conclusion drawn from the current study clearly challenges the suggestion that the formation of SSRs, in general, bypasses limited capacity aspects of our perceptual and cognitive systems (e.g., Alvarez, 2011; Chong & Treisman, 2005). Other more recent evidence, however, has also brought into question the automaticity of perceptual averaging processes in general (Brand et al., 2012; Jacoby et al., 2013; Marchant et al., 2013; Whiting & Orient, 2011). For example, SSRs are susceptible to object-substitution masking which operates at stages of visual processing beyond the initial registration of features, suggesting that SSRs rely, at least in part, on later processing stages (Jacoby et al., 2013). The finding that capacity limitations for SSR formation depends on whether SSRs are being formed across multiple ensembles (fixed capacity) or single ensembles (unlimited capacity) could be a critical piece to the puzzle of resolving these and other apparent conflicts.

Beyond the specific question of whether SSRs require limited-capacity processes, the current application of the extended simultaneous-sequential method contributes to a developing picture regarding capacity limitations in perceptual processing more generally. Processes found to engage only unlimited-capacity processes using this method include, but are not limited to, contrast discrimination (Scharff et al., 2011a), image shape (Scharff et al., 2013), size discrimination of individual items (Huang & Pashler, 2005), modal and amodal surface completion (Attarha & Moore, 2010; Attarha, Moore, Scharff, & Palmer, 2013), symmetry detection (Huang, Pashler, & Junge, 2004), and letter identification (Shiffrin & Gardner, 1972). Processes that yield results consistent fixed-capacity include object categorization (Scharff et al., 2011b), object shape identification (Scharff et al., 2013), word categorization (Scharff et al., 2011a), and now summary statistics. These processes constitute extreme conditions, with unlimited-capacity processing on the one hand and maximally limited-capacity (i.e., fixed-capacity) processing on the other. Together the results indicate that at some point (or points) within the stream of visual processing between contrast discrimination and object identification, severe limitations ensue. When drawing similarities between the processes at each extreme, it appears as though sensory and segmentation processes have unlimited capacity while object and semantic processes have fixed capacity. The present study suggests that the formation of multiple summary statistic representations is more like object and semantic processing than it is like sensory or organizational processing.

Finally, we end with a discussion of the contrast between processing capacity (the degree to which a process can be engaged independently by multiple stimuli; Broadbent, 1958; Estes & Taylor, 1964; Rumelhart, 1970; Shiffrin & Gardner, 1972) and storage capacity (the amount of information that can be maintained in memory; Alvarez & Cavanagh, 2004; Awh, Barton, & Vogel, 2007; Luck & Vogel, 1997). Many recent studies have investigated the storage capacity of visual working memory. In an initial paper, Luck and Vogel (1997) used a simple change-detection method to estimate that observers were able to hold approximately three stimuli in visual working memory. This study lead to a flurry of follow-up studies asking questions about the nature of this capacity limitation, such as whether it is limited by the number of objects that can be held or the degree of precision with which stimuli can be remembered, or both. Because estimates of storage capacity from these studies tend to be on the order of 2.5–4 items (see Brady, Konkle, & Alvarez, 2011 for a review), there has been a tendency to criticize the use of the simultaneous-sequential method with conditions that vary from two-at-a-time presentations (sequential) to four-at-a-time presentations (simultaneous) because both 2 and 4 fall within the range of most people’s ‘capacity.’ It is critical to note, however, that the simultaneous-sequential method is assessing processing independence versus dependence, not storage capacity. If stimulus presentation conditions are such that performance is limited by how much information can be extracted from the display (e.g., because stimuli are presented briefly), then limited-capacity processing predicts a difference between simultaneous versus sequential even for one versus two items. Two versus four has been used in order to minimize contamination from differences in eye movements across conditions and to minimize contamination from sensory effects like crowding, but the logic is identical. Finally, if the criticism regarding four items is too small were valid, evidence of limited-capacity should never attain. Yet it has for many different tasks, including shape identification, spatial configuration, object categorization, and word categorization (Huang & Pashler, 2005; Scharff et al., 2011a, 2011b, 2013). Thus, while Experiment 3 clearly demonstrates unlimited capacity even for 16 items, it is important to note that the logic of the simultaneous-sequential method does not depend on this extension.

In summary, the current results indicate that the formation of SSRs across multiple ensembles of stimuli depends on limited-capacity processes (i.e., ensembles are not processed independently), whereas the formation of a single SSR for one ensemble of stimuli seems to be unlimited capacity (i.e., stimuli within an ensemble are processed independently). It has been proposed that a compressed representation of the environment that is established through the formation of multiple SSRs bypasses limited-capacity components of our perceptual and cognitive systems and serves to guide later visual processes. We suggest that this cannot be the case given the highly limited nature of forming multiple SSRs.

References


Received September 19, 2013
Revision received January 15, 2014
Accepted February 5, 2014