

# Where do we look when we count? The role of eye movements in enumeration

XINGSHAN LI

*Institute of Psychology, Chinese Academy of Sciences, Beijing, China  
and Vanderbilt University, Nashville, Tennessee*

AND

GORDON D. LOGAN AND N. JANE ZBRODOFF

*Vanderbilt University, Nashville, Tennessee*

Two experiments addressed the coupling between eye movements and the cognitive processes underlying enumeration. Experiment 1 compared eye movements in a counting task with those in a “look” task, in which subjects were told to look at each dot in a pattern once and only once. Experiment 2 presented the same dot patterns to every subject twice, to measure the consistency with which dots were fixated between and within subjects. In both experiments, the number of fixations increased linearly with the number of objects to be enumerated, consistent with tight coupling between eye movements and enumeration. However, analyses of fixation locations showed that subjects tended to look at dots in dense, central regions of the display and tended not to look at dots in sparse, peripheral regions of the display, suggesting a looser coupling between eye movements and enumeration. Thus, the eyes do not mirror the enumeration process very directly.

The coupling between eye movements and cognition varies with the task at hand. In tasks like reading, the eyes are “the window to the soul,” moving with each step of the underlying process and focusing on the object of the current computation (Rayner, 1998). In tasks like memory retrieval, the coupling is looser. People tend to look at certain locations, but the sequence of eye movements is not completely predictable (Ryan, Hannula, & Cohen, 2007). The present article addresses the coupling between eye movements and enumeration, monitoring eye position while subjects count the objects in a display and report the numerosity. We ask whether the eyes mirror the underlying counting process, and we distinguish three hypotheses about the coupling between eye movements and enumeration. The indexing hypothesis assumes that coupling is tight, so that the eyes move with each step of the counting process, fixating each object as it is included in the count. The control hypothesis assumes that coupling is looser, so that eye movements contribute to the control of counting by preventing refixations of previously viewed objects but are not involved in each counting step. The perception hypothesis assumes very loose coupling. Eye movements are not coupled directly to the counting process, but instead are directed to dense regions of the display where foveal vision can increase perceptual resolution. What is at stake is the utility of eye movements as a window to the soul of the counting process. The window is more revealing the tighter the coupling.

Enumeration is an elementary mathematical skill that is a cornerstone of mathematical literacy (Geary, 2000). It is acquired at an early age and retained throughout the life span (Beckwith & Restle, 1966; Chi & Klahr, 1975; Geary, 2000; Trick, Enns, & Brodeur, 1996). It is well understood formally and psychologically. Formally, enumeration involves establishing correspondence between the objects in a display and the set of integers, beginning with 1, and declaring the numerosity of the display to equal the largest corresponding integer, following the cardinality principle. Psychologically, enumeration involves a serial counting process that establishes correspondence between visual objects and an ordered mental representation of the set of integers (the number line). The process is serial, indexing each object one at a time and stepping through the number line as each object is indexed. When every object in the display has been indexed, the current value of the number line is reported as the numerosity (Beckwith & Restle, 1966).

The counting theory of enumeration is consistent with people’s intuitions when they count, and it makes specific behavioral predictions. First, people who employ the counting procedure will count accurately in all ranges of numerosity. The counting procedure is an algorithm that is guaranteed to produce a correct result if it is applied correctly. It works accurately even with large numerosities where magnitude estimation is inaccurate (Mandler & Shebo, 1982). Second, the time required to enumerate

---

G. D. Logan, [gordon.logan@vanderbilt.edu](mailto:gordon.logan@vanderbilt.edu)

---

a display will increase monotonically with increases in the number of objects in the display. If each counting step takes the same time, on average, then the increase in response time (RT) with increases in the number of objects will be linear, with a slope that reflects the duration of the processes involved in each step. This prediction has been confirmed in many studies (Jensen, Reese, & Reese, 1950; Klahr, 1973; Logan & Zbrodoff, 2003; Mandler & Shebo, 1982; Saltzman & Garner, 1948; Trick & Pylyshyn, 1994). For displays with 5 or more objects, RT increases linearly with numerosity with a slope around 300 msec per object. However, there are two important qualifications: First, displays with numerosities between 1 and 4 or 5 seem to be enumerated by a special process, called *subitizing*, that is relatively insensitive to the number of objects in the display, as if small numerosities can be apprehended in a single glance (Kaufman, Lord, Reese, & Volkman, 1949; Logan & Zbrodoff, 2003; Mandler & Shebo, 1982). We focus on displays of 5–12 objects in order to capture the counting process. Second, if subjects enumerate the same displays many times, counting may be replaced by memory retrieval, in which the display retrieves the numerosity associated with it in the past. Memory retrieval does not depend strongly on the number of objects, so the slope of the function relating RT to numerosity becomes shallower with practice, as subjects shift from counting to memory retrieval (Lassaline & Logan, 1993; Palmeri, 1997; Wolters, van Kempen, & Wijlhuizen, 1987). We present displays only once or twice in order to capture the counting process.

The counting model provides a framework in which to articulate our investigation of the role of eye movements in enumeration. The *indexing* hypothesis assumes tight coupling between eye movements and the underlying counting process, in which eye movements implement the process of indexing objects in the display. It predicts that each object is fixated once and only once because each object is indexed once and only once, so the number of fixations should equal the number of objects in the display. There are no refixations, and every object is fixated. This is the strong version of the indexing hypothesis. A weaker version assumes that subjects first subitize the display and then count the objects that remain, fixating once or twice in the subitized region (Watson, Maylor, & Bruce, 2007) and then fixating the remaining objects once and only once. The weaker indexing hypothesis predicts that there are no refixations, and the number of fixations increases linearly with the number of objects in the display with a slope of 1.0 (i.e., one extra fixation for each object beyond the subitizing range).

The *control* hypothesis assumes a looser coupling. Subjects may not fixate each object in the display, but they exploit the control system for eye movements to keep themselves from refixating objects they have fixated before. Whereas the indexing hypothesis is concerned with fixating objects once and only once, the control hypothesis is concerned with fixating objects only once. The control hypothesis exploits properties of the eye-movement control system that produce a phenomenon called *inhibition of return*, whereby covert attention (Pos-

ner & Cohen, 1984) and overt eye movements (Beck, Peterson, & Vomela, 2006; Peterson, Kramer, Wang, Irwin, & McCarley, 2001) are less likely to return to previously visited locations (Watson et al., 2007). Thus, the control hypothesis predicts that no object will be refixated. It does not predict that every object will be fixated (unlike the indexing hypothesis). Following precedent in the visual search literature (Beck et al., 2006; Peterson, Beck, & Wong, 2008; Peterson et al., 2001), immediate refixations that are caused by corrective saccades do not contradict the control hypothesis, but nonimmediate refixations that are not caused by corrective saccades do contradict the control hypothesis. Corrective saccades can be identified by the duration of the fixation preceding the saccade, which is shorter than usual, and by the landing point of the saccade, which is closer to the target object than is the original saccade (Henson, 1978; Rayner, 1998).

This is the strong version of the control hypothesis. A weaker version allows for nonimmediate refixations (Beck et al., 2006; Henson, 1978; Peterson et al., 2008; Peterson et al., 2001). According to the weak version of the control hypothesis, subjects tend to avoid fixating at the dots they have looked at before, although they cannot do so perfectly. The weak version allows some nonimmediate refixations, but the proportion of refixations must be smaller than that the amount predicted by a no-memory model. The coupling between eye movements and cognition is even weaker in the weak version of the control hypothesis.

The *perception* hypothesis assumes the loosest coupling. It assumes that the eyes are driven by the need for perceptual resolution, so it predicts that eye movements will be directed to the most dense parts of the display, where resolution is the poorest. It does not predict that every object will be fixated (unlike the indexing hypothesis) or that no object will be refixated (unlike the control hypothesis).

The literature provides some support for each hypothesis. The indexing hypothesis is supported by many studies that show that eye movements are important in enumeration (Atkinson, Campbell, & Francis, 1976; Gippenreiter, Romanov, & Smirnov, 1969; Kowler & Steinman, 1977, 1979; Noro, 1980; Simon & Vaishnavi, 1996; van Oeffelen & Vos 1984). The clearest evidence comes from Watson et al. (2007), who found that subjects make one eye movement for each object in the display, beyond the subitizing range. In the counting range, the slope of the function relating the number of eye movements to the number of objects was 1.0—essentially, a perfect confirmation of the indexing hypothesis. However, some evidence suggests that indexing with eye movements may not be necessary. People can count without moving their eyes, albeit with higher error rates (Simon & Vaishnavi, 1996) and slower counting rates (Watson et al., 2007). The present experiments sought more direct evidence for the indexing hypothesis by examining where subjects look while they are counting. If they are indexing display objects with eye movements, they should fixate every object and never refixate any of them.

Support for the control hypothesis was provided by Watson et al. (2007), who found that fixation duration

was longer when the next saccade was in a direction opposite to the previous saccade. They interpreted this as evidence for inhibition of return (Posner & Cohen, 1984), which is a phenomenon that reflects the operation of the eye-movement control system (Klein, 2000). The present experiments sought more direct evidence of inhibition of return, asking whether subjects refixated any objects and if their tendency to refixate was modulated by the lag between the first fixation and the second (Beck et al., 2006; Peterson et al., 2001).

Support for the perception hypothesis comes from studies that show that subjects tend to fixate groups of objects instead of individual objects (Gippenreiter et al., 1969; van Oeffelen & Vos, 1984). When displays become small and dense, subjects no longer fixate each object and accuracy fails (Gippenreiter et al., 1969). The present experiments sought more direct evidence for the perception hypothesis, asking whether subjects were more likely to fixate objects in dense versus sparse regions of the displays, that is, more likely to fixate objects that had many near neighbors than objects that were isolated.

We conducted two experiments to distinguish between the three hypotheses by observing the eye movements while subjects enumerated displays of 5–12 dots. Experiment 1 compared eye movements in the counting task with those in a “look” task, in which subjects were told to look at each dot in a pattern once and only once. The look task requires the indexing component of the count task and requires subjects to use eye movements to implement the indexing. Thus, the look task provides an estimate of what performance would be like if the indexing hypothesis were true and subjects actually looked directly at each dot. If subjects use eye movements to index dots in the counting task, their performance should resemble that in the look task in important ways.

Experiment 2 presented the same dot patterns to every subject twice, so we could measure the consistency with which dots were fixated between and within subjects. We were interested in how consistent the eye-movement patterns are within subjects, when subjects view the same pattern twice, and how consistent they are across subjects. Both versions of the indexing hypothesis predict that the number of fixations should be quite consistent when subjects viewed the same pattern twice. Hence, different eye movement patterns, especially different numbers of fixations, will reject the indexing hypothesis.

We quantified the eye movements in three ways. (1) We counted the *number of fixations* in each display and calculated the number of dots that were fixated. The indexing hypothesis predicts that the number of fixations should equal the number of dots and that every dot should be fixated. The control and perception hypotheses allow fewer fixations than dots. (2) We counted the *number of refixations* and the lag between refixations of the dots. The indexing and control hypotheses predict there should be no refixations other than immediate refixations on corrective saccades. The perception hypothesis allows refixations, particularly in dense regions. (3) We calculated the *locations of fixations* relative to the dots to determine whether each dot was fixated and to determine the characteristics

of dots that were and were not fixated. The indexing hypothesis predicts that each and every dot will be fixated. The control hypothesis allows fewer fixations than dots. The perception hypothesis predicts that isolated dots with far neighbors are less likely to be fixated than dense dots with near neighbors.

What is the relation between the three hypotheses? We introduced them as alternatives and discussed them as if they were mutually exclusive, but they need not be. The three hypotheses may reflect different factors that compete for control of eye movements. The cognitive counting procedure promotes eye movements to the object it is currently indexing; the eye-movement control system promotes eye movements to objects that have not been fixated recently; and the perceptual sampling process promotes eye movements to regions where resolution is most limited. Each fixation may reflect the result of a different competition between these factors, so no single factor may govern the entire sequence of fixations. On the one hand, the implications would be clearest if one factor won all the competitions, as we require in the look condition of Experiment 1. If that factor was the one that underlies the indexing hypothesis, then eye movements would mirror the soul of the counting process. If it was perception, then eye movements would not mirror counting and researchers might look elsewhere for its reflection. On the other hand, the implications of a competition between different factors that determine eye movements are intriguing. If the results suggest that several factors are at work, we will learn which ones they are and, perhaps, the conditions under which they exert greater and lesser influence.

## EXPERIMENT 1

The first experiment evaluated the indexing, control, and perception hypotheses by comparing eye movements, RTs, and accuracy between two conditions: a *count* condition, in which subjects were asked to report the numerosity of the display, and a *look* condition in which subjects were asked to fixate each and every dot without reporting the numerosity. The look condition should require the indexing and control processes in the counting model. If eye movements are tightly coupled to indexing and control processes, the count condition should show the same pattern of eye movements as the look condition. If eye movements are not tightly coupled, then the pattern of eye movements should diverge, and the divergence should be greater the less tight the coupling.

### Method

**Subjects.** Ten undergraduate students from Vanderbilt University were paid to participate in the experiment. They all had normal uncorrected vision or their vision was corrected via contact lenses or glasses.

**Apparatus.** Stimuli were presented on a 21-in. CRT monitor with a resolution of  $1,024 \times 768$  pixels and a refresh rate of 85 Hz, controlled by a personal computer. A chinrest located 57 cm away from the monitor was used to minimize head movements. Eye movements were monitored via an EyeLink 2 tracker with eye position sampled at 250 Hz. Eye movements were tracked by the positions of the pupil and corneal reflection of the right eye. Drift correction

was conducted at the beginning of each trial. The eye tracker was calibrated at the beginning of the experiment. A validation was run immediately after calibration to make sure the maximal error was less than 1° and the average error was less than 0.5°. During calibration and validation, subjects looked at a dot presented in a 3 × 3 grid in a random order.

**Materials.** Twenty different dot patterns were randomly generated for each numerosity (5–12). All of the subjects looked at the same set of patterns. To generate a pattern, the location of each dot was sampled randomly from a 460 × 410 pixel array. Dots were sampled sequentially and added to the array with the constraint that the minimal distance between the other dots was larger than 1°. The dots in each pattern were displayed on an 18 × 16 cm area at the center of the display, which subtended about 18° × 16° visual angle. All dots had a diameter of 2.4 mm (0.24°) and were at least 1 cm (1°) away from each other. Dots were displayed as white [RGB = (255,255,255)] on a black background [RGB = (0,0,0)].

**Procedure.** There were two 160-trial experimental blocks: one count block and one look block. There were 16 practice trials at the beginning of each block. Half of the subjects started with the count block; the other half started with the look block. Each pattern was shown once in each experimental block. The 16 practice trials included two patterns for each numerosity, which were different from those used in the experimental blocks. The order of the trials was randomized within a block. Each trial started with drift correction, followed by the display with dots. Drift correction required subjects to look at the central fixation point and press one button on a Microsoft SideWinder Game Pad. In the look block, subjects looked at individual dots without counting them. They pressed a button when they finished, and the next trial started 1 sec later. In the count task, subjects counted the dots and pressed a button when they finished. To decrease head movements, subjects reported the number of dots in the display by fixating a number from 5 to 12 (presented as digits in the center of a 4° circle) that appeared in a row above and a row below the region in which the dots appeared. Subjects were required to look at the correct number for 300 msec, whereupon the fixated number changed from black to green. Subjects needed to fixate that number for an additional 500 msec to confirm the selection, after which the number turned red. An error response was signaled by a sound (Utopia Program Error clip; Microsoft Windows 2000). The experiment took about 45 min. Subjects were allowed breaks after the practice blocks and after every 80 experimental trials. Instructions were given at the beginning of each task.

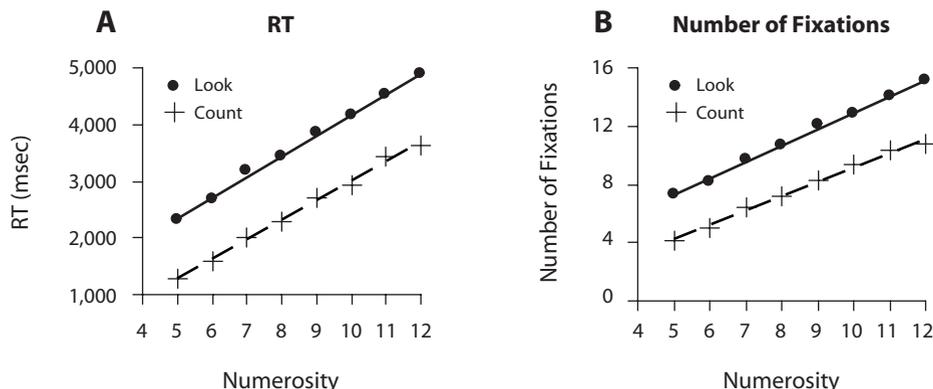
**Results and Discussion**

**Reaction time and accuracy.** In the count task, accuracy was .97 overall, and was not different for different numerosities or blocks ( $p > .1$ ). This suggests that subjects

enumerated the displays by counting, which is accurate, rather than estimation, which is inaccurate. RT for correct trials increased with numerosity. The increase was linear, with a slope of 354 msec/dot (calculated after excluding the extreme numerosities, 5 and 12). This also suggests that subjects enumerated the displays by counting, employing a serial process that successively indexes the dots and increments a counter. In these respects, the data from the present experiment replicate standard results in past studies of enumeration (Jensen et al., 1950; Klahr, 1973; Logan & Zbrodoff, 2003; Mandler & Shebo, 1982; Saltzman & Garner, 1948; Trick & Pylyshyn, 1994).

In the look task, RT also increased linearly with numerosity with a slope of 360 msec/dot. This slope was almost identical to the slope for the count task, which is consistent with the hypothesis that the look task and the count task involve the same indexing process. The slopes of the linear functions relating RT to numerosity were significantly greater than 0 [ $ts(9) = 13.53$  and  $11.61$ , respectively, for the look task and the count task;  $ps < .001$ ]. The intercept of the linear function relating RT to numerosity was 576 msec for the look task and -695 msec for the count task. Both intercepts were significantly different from 0 [ $ts(9) = 4.59$  and  $-3.41$  for look and count, respectively;  $ps < .01$ ]. One interpretation of the difference in intercepts is in terms of subitizing: Subjects index fewer dots in the count condition because they subitize 3–5 dots and only index the remaining ones. By contrast, subjects in the look condition must index all the dots. We can estimate the number of dots that were subitized by dividing the difference in intercepts (1,271 msec) by the estimated counting rate in the count condition (354 msec/dot). This yields a value of 3.6 dots, which is within the subitizing range (Logan & Zbrodoff, 2003; Mandler & Shebo, 1982). We will address this issue further in the General Discussion.

These conclusions were confirmed by a 2 (task: look, count) × 6 (numerosity) within-subjects ANOVA on mean RTs (Figure 1A). The extreme numerosities (5 and 12) were excluded in this and the following ANOVAs, following common practice in experiments on enumeration (Trick & Pylyshyn, 1993; Watson et al., 2007). Because



**Figure 1.** Mean response times (RTs) (A) and number of fixations (B) in Experiment 1 as a function of the number of objects presented.

preliminary analysis showed no main effect or interaction of task order on RT and number of fixations ( $F_s < 1$ ), the data from the two task orders were pooled. There was a main effect of numerosity [ $F(5,45) = 163$ ,  $MS_e = 54,846$ ,  $p < .001$ ,  $\eta_p^2 = .95$ ] and a main effect of task [ $F(1,9) = 33$ ,  $MS_e = 1,200,683$ ,  $p < .001$ ,  $\eta_p^2 = .79$ ]. There was no interaction between numerosity and task ( $F < 1$ ).

So far, the results are consistent with the hypothesis that subjects enumerated the displays with a serial indexing process, which predicts accurate enumeration and a linear increase in RT with numerosity. Moreover, the RT data in the count task were similar to RT data in the look task (except for an intercept shift), which also required indexing. Now we ask whether subjects' eye movements were coupled to that serial indexing process.

**Number of fixations.** Following Watson et al. (2007), we counted the number of fixations on each trial. Number of fixations increased linearly with numerosity in the count task with a slope of 1.03 fixations/dot (see Figure 1B). This is consistent with the indexing hypothesis. Number of fixations also increased linearly with numerosity in the look task with a slope of 1.14 fixations per dot, which was significantly larger than the slope in the count task [ $t(9) = 3.00$ ,  $p < .05$ ]. Possibly, subjects refixated some dots in the look task to be sure the computer would register their "look" at them. The intercepts of the linear functions relating number of fixations to numerosity were substantially different between conditions ( $M_s = 1.61$  for look and  $-0.96$  for count). As with the RT data, this may reflect the differential opportunity to subitize in the look and count conditions. We can estimate the number of dots that were subitized in the count condition by calculating the difference between intercepts. The value, 2.57 dots, is smaller than the 3–5 expected for subitizing, but not by much.

Mean numbers of fixations were submitted to a 2 (task: count, look)  $\times$  6 (numerosity) within-subjects ANOVA. The main effect of numerosity [ $F(5,45) = 246$ ,  $MS_e = 0.33$ ,  $p < .001$ ,  $\eta_p^2 = .97$ ], the main effect of task [ $F(1,9) = 95$ ,  $MS_e = 4$ ,  $p < .001$ ,  $\eta_p^2 = .91$ ], and the interaction between task and numerosity [ $F(7,63) = 3.47$ ,  $MS_e = 0.22$ ,  $p < .01$ ,  $\eta_p^2 = .28$ ] were all significant.

So far, the results are consistent with the indexing hypothesis, in which eye movements are tightly coupled to the serial indexing process that underlies enumeration. The indexing hypothesis predicts one eye movement for each object in the display. The count-task data disconfirmed a strong version of this hypothesis, showing that the number of fixations was always smaller than the number of dots in the display. The count-task data are consistent with a weaker version of this hypothesis, in which subjects first subitize and then index. This hypothesis predicts a slope of 1.0 in the linear function relating number of fixations to numerosity and allows the intercept to be less than 0, as we observed (and as observed by Watson et al., 2007). The indexing hypothesis also predicts that subjects will index every dot in the display, excluding those that are subitized. The remaining analyses focus on the location of the fixations—where we look when we count—to test that hypothesis.

Note that the observed slopes of 1.03 (count) and 1.14 (look) relate the mean number of fixations to numerosity. There was considerable variation in the number of fixations within each numerosity. To assess the importance of this variability, we calculated the correlation between the number of fixations on each trial and the numerosity of the display. The correlation was .62 for the count task and .85 for the look task. The difference between these correlations was significant [ $t(9) = 8.15$ ,  $p < .01$ ], suggesting that the patterns of eye movements were different for the look task and the count task. The correlation for the count task was significantly less than 1.0 [ $t(9) = 16.65$ ,  $p < .01$ ], suggesting that other factors besides indexing may influence the number of fixations. The correlation for the look task was also significantly less than 1.0 [ $t(9) = 10.54$ ,  $p < .01$ ].

**Do subjects fixate dots?** The first point to establish is whether fixations in the count and look tasks are directed toward the dots. An extreme version of the perception hypothesis might predict that the eyes roam around the densest regions of the display to improve resolution, independent of the counting process. To determine whether subjects fixated dots, we calculated the distribution of the distance between each fixation and the dot that was nearest to it (see Figure 2). The modal distance was  $0.5^\circ$ , which is the resolution of the eyetracker. For comparison, we calculated the distribution of the distance between each dot and the nearest neighboring dot for comparison. The modal distance was between  $2^\circ$  and  $3.5^\circ$ . The distributions of fixations for the count and look tasks were quite similar, except that the look-task distribution was shifted more to the left, reflecting slightly more precise fixations.

**Number of dots fixated.** We know that subjects fixate dots. The indexing hypothesis predicts that subjects fixate each and every dot once and only once, and the control hypothesis predicts that subjects fixate each dot only once. The numbers of fixations are consistent with these predictions, but only if each fixation was directed to a different dot. A more stringent test of the predictions involves identifying the dots that were actually fixated and determining whether each of those dots was fixated once and only once. To carry out this test, we counted the number of dots fixated in each display. For each fixation, we calculated its screen location and chose the dot that was closest to that location as a fixated dot. Then we calculated the number of dots that were fixated, excluding refixations on a dot that had been fixated previously. The means across subjects are plotted in Figure 3A.

Subjects fixated fewer dots than the number presented in both tasks. The average number of objects in the displays was 8.5 (mean of numerosities 5–12). Subjects fixated 5.24 dots in the count task and 7.88 dots in the look task. The former was significantly different from the number presented [ $t(9) = 8.54$ ,  $p < .001$ ], but the latter was not [ $t(9) = 1.62$ ,  $p > .1$ ]. The result with the count task challenges the indexing hypothesis, which predicts that each dot should be fixated once and only once. The number of fixated dots shows a slope shift rather than an intercept shift: The slope was 0.67 for the count task and 0.89 for the look task. These slopes were significantly smaller

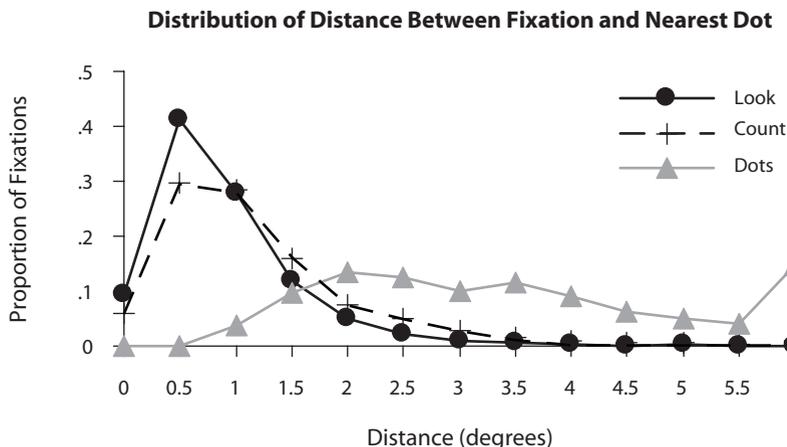


Figure 2. Distribution of the distances between the fixations and the nearest dot in Experiment 1. The gray line is the distribution of the distances between a dot and its nearest neighbor.

than 1.0 [ $t(9) = 11.51, p < .001$ , for the count task;  $t(9) = 3.83, p < .01$ , for the look task] and significantly different from each other [ $t(9) = 8.00, p < .001$ ]. The intercept was  $-0.41$  for the count task and  $0.30$  for the look task. The intercepts were not significantly different from 0 [ $t(9) =$

$1.42, p > .05$ , for the count task;  $t(9) = 2.11, p > .05$ , for the look task], though they were different from each other [ $t(9) = 2.94, p < .05$ ].

The mean numbers of fixated dots were submitted to a  $2$  (task: look, count)  $\times 6$  (numerosity)  $\times 2$  (order: look

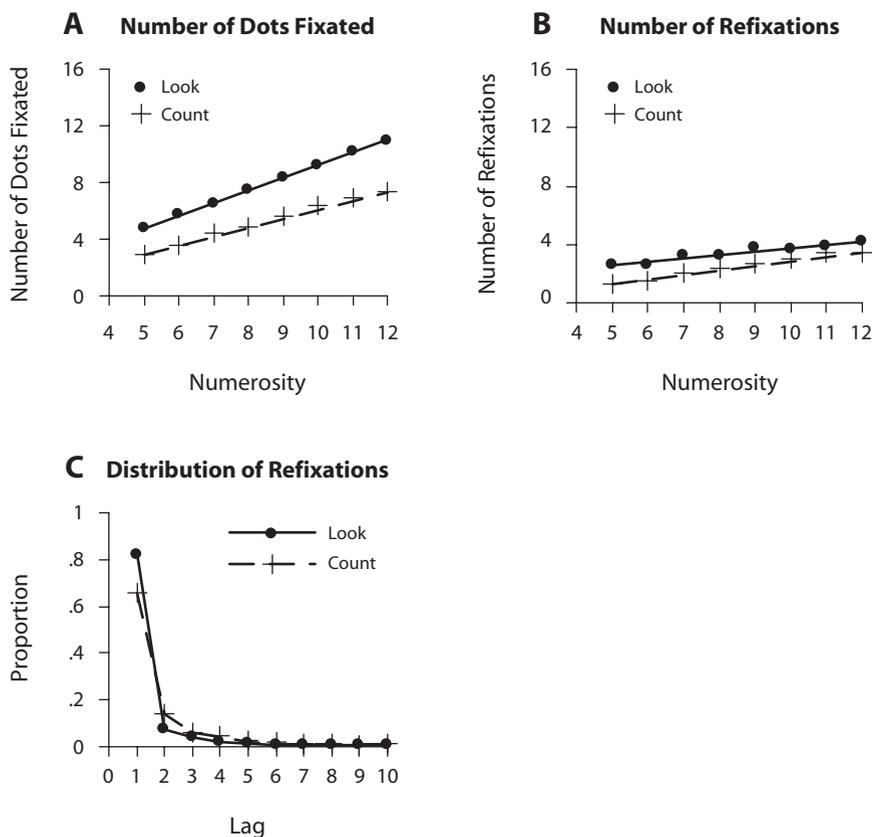


Figure 3. Refixation properties in Experiment 1. (A) Number of dots fixated as a function of the number of objects presented. (B) Number of refixations as a function of the number of objects presented. (C) Distribution of refixations as a function of lag between the two fixations.

task first or count task first) ANOVA. There were significant main effects of numerosity [ $F(5,40) = 514$ ,  $MS_e = 0.09$ ,  $p < .001$ ,  $\eta_p^2 = .99$ ] and task [ $F(1,9) = 325$ ,  $MS_e = 0.64$ ,  $p < .001$ ,  $\eta_p^2 = .98$ ] and a significant interaction between task and numerosity [ $F(5,40) = 21$ ,  $MS_e = 0.05$ ,  $p < .001$ ,  $\eta_p^2 = .72$ ]. The number of fixations increased with numerosity, but there were fewer fixations in the count task than in the look task, and that difference increased with numerosity. There was also an interaction between task and order [ $F(1,8) = 26$ ,  $MS_e = 0.64$ ,  $p < .001$ ,  $\eta_p^2 = .77$ ]. More items were fixated in the count task when the look task was conducted first (5.9 items) than when the count task was conducted first (4.7 items), suggesting a carryover effect. In comparison, an average of 7.8 items were fixated in the look task when the look task was conducted first, and 8.0 items were fixated when the count task was conducted first. Since this interaction did not change the pattern of our major results, we did not explore this effect further.

These data suggest that subjects did not fixate every dot in each display, which contradicts the indexing hypothesis. These data also suggest that subjects did not fixate each dot only once, which contradicts the control hypothesis.

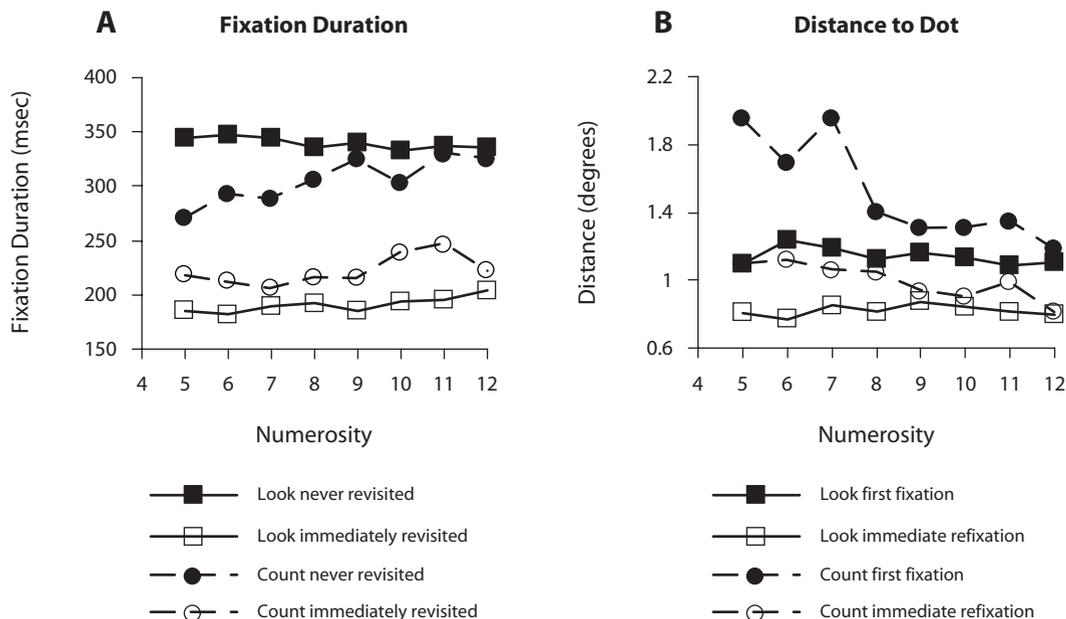
**Number of refixations.** The indexing and control hypothesis predict that subjects should fixate each dot only once and that subjects should not refixate any dots. The data show that there are more fixations than fixated dots in each display, which implies that some dots must have been refixated. Indeed, 29% of the fixations in the look task were refixations and 27% of the fixations in the count task were refixations. We calculated the number of refixations by counting the number of times the same dot was fixated in each display. The number of refixations was greater in the look task (3.36) than in the count task (2.48) and increased with numerosity in each task (see Figure 3B). Number of refixations increased linearly with numerosity with a slope significantly greater than 0 in both tasks [look task, 0.24 refixations/numerosity,  $t(9) = 5.22$ ,  $p < .001$ ; count task, 0.36 refixations/numerosity,  $t(9) = 6.84$ ,  $p < .001$ ]. The slopes were significantly different from each other [ $t(9) = 3.03$ ,  $p < .05$ ]. Mean numbers of refixations were submitted to a 2 (task: count, look)  $\times$  6 (numerosity)  $\times$  2 (order: look task first or count task first) ANOVA. The main effects of numerosity [ $F(5,40) = 24.20$ ,  $MS_e = 0.27$ ,  $p < .001$ ,  $\eta_p^2 = .75$ ] and task [ $F(1,8) = 10.77$ ,  $MS_e = 2.2$ ,  $p < .05$ ,  $\eta_p^2 = .57$ ] were significant. There was also an interaction between task and task order [ $F(1,8) = 5.74$ ,  $MS_e = 2.2$ ,  $p < .05$ ,  $\eta_p^2 = .42$ ]. Task order affected the number of refixations in the look task more than that in the count task. For the look task, there were 3.8 refixations when the look task was first and 2.9 refixations when the count task was first. In contrast, for the count task, there were 2.3 refixations when the look task was first and 2.7 refixations when the count task was first. Again, since this interaction did not change the pattern of our major results, we did not explore this effect further.

To separate corrective saccades from other refixations, we sorted the data according to the lag between the first fixation and the refixation. The proportions of the total number of refixations (see Figure 3C) were higher for im-

mediate refixations than for longer lags: .66 in the count task and .82 in the look task were immediate refixations. The immediate refixations may have been corrective saccades (Henson, 1978; Rayner, 1998). Two additional analyses showed that the immediate refixations were likely to be corrective saccades. First, fixation duration was shorter when it preceded an immediate refixation (206 msec; see Figure 4A) than when it did not (322 msec). The observation was supported by a main effect of fixation type [ $F(1,9) = 69.72$ ,  $MS_e = 11,780$ ,  $p < .001$ ,  $\eta_p^2 = .89$ ] in a 2 (task)  $\times$  6 (numerosity)  $\times$  2 (fixation type: immediate refixation or not) ANOVA. Because preliminary analysis showed that there was no main effect or interaction of task order ( $F_s < 1$ ), it was not included in the analysis. There was also a main effect of numerosity [ $F(5,45) = 2.70$ ,  $MS_e = 779$ ,  $p < .05$ ,  $\eta_p^2 = .23$ ]. Fixation duration increased with numerosity. There was an interaction between task and numerosity [ $F(5,45) = 4.36$ ,  $MS_e = 498$ ,  $p < .001$ ,  $\eta_p^2 = .33$ ] and an interaction between task and fixation type [ $F(1,9) = 22.52$ ,  $MS_e = 2,806$ ,  $p < .001$ ,  $\eta_p^2 = .71$ ]. The interaction between task and numerosity reflects the fact that fixation duration increased with numerosity in the count task but not in the look task. The interaction between task and fixation type reflects the fact that the difference between two fixation types was larger in the look task (149 msec) than in the count task (84 msec). The difference was significant in both the look task [ $F(1,9) = 118.68$ ,  $MS_e = 2,806$ ,  $p < .001$ ,  $\eta_p^2 = .93$ ] and the count task [ $F(1,9) = 37.72$ ,  $MS_e = 2,806$ ,  $p < .001$ ,  $\eta_p^2 = 0.81$ ].

Second, the immediate refixation was closer to the nearest dot than the previous fixation ( $0.92^\circ$  vs.  $1.32^\circ$ ; see Figure 4B). This was confirmed by a main effect of fixation type [ $F(1,9) = 10.06$ ,  $MS_e = 0.19$ ,  $p < .001$ ,  $\eta_p^2 = .86$ ] in a 2 (task)  $\times$  6 (numerosity)  $\times$  2 (fixation type: first fixation or immediate refixation) ANOVA on the distance to the closest dot. There was also a main effect of task [ $F(1,9) = 8.06$ ,  $MS_e = 0.49$ ,  $p < .05$ ,  $\eta_p^2 = .47$ ]. Fixations were closer to the dots in the look condition ( $1.0^\circ$ ) than in the count condition ( $1.25^\circ$ ). In summary, these analyses suggest that immediate refixations were caused by corrective saccades.

After excluding immediate refixations, there were still 5% nonimmediate refixations in the look task and 9% in the count task (see Figure 5). These nonimmediate refixations contradict the indexing hypothesis and the strong version of the control hypothesis, which predict there should be no nonimmediate refixations. From the analyses of fixation locations, we know that many dots were revisited. Although these results suggest that the eyes do not have a perfect memory, it is possible that they have some memory for previous fixations (Beck et al., 2006; Peterson et al., 2008; Peterson et al., 2001). To examine this possibility, we compared the observed proportion of nonimmediate refixations with that predicted by a no-memory model. The no-memory model assumes that eye movements have no memory at all, and the target of each fixation is selected randomly from all of the dots (with replacement). If the eye movements do not have any memory at all, we expect that the proportion of refixations will



**Figure 4.** Some fixation properties in Experiment 1. (A) Fixation duration as a function of whether the fixation has an immediate refixation or not. (B) Distance to the closest dot as a function of whether the fixation is the first fixation on a dot or a refixation.

be similar to that predicted by the no-memory model. If eyes have some memory during counting, we expect that the observed proportion of nonimmediate refixations will be less than that predicted by the no-memory model.

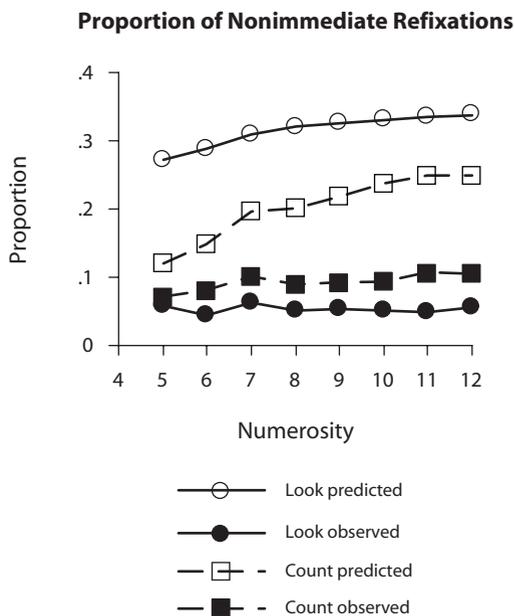
The expected results of no-memory eye movements were generated by a Monte Carlo simulation (see Peterson et al., 2001, for a similar approach). The number of fixations in a trial is the closest integer to a number sampled from a normal distribution with a mean equal to the observed mean number of fixations and a standard deviation equal to the observed standard deviation for that numerosity. The simulation was run using 10 “subjects,” with each subject receiving 20 trials for each numerosity in each condition, which was identical to the design of Experiment 1. The results are shown in Figure 5.

The predicted proportion of nonimmediate refixations is greater than the observed values. The observed and predicted data were compared using multiple *t* tests with a Bonferroni correction for the number of tests performed. The observed data were significantly different from those predicted by the no-memory model except the count condition when the numerosity was 5 [ $t(18) = 3.10, p = .006$ ]. This suggests that eye movements have some memory, which supports the weak version of the control hypothesis.

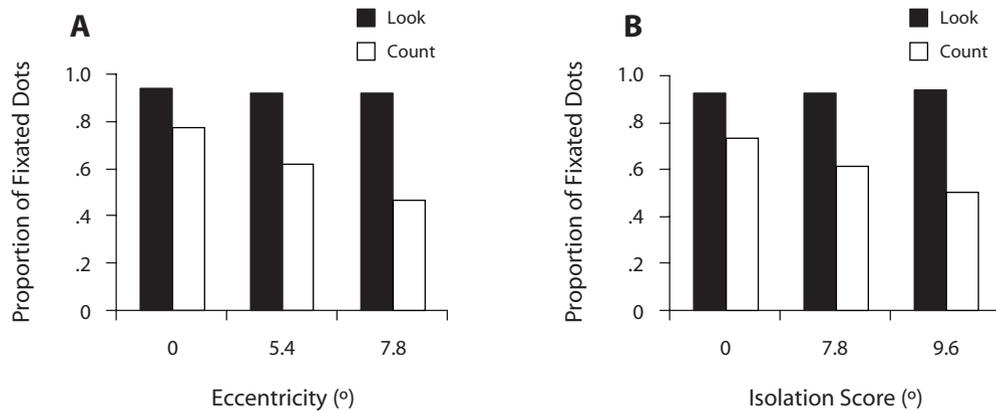
**Isolation and eccentricity.** The analyses so far have shown that subjects do not fixate every dot in the count task. This raises the question, which dots do they not fixate? The perception hypothesis predicts that subjects will be likely to fixate dots in dense regions in the center of the display and will not be likely to fixate dots that are isolated or in peripheral regions of the display. We conducted two related tests of these predictions. First, we assessed

the effects of eccentricity. We classified the eccentricities of dots (distance from the central fixation point) into three categories (low, 0°–5.4°; medium, 5.4°–7.8°; high, larger than 7.8°). The boundaries were chosen so that there were equal numbers of dots in each cell.

Subjects in the count condition showed strong effects of eccentricity (see Figure 6A). They were increasingly less



**Figure 5.** Proportions of nonimmediate refixations in Experiment 1 as a function of the number of objects presented.



**Figure 6.** (A) Proportions of dots that were fixated as a function of eccentricity in Experiment 1. (B) Proportions of dots that were fixated as a function of the average distance between a dot and all of the other dots.

likely to fixate dots as eccentricity increased. By contrast, subjects in the look condition were unaffected by eccentricity. They were equally likely to fixate dots regardless of eccentricity, consistent with their instructions.

The proportions of dots that were fixated were submitted to a 2 (task: look, count)  $\times$  3 (eccentricity) within-subjects ANOVA. The proportion was higher in the look task ( $M = .93$ ,  $SE = .02$ ) than the count task ( $M = .62$ ,  $SE = .04$ ) [ $F(1,9) = 91.44$ ,  $MS_e = 0.02$ ,  $p < .001$ ,  $\eta_p^2 = .91$ ]. There was a significant main effect of eccentricity [ $F(2,18) = 104.29$ ,  $MS_e = 0.001$ ,  $p < .001$ ,  $\eta_p^2 = .92$ ] and a significant interaction between task and eccentricity [ $F(2,18) = 115.53$ ,  $MS_e = 0.001$ ,  $p < .001$ ,  $\eta_p^2 = .93$ ]. Eccentricity had strong effects in the count task but none in the look task.

The perception hypothesis predicts that isolation is also a critical variable. Subjects must fixate on dots in dense neighborhoods to resolve them perceptually, but isolated dots may be resolved without being fixated. To test this prediction, we calculated an isolation score for each dot based on the mean distance to the other dots. The isolation scores of the dots were categorized into three cells (low,  $0^\circ$ – $7.8^\circ$ ; medium,  $7.8^\circ$ – $9.6^\circ$ ; high, larger than  $9.6^\circ$ ). The data from the count condition show a strong effect of isolation (Figure 6B). Subjects were increasingly less likely to fixate dots as isolation increased. By contrast, subjects in the look condition were unaffected by isolation, fixating every dot regardless of its location, in accord with their instructions.

The proportions of dots that were fixated were submitted to a 2 (task: look, count)  $\times$  3 (isolation: high, medium, low) within-subjects ANOVA. The proportion was higher in the look task ( $M = .93$ ,  $SE = .02$ ) than the count task ( $M = .62$ ,  $SE = .03$ ) [ $F(1,9) = 91.13$ ,  $MS_e = 0.02$ ,  $p < .001$ ,  $\eta_p^2 = .91$ ]. There was a significant main effect of isolation [ $F(2,18) = 42.92$ ,  $MS_e = 0.001$ ,  $p < .001$ ,  $\eta_p^2 = .83$ ]. There was an interaction between task and isolation [ $F(2,18) = 67.47$ ,  $MS_e = 0.001$ ,  $p < .001$ ,  $\eta_p^2 = .88$ ]. Isolation had strong effects in the count task but none in the look task.

Note that the eccentricity and the isolation were not entirely independent. Dots closer to the fixation point were also more likely to be closer to the other dots. Indeed, the correlation between isolation and eccentricity was .24. However, eccentricity and isolation measured different aspects of the spatial distribution of the dots. Dots may be close to the fixation but isolated from the other dots or eccentric with near neighbors. The present study cannot determine which factor played a more important role in capturing the eyes.

## Conclusions

The experiment replicated classic effects in the enumeration literature. Subjects enumerated very accurately and did so with a process whose duration increased linearly with numerosity. These data suggest that subjects enumerated by counting. The eye movement data addressed the coupling between eye movements and enumeration and the theories that made predictions about it. (1) The number of fixations increased linearly with numerosity with a slope of 1, which suggests that subjects might fixate every dot, but (2) analysis of the number of fixated dots showed a linear relation with a slope of 0.67, indicating that subjects did not fixate every dot. (3) Subjects had a tendency to avoid fixating dots that they had fixated before, although they could not do so perfectly. (4) Analysis of eccentricity and isolation suggests that subjects fixated dots in dense, central regions of the display and tended not to fixate dots in sparse, eccentric regions of the display.

Result 1 is consistent with the indexing hypothesis, which predicts that subjects fixate every dot once and only once, but Results 2–4 are inconsistent with it. Results 1, 2, and 4 are consistent with a strong version of the control hypothesis, which predicts that subjects fixate dots only once, but Result 3 is consistent with a weak version of the control hypothesis. The perception hypothesis is consistent with all four results. It makes no strong predictions about Results 1, 2, and 3 but it predicts the observed eccentricity and isolation effects in Result 4. In sum, the results suggest a loose coupling of eye move-

ments and the cognitive processes that underlie enumeration. These conclusions are strengthened by the results of the look task, which required indexing. It produced a slope slightly larger than 1 in the number-of-fixations analysis and a slope just less than 1 in the number-of-fixated-dots analysis, and showed no effects of isolation or eccentricity, all of which are consistent with indexing. It produced a tendency toward immediate refixation, like the look task, contradicting a strong version of the control hypothesis. Thus, eye movements can show evidence of indexing when subjects are required to index. This makes the fact that they do not show such evidence when they are not required to index all the more impressive.

## EXPERIMENT 2

Experiment 1 suggests a loose coupling between eye movements and enumeration. The RT and accuracy data suggest that subjects were counting, but the eye-movement data suggested that their fixations were not driven exclusively by the enumeration process but instead were attracted to dense regions in the center of the display. That is, the data contradicted the indexing hypothesis, supported a weak version of the control hypothesis, and supported the perception hypothesis. This outcome was surprising, although we will rationalize it in the General Discussion, so we sought to replicate it in a new experiment. Experiment 2 examined the coupling between eye movements and enumeration in two additional ways. First, Experiment 2 compared counting in conditions that were designed to highlight the effects predicted by the perception hypothesis. Subjects saw the same displays twice, and their performance on the two exposures was compared. If eye movement behavior was determined entirely by perceptual factors, then it should be the same on the two exposures because the displays were the same. Moreover, every subject saw the same displays. If eye movement behavior was determined entirely by properties of the display, then it should be the same for different subjects viewing the same displays. There should be as much consistency between subjects as within subjects. Second, Experiment 2 tested the indexing hypothesis further. Both versions of the indexing hypothesis predict that the number of fixations should be the same when subjects view the same pattern twice. Hence, different numbers of fixations will be strong evidence against the indexing hypothesis.

These are strong predictions that are unlikely to be true. Compton and Logan (1999) investigated the reliability of perceptual grouping of random patterns of 6–11 dots, similar to those in the present experiments. They presented the same dot patterns twice and had subjects draw circles around the perceptual groups they saw. Subjects chose the same perceptual groups on the two exposures on only .34 of the trials, suggesting weak determination from the stimulus. If grouping processes drive eye movements (Zelinsky, Rao, Hayhoe, & Ballard, 1997), then subjects may look differently at the same display on different occasions. Nevertheless, failures of perfect determination will be interesting.

We presented displays twice in Experiment 2, so it is possible that memory for the first presentation could influence eye movement behavior on the second exposure. We think that is unlikely. Subjects saw 160 different patterns in each exposure, so the number of patterns to be remembered was large. Previous studies showed that subjects are unlikely to use memory retrieval when they view patterns only twice (Lassaline & Logan, 1993; Palmeri, 1997). Compton and Logan (1999) presented dot patterns twice and had people perform recognition judgments on the second presentation. Memory performance was barely above chance.

## Method

**Subjects.** Twelve undergraduate students from Vanderbilt University were paid to participate in the experiment. None of them had participated in Experiment 1.

**Apparatus.** The same apparatus was used as that in Experiment 1.

**Materials.** The patterns were constructed in the same way as in Experiment 1, and the same set of 20 patterns for each numerosity was used for all of the subjects.

**Procedure.** There were two 160-trial experimental blocks. Each pattern was shown to subjects twice, once in each block. There were 16 practice trials at the beginning of the experiment. The order of the trials was randomized within a block. The procedure for each trial was the same as that in the count task in Experiment 1. Each subject took about 45 min.

## Results and Discussion

**Reaction time and accuracy.** Accuracy was .98 and was not different for different numerosities or blocks ( $ps > .1$ ). RTs increased linearly with numerosity, with a slope of 315 msec/dot in Block 1 and 330 msec/dot in Block 2 (Figure 7A), and were not significantly different from each other [ $t(11) = 1.39, p > .1$ ]. The intercepts were  $-320$  msec for Block 1 and  $-519$  for Block 2, and were significantly different from each other [ $t(11) = 2.46, p < .05$ ]. Both were significantly smaller than 0 [ $t(11) = 2.72, p < .05$ , for Block 1;  $t(11) = 5.67, p < .001$ , for Block 2]. The RT and accuracy data are consistent with an enumeration process that relies on serial indexing (Jensen et al., 1950; Klahr, 1973; Logan & Zbrodoff, 2003; Mandler & Shebo, 1982; Saltzman & Garner, 1948; Trick & Pylyshyn, 1994). The question is how tightly the eye movements are coupled to this process.

These conclusions were supported by a 2 (block)  $\times$  8 (numerosity) within-subjects ANOVA on the mean RTs. There was a significant main effect of numerosity [ $F(7,77) = 160.90, MS_e = 81,030.42, p < .001, \eta_p^2 = .94$ ]. The main effect of block [ $F(1,11) = 1.72, p > .1$ ] and the interaction between block and numerosity ( $F < 1$ ) were not significant.

**Number of fixations.** Number of fixations increased linearly with numerosity with a slope of 0.90 fixations/dot for Block 1 and 0.88 fixations/dot for Block 2 (Figure 7B). These slopes were not significantly different from each other [ $t(11) = 0.55, p > .1$ ] but were significantly larger than 0 [ $ts(11) = 21.20$  and  $19.74$ , respectively, for Block 1 and Block 2,  $ps < .001$ ], and significantly smaller than 1 [ $ts(11) = 2.45$  and  $2.75$  for Blocks 1 and 2,

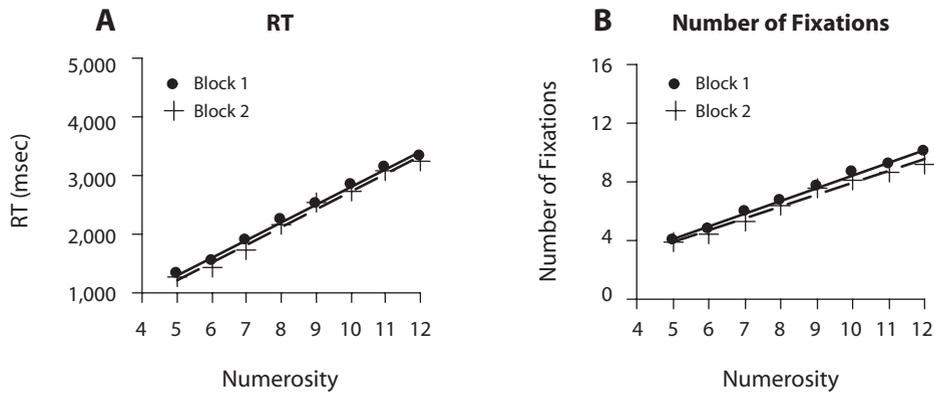


Figure 7. Mean response times (RTs) (A) and number of fixations (B) in Experiment 2 as a function of the number of objects presented.

$p < .05$ ]. The intercepts ( $-0.47$  and  $-0.73$  for Blocks 1 and 2, respectively) were not significantly different from 0 [ $t(11) = 1.40$  ( $p > .1$ ) and  $2.01$  ( $p < .1$ ) for Blocks 1 and 2, respectively] and were not significantly different from each other [ $t(11) = 0.90$ ,  $p > .1$ ]. These results contradict the indexing hypothesis, which predicts a slope of 1.0 in the function relating number of fixations to numerosity. The observed slopes were close to 1.0, however, so the contradiction is not severe.

Mean number of fixations was submitted to a 2 (block)  $\times$  8 (numerosity) within-subjects ANOVA. There were significant main effects of numerosity [ $F(7,77) = 229.09$ ,  $MS_e = 0.45$ ,  $p < .001$ ,  $\eta_p^2 = .95$ ] and block [ $F(1,11) = 5.34$ ,  $MS_e = 1.71$ ,  $p < .05$ ,  $\eta_p^2 = .33$ ], but the interaction between block and numerosity was not significant [ $F(7,77) = 1.72$ ,  $p > .1$ ].

We calculated the correlation between the number of fixations on each trial and the numerosity of the display. The mean value across subjects was .62 collapsed over Blocks 1 and 2 (.62 for Block 1, and .59 for Block 2), which is substantially less than 1.0, suggesting that something other than indexing contributed to the number of fixations.

**Distribution of fixations.** The distribution of the distance between each fixation and its nearest dot is shown in Figure 8. The modal distance was between  $0.5^\circ$  and  $1^\circ$ , which is consistent with Experiment 1.

**Number of dots fixated.** The number of fixated dots increased linearly with numerosity. The slopes (0.60 and 0.57 fixations/dot) were about 65% as large as the slopes of the dots were fixated (Figure 9A). This contradicts the strong version of the indexing hypothesis, which says that every dot should be fixated, and it contradicts a weaker version, which says that subjects subitize three to five dots and fixate the rest.

Mean number of unique dots was submitted to a 2 (block)  $\times$  8 (numerosity) within-subjects ANOVA. There were significant main effects of numerosity [ $F(7,77) = 271.26$ ,  $MS_e = 0.16$ ,  $p < .001$ ,  $\eta_p^2 = .96$ ] and block [ $F(1,11) = 7.20$ ,  $MS_e = 0.45$ ,  $p < .05$ ,  $\eta_p^2 = .40$ ]. The interaction between block and numerosity was not significant [ $F(7,77) = 1.45$ ,  $p > .1$ ].

**Number of refixations.** About 0.29 fixations in Block 1 and 0.28 fixations in Block 2 revisited a dot that

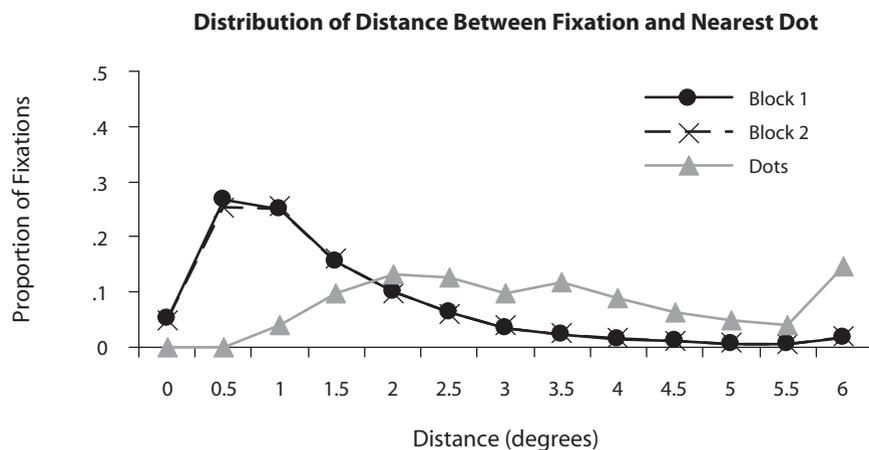
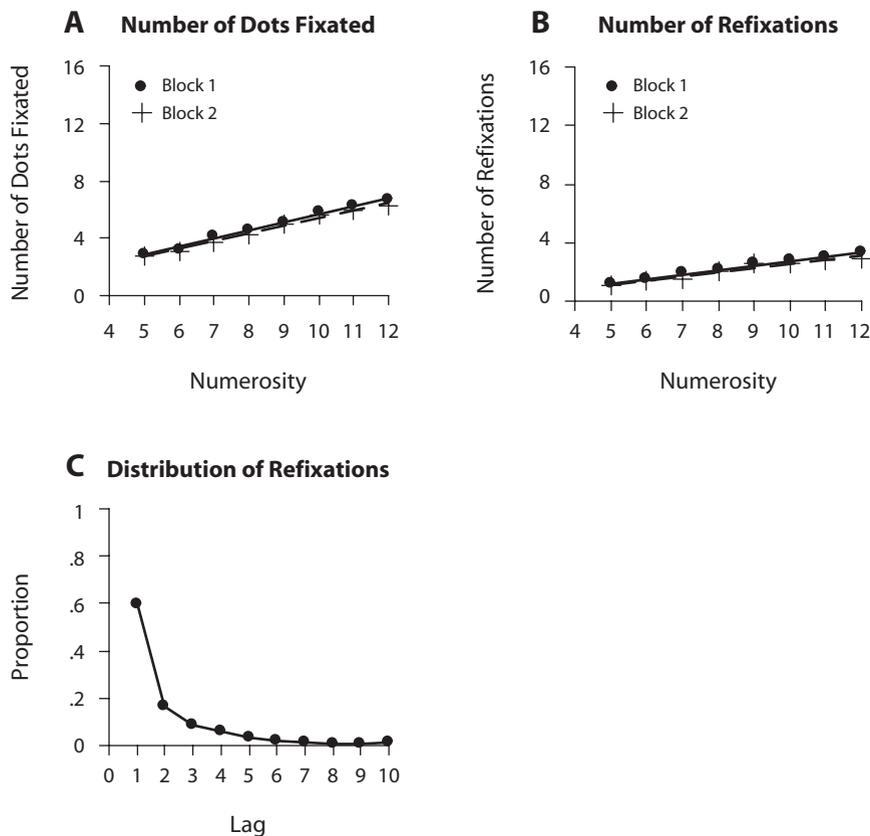


Figure 8. Distribution of the distance between the fixations and the nearest dot in Experiment 2. The gray line is the distribution of the distances between a dot and its nearest neighbor.



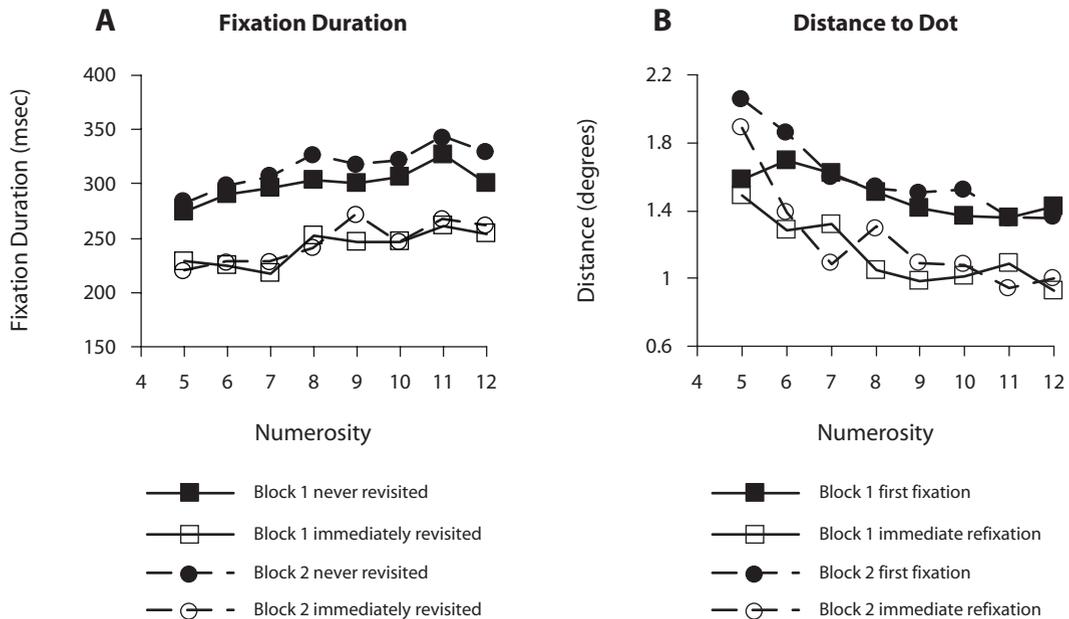
**Figure 9. Properties of refixations in Experiment 2. (A) Number of dots fixated as a function of the number of objects presented. (B) Number of refixations as a function of the number of objects presented. (C) Distribution of refixations as a function of lag between the two fixations.**

had been fixated before (Figure 9B). Number of refixations increased linearly with numerosity, with a slope significantly greater than 0 in both blocks [0.30 refixations/numerosity for each block;  $t_s(11) = 7.54$  and  $7.52$  for Blocks 1 and 2, respectively;  $p_s < .001$ ]. The slopes were not significantly different from each other [ $t(11) = 0.82$ ,  $p > .1$ ]. This is inconsistent with the indexing and control hypotheses.

The proportion of refixation as a function of lag between the current and previous fixations is shown in Figure 9C. Two additional analyses showed that immediate refixations were likely to be corrective saccades. First, fixation duration was shorter when it preceded immediate refixation (244 msec; see Figure 10A) than when it did not (311 msec). The observation was supported by a main effect of fixation type [ $F(1,11) = 33.14$ ,  $MS_e = 9,878$ ,  $p < .001$ ,  $\eta_p^2 = .75$ ] in a 2 (block)  $\times$  6 (numerosity)  $\times$  2 (fixation type: immediate refixation or not) ANOVA. There was a main effect of numerosity [ $F(5,55) = 9.87$ ,  $MS_e = 1,069$ ,  $p < .001$ ,  $\eta_p^2 = .47$ ]. Fixation durations increased with numerosity. There was also a main effect of block [ $F(1,11) = 8.00$ ,  $MS_e = 765$ ,  $p < .05$ ,  $\eta_p^2 = .42$ ]. Fixation durations were longer in Block 2 than in Block 1. Since the main effect of numerosity and block were not related to the purpose of this analysis, we will not discuss them further.

Second, immediate refixations were usually closer to the nearest dot than were previous fixations ( $1.13^\circ$  vs.  $1.53^\circ$ , see Figure 10B). This was confirmed by a main effect of fixation type [ $F(1,11) = 107.41$ ,  $MS_e = 0.10$ ,  $p < .001$ ,  $\eta_p^2 = .91$ ] in a 2 (block)  $\times$  6 (numerosity)  $\times$  2 (fixation type: first fixation or immediate refixation) ANOVA on the distance to the closest dot. There was also a main effect of numerosity [ $F(5,55) = 7.87$ ,  $MS_e = 0.12$ ,  $p < .001$ ,  $\eta_p^2 = .42$ ]. Distance to the nearest dot decreased with increasing numerosity, reflecting the construction of the displays (objective distance between dots decreased when numerosity increased). These results suggest that immediate refixations were likely to be corrective saccades.

After excluding immediate refixations, there were still 12% nonimmediate refixations in Block 1 and 10% in Block 2 (see Figure 5). These nonimmediate refixations contradict the indexing hypothesis and the strong version of the control hypothesis, which predict there should be no nonimmediate refixations. To examine whether eye movements have memory in counting, we calculated proportions of nonimmediate refixations predicted by a no-memory model using Monte Carlo simulation, as we did in Experiment 1. The simulation was run using 12 “subjects,” with each subject receiving 20 trials for each numerosity in each of the two blocks. The results are shown



**Figure 10.** Some fixation properties in Experiment 2. (A) Fixation duration as a function of whether the fixation has an immediate refixation. (B) Distance to the closest dot as a function of whether the fixation is the first fixation on the dot or a refixation.

in Figure 11. The predicted proportion of nonimmediate refixations was greater than the observed values. The observed and predicted data were compared using multiple  $t$  tests with a Bonferroni correction for the number of tests preformed. The differences were significant at all of the data points for numerosities larger than 7. This suggests that eye movements have some memory, but their memory is not perfect. Consistent with Experiment 1, the refixation analysis supports a weak version of the control hypothesis.

**Eccentricity and isolation.** Eccentric dots were less likely to be fixated than central ones [Figure 12A;  $F(2,22) = 281.62$ ,  $MS_e = 0.003$ ,  $p < .001$ ,  $\eta_p^2 = .96$ ]. Isolated dots were less likely to be fixated than dots in dense neighborhoods [Figure 12B;  $F(2,22) = 205.55$ ,  $MS_e = 0.002$ ,  $p < .001$ ,  $\eta_p^2 = .95$ ]. These results replicate Experiment 1, confirm the perception hypothesis, and disconfirm the indexing hypothesis, which says that all dots—even isolated and eccentric ones—should be fixated.

**Consistency of fixations across repetitions.** Each subject viewed each pattern twice. How consistent were the eye movement patterns in the two exposures to a pattern? We measured the similarity of eye movements in three ways. First, we measured the difference in the number of fixations between the first and second exposures to the same pattern. The distribution of the differences is shown in Figure 13A. If the eye movement patterns were similar when subjects viewed the same pattern twice, the mean of the distribution should be very close to zero, and the standard deviation should be very small. However, as shown in Figure 13A, the difference in the number of fixations was zero only in about .20 of the patterns, and there was substantial variability in the distribution. The

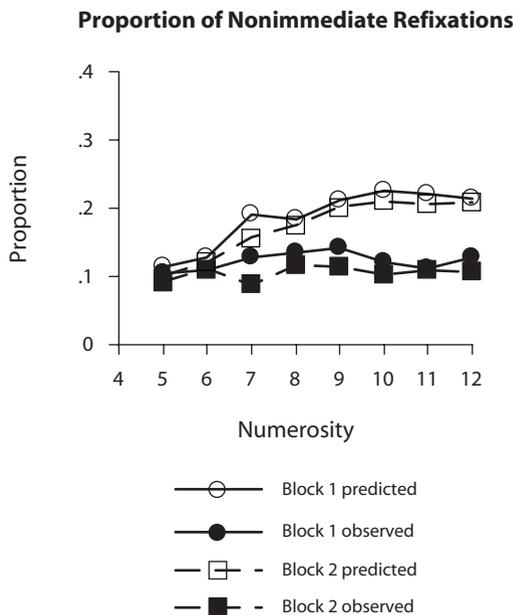
mean of the distribution was 0.3, and the average standard deviation across subjects was 2.8, which were both significantly larger than 0 [ $t(11) = 2.50$ ,  $p < .05$ , and  $t(11) = 8.49$ ,  $p < .001$ , respectively]. This suggests that the number of fixations differed from one exposure to the pattern to the next.

Second, we calculated similarity by dividing the number of dots that were fixated in both exposures by the average number of dots fixated in two exposures. The similarity between the two blocks was .68 and was not different between numerosities ( $F < 1$ ). This suggests that subjects do not always look at the same objects when they view the pattern twice.

Third, we measured the similarity of the scan path of the eye movements. We expressed the scan path on each exposure as a sequence of dots, and we calculated the Levenshtein distance between the sequences for the two exposures (Levenshtein, 1966; Privitera & Stark, 2000; Wagner & Fischer, 1974). The Levenshtein distance between two strings is the minimum number of operations needed to transform one string into the other using insertion, deletion, or substitution of a single element. We calculated the similarity of scan paths as follows:

$$\text{similarity} = 1 - \frac{\text{Levenshtein distance}}{\max(\text{nfix1}, \text{nfix2})},$$

where  $\max(\text{nfix1}, \text{nfix2})$  is the larger number of fixations in the two exposures. If the scan paths are the same for the two exposures, similarity is 1; if the scan paths are totally different, similarity is 0. Mean similarity of scan paths is shown as a function of numerosity in Figure 13B. Similarity was always less than .5 and decreased as numerosity increased [ $F(7,77) = 12.99$ ,  $MS_e = 0.002$ ,  $p < .001$ ,  $\eta_p^2 = .54$ ].



**Figure 11. Proportion of nonimmediate refixations in Experiment 2 as a function of the number of objects presented.**

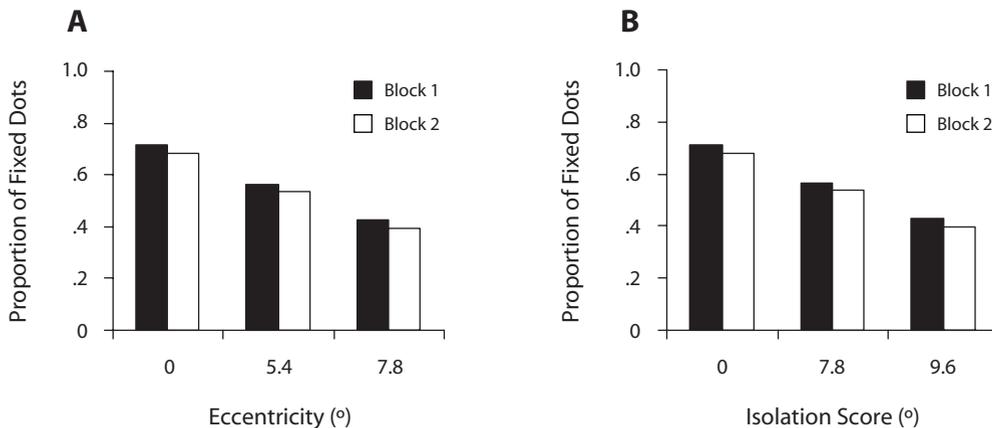
All three measures show marked differences between the fixations on two exposures to the same pattern. This is inconsistent with a strong version of the perception hypothesis, which assumes that eye movement behavior is driven entirely by properties of the pattern, but it is consistent with a weaker version, which assumes that eye movement behavior is driven probabilistically by properties of the pattern, among other influences.

**Consistency of fixations between subjects.** All subjects viewed the same set of dot patterns. We calculated the Levenshtein distance between the scan paths of each pair of subjects viewing the same display. Then we averaged the distances within and across dot patterns. Figure 14 shows the similarity of scan path across subjects as a function of numerosity. The similarity is low (about .3) compared with

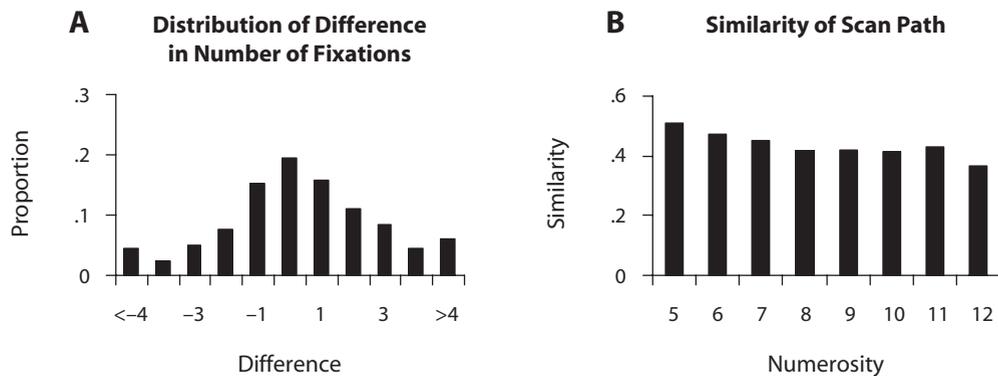
the within-subjects similarity (about .4). The low similarity is inconsistent with a strong version of the perception hypothesis that assumes that eye movements are driven entirely by the properties of the pattern, and so should be the same for all subjects viewing the same pattern. It is consistent with a weaker version of the perception hypothesis that assumes that eye movements are driven probabilistically by properties of the display. The higher similarity within subjects than between subjects suggests that there are individual differences in the susceptibility to display properties.

**Conclusions**

Experiment 2 replicated the main results from the count condition of Experiment 1: Accuracy was high and RT increased linearly with numerosity, with a substantial slope



**Figure 12. (A) Proportions of dots that were fixated as a function of eccentricity in Experiment 2. (B) Proportions of dots that were fixated as a function of the average distance between a dot and all of the other dots.**



**Figure 13.** Consistency of eye movement patterns when a pattern was viewed twice. (A) Distribution of differences in the number of fixations when a pattern was viewed twice in Experiment 2. (B) Similarity of scan paths within subjects in Experiment 2 as a function of the number of objects presented.

that is consistent with serial counting. The number of fixations increased with numerosity, but not with a slope of 1.0, which suggests that some dots were not fixated. The number of fixated dots was smaller than the number of fixations, and about one third of the dots were refixated. Subjects tended to look at dots in dense, central regions and ignore dots in isolated, eccentric regions. Experiment 2 also showed that the number of fixations differed from one exposure to the pattern to the next. Together, these results are inconsistent with the indexing hypothesis, which predicts that every dot will be fixated once and only once. The results are not consistent with a strong version of the control hypothesis, which predicts that no dot will be refixated. The results are consistent with a weak version of the control hypothesis, which predicts that eye movements tend to avoid already fixated dots, but cannot do so perfectly. The results are most consistent with the perception hypothesis, which predicts more fixations to dense, central regions.

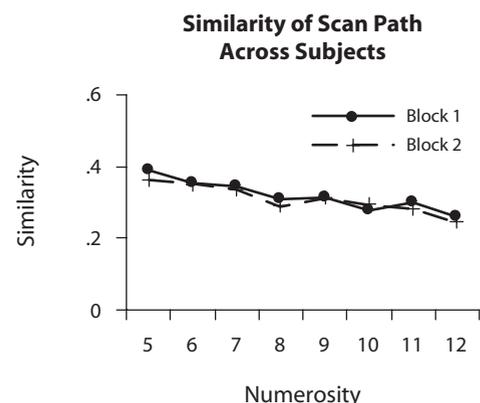
The consistency data ask how strongly the properties of the display determine eye-movement behavior—in essence, how tightly the eyes are coupled to perceptual processes that increase resolution. The within-subjects data show that the same subject looks at the same pattern in different ways on different exposures; the between-subjects data show that different subjects look in different ways. Thus, the coupling between eye movements and perceptual demands does not appear to be very tight.

## GENERAL DISCUSSION

Two experiments addressed the role of eye movements in enumeration. We assessed three hypotheses about the coupling between eye movements and the enumeration process. The indexing hypothesis predicts tight coupling, with one eye movement for every dot in the display. The control hypothesis predicts a looser coupling, in which the eye movement control system prevents refixation of dots that have already been fixated. The perception hypothesis predicts the loosest coupling, with eye movements directed to dense, central regions of the display to resolve perceptual confusion. The data from both experiments were inconsistent with the indexing hypothesis.

Subjects did not fixate every dot in the display, and the number of unique dots that were fixated was about half of the number of dots in the display. The eyes sometimes revisited already fixated dots. Even after excluding immediate refixations, there were still about 10% nonimmediate refixations, which is inconsistent with a strong version of the control hypothesis. However, there was some evidence that subjects avoided looking at already fixated dots. Monte Carlo analysis showed that the observed proportion of nonimmediate refixations was smaller than that predicted by a no-memory model. These results supported a weak version of the control hypothesis. The data were most consistent with the perception hypothesis. In both experiments, subjects tended to look at dots in dense, central regions of the display and tended not to look at dots in sparse, peripheral regions of the display.

In theory, each dot can only be counted once for accurate counting. Then why were there so many nonimmediate refixations during counting? Visual search studies provide some insight into this question. Those studies found that people usually avoid looking at already fixated items during visual search, though there are a small number of refixations (Beck et al., 2006; Peterson et al., 2001). Other studies found that increasing memory load could increase



**Figure 14.** Similarity of scan paths between subjects in Experiment 2 as a function of the number of objects presented.

the proportion of refixation and reduce search efficiency (Han & Kim, 2004; Peterson et al., 2008). Peterson et al. (2008) suggested that attention may move away from one item before processing is completed, so a refixation is necessary to complete processing. We suspect that this is also the case for eye movements in counting.

The conclusion that the coupling between eye movements and enumeration is loose is surprising. One possibility is that our subjects did not use a counting process to enumerate the displays. That is unlikely because accuracy was very high. Alternatives to counting, such as estimation, produce low accuracies in the range of numerosities we employed (Kaufman et al., 1949; Mandler & Shebo, 1982). Another possibility is that the counting process works independently of eye movements, moving covert attention to the locations of the dots that are indexed. This is consistent with our data but inconsistent with previous experiments showing a tight coupling between shifts of covert attention and eye movements (Deubel & Schneider, 1996; Irwin & Andrews, 1996). A third possibility is that eye movements are determined jointly by the enumeration process and the need for perceptual clarity. Subjects may tend to look at the dot they are indexing and, at the same time, tend to look at dots in dense, central regions of the display. These tendencies may compete for control of each eye movement. Sometimes, one tendency may win; other times, the other tendency may win. The competition would loosen the coupling between eye movements and enumeration. A quantitative model would be required to

determine whether it could account for the looseness of the coupling we observed in our experiments.

Another perspective on this “competing tendencies” hypothesis is provided by the fact that eye movements are not necessary for accurate enumeration, but they help. Counting is faster and more accurate when eye movements are allowed than when they are prohibited (Simon & Vaishnavi, 1996; Watson et al., 2007). Eye movements may be helpful when the dot that is currently being indexed is fixated; they may not be harmful when another dot is fixated. Indeed, Carlson, Avraamides, Cary, and Strasberg (2007) showed that counting was facilitated if subjects were allowed to point to the objects they were counting. When pointing was prevented, subjects nodded in time with their counting. When nodding was prevented, counting suffered. Carlson et al. suggested that pointing and nodding were beneficial because they externalized the counting process, allowing the motor control system to participate in the cognitive control required for counting. Each touch and nod marks a step in the indexing process, making it more prominent and making it easier to move on to the next step. Perhaps eye movements play a similar role, marking the indexing steps when fixations coincide with the current index. Indexing can proceed without eye movements, but the occasional coordination may be facilitatory. Further research will be required to evaluate this hypothesis.

When interpreting the difference in intercepts of RT in Experiment 1, we suggested an interpretation in terms of subitizing: Subjects index fewer dots in the count condi-

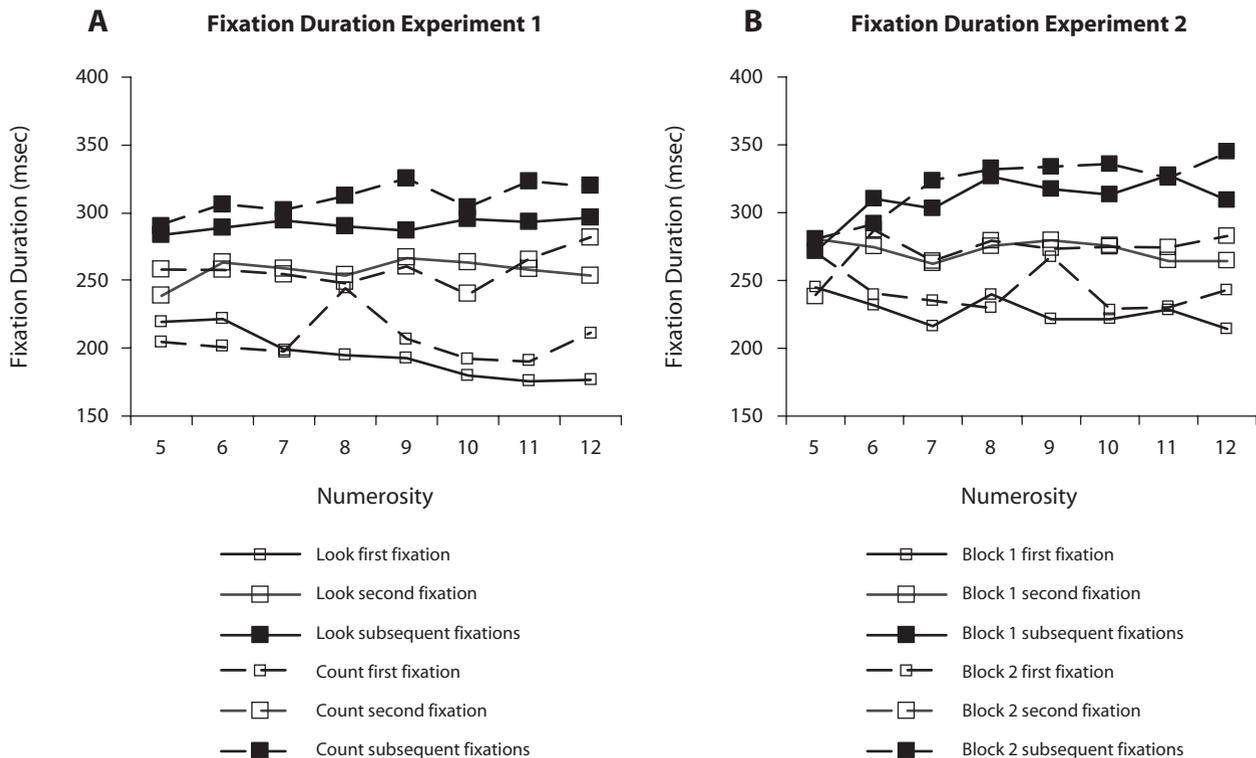


Figure 15. Fixation duration of first, second, and subsequent fixations as a function of numerosity for Experiment 1 (A) and Experiment 2 (B).

tion because they subitize three to five dots and only index the remaining ones. To test this hypothesis, we examined the durations of the first two fixations and the isolation scores of the first two fixated dots. If subjects subitize three to five dots during the first one or two fixations and then index the other dots, then fixation duration for the first one or two fixations should be shorter than for subsequent fixations. The results of both experiments are consistent with this prediction (Figure 15). In Experiment 1, the duration of the first fixation (200 msec) was shorter than the second fixation (257 msec), which was shorter than subsequent fixations (301 msec). This effect was confirmed by a main effect of fixation type [ $F(2,18) = 67.86$ ,  $MS_e = 4,644$ ,  $p < .001$ ,  $\eta_p^2 = .88$ ] in a repeated measures 2 (task)  $\times$  3 (fixation type: first fixation, second fixation, subsequent fixations)  $\times$  6 (numerosity) ANOVA. Planned contrasts showed that the difference between the first fixation and the second was significant [ $F(1,9) = 73.51$ ,  $MS_e = 2,723$ ,  $p < .001$ ,  $\eta_p^2 = .89$ ], as was the difference between the second fixation and subsequent fixations [ $F(1,9) = 32.07$ ,  $MS_e = 3,697$ ,  $p < .001$ ,  $\eta_p^2 = .78$ ]. In Experiment 2, the duration of the first fixation (233 msec) was shorter than that of the second fixation (273 msec), which was shorter than that of the subsequent fixations (325 msec). This effect was confirmed by a main effect of fixation type [ $F(2,22) = 35.57$ ,  $MS_e = 9,379$ ,  $p < .001$ ,  $\eta_p^2 = .75$ ] in a repeated measures 2 (block)  $\times$  3 (fixation type: first fixation, second fixation, subsequent fixations)  $\times$  6 (numerosity) ANOVA. These results are consistent with the suggestion that subjects subitize during the first one or two fixations and then count the other dots.

We also looked at the isolation score for the dots that were fixated first, second, and subsequently. Subjects may subitize isolated dots and then count the remaining dots. If they do, then the isolation score for the dots that were fixated first and second should be larger than the mean isolation score for the display. The results disconfirmed this prediction. The isolation scores of the dots that were fixated first and second were smaller than the mean isolation score. The mean isolation score of the display was 8.8°. In the count condition of Experiment 1, the isolation scores for the first and second fixated dots were 8.5° and 8.2°, respectively. Both values were smaller than the mean isolation score for the display (8.8°) [ $t(9) = 1.86$ ,  $p < .1$ ;  $t(9) = 7.05$ ,  $p < .001$ ]. In Experiment 2, the isolation scores of the first and second fixated dots were 7.8° and 7.9°, respectively, for Block 1 and were 7.9° and 7.9° for Block 2. These scores were significantly smaller than the mean isolation score of the display [8.8°;  $ts(11) = 11.49$ , 11.26, 7.89, and 12.08, respectively;  $ps < .001$ ]. These results are inconsistent with the hypothesis that subjects first fixate isolated dots and subitize them.

## Conclusions

We examined three hypotheses regarding eye movements in counting. The indexing hypothesis, which assumes a strong coupling between cognition and eye movements, was eliminated with strong evidence. The strong version of the control hypothesis was also eliminated, but a weak

version was supported. Subjects tended to avoid looking at already fixated dots, but they could not do so perfectly. The perceptual hypothesis was supported by strong evidence. These results suggest a loose coupling between eye movements and cognition in counting. It is likely that a mixed strategy of enumeration, control, and perception could account for the eye movements we observed.

## AUTHOR NOTE

This research was supported by Grant BCS 0646588 from the National Science Foundation and Grant R01-MH073879-01 from the National Institute of Mental Health. Correspondence concerning this article should be addressed to G. D. Logan, Department of Psychology, Vanderbilt University, Nashville, TN 37203 (e-mail: gordon.logan@vanderbilt.edu).

## REFERENCES

- ATKINSON, J., CAMPBELL, F. W., & FRANCIS, M. R. (1976). The magic number 4±0: A new look at visual numerosity judgements. *Perception*, *5*, 327-334.
- BECK, M. R., PETERSON, M. S., & VOMELA, M. (2006). Memory for where, but not what, is used during visual search. *Journal of Experimental Psychology: Human Perception & Performance*, *32*, 235-250. doi:10.1037/0096-1523.32.2.235
- BECKWITH, M., & RESTLE, F. (1966). Process of enumeration. *Psychological Review*, *73*, 437-444.
- CARLSON, R. A., AVRAAMIDES, M. N., CARY, M., & STRASBERG, S. (2007). What do the hands externalize in simple arithmetic. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *33*, 747-756. doi:10.1037/0278-7393.33.4.747
- CHI, M. T. H., & KLAHR, D. (1975). Span and rate of apprehension in children and adults. *Journal of Experimental Child Psychology*, *19*, 434-439.
- COMPTON, B. J., & LOGAN, G. D. (1999). Judgments of perceptual groups: Reliability and sensitivity to stimulus transformation. *Perception & Psychophysics*, *61*, 1320-1335.
- DEUBEL, H., & SCHNEIDER, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research*, *36*, 1827-1837.
- GEARY, D. C. (2000). From infancy to adulthood: The development of numerical abilities. *European Child & Adolescent Psychiatry*, *9*, II/11-II/16.
- GIPPENREITER, Y. B., ROMANOV, V. Y., & SMIRNOV, S. D. (1969). Eye and hand movements in the process of counting the elements of a test object. *Psikholgocjesloe Issledovaniya*, *1*, 51-56.
- HAN, S. H., & KIM, M. S. (2004). Visual search does not remain efficient when executive working memory is working. *Psychological Science*, *15*, 623-628.
- HENSON, D. B. (1978). Corrective saccades: Effects of altering visual feedback. *Vision Research*, *18*, 63-67.
- IRWIN, D. E., & ANDREWS, R. V. (1996). Integration and accumulation of information across saccadic eye movements. In T. Inui & J. I. McClelland (Eds.), *Attention and performance XVI: Information integration in perception and communication* (pp. 125-155). Cambridge, MA: MIT Press, Bradford Books.
- JENSEN, E. M., REESE, E. P., & REESE, T. W. (1950). The subitizing and counting of visually presented fields of dots. *Journal of Psychology*, *30*, 363-392.
- KAUFMAN, E. L., LORD, M. W., REESE, T. W., & VOLKMAN, J. (1949). The discrimination of visual number. *American Journal of Psychology*, *62*, 498-525.
- KLAHR, D. (1973). A production system for counting, subitizing, and adding. In W. G. Chase (Ed.), *Visual information processing* (pp. 527-546). New York: Academic Press.
- KLEIN, R. M. (2000). Inhibition of return. *Trends in Cognitive Sciences*, *4*, 138-147.
- KOWLER, E., & STEINMAN, R. M. (1977). Role of small saccades in counting. *Vision Research*, *17*, 141-146.
- KOWLER, E., & STEINMAN, R. M. (1979). Miniature saccades: Eye movements that do not count. *Vision Research*, *19*, 105-108.

- LASSALINE, M. E., & LOGAN, G. D. (1993). Memory-based automaticity in the discrimination of visual numerosity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 561-581.
- Левенштейн, В. И. (1965). Двоичные коды с исправлением выпадений, вставок и замещений символов. *Доклады Академии Наук СССР*, **163**, 845-848. Appeared in English as Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady*, **10**, 707-710.
- LOGAN, G. D., & ZBRODOFF, N. J. (2003). Subitizing and similarity: Toward a pattern-matching theory of enumeration. *Psychonomic Bulletin & Review*, **10**, 676-682.
- MANDLER, G., & SHEBO, B. J. (1982). Subitizing: An analysis of its component processes. *Journal of Experimental Psychology: General*, **111**, 1-22.
- NORO, K. (1980). Determination of counting time in visual inspection. *Human Factors*, **22**, 43-55.
- PALMERI, T. J. (1997). Exemplar similarity and the development of automaticity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **23**, 324-354.
- PETERSON, M. S., BECK, M. R., & WONG, J. H. (2008). Were you paying attention to where you looked? The role of executive working memory in visual search. *Psychonomic Bulletin & Review*, **15**, 372-377. doi:10.1080/03610730802070068
- PETERSON, M. S., KRAMER, A. F., WANG, R. X. F., IRWIN, D. E., & MCCARLEY, J. S. (2001). Visual search has memory. *Psychological Science*, **12**, 287-292.
- POSNER, M. I., & COHEN, Y. (1984). Components of visual orienting. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 531-556). Hillsdale, NJ: Erlbaum.
- PRIVITERA, C. M., & STARK, L. W. (2000). Algorithms for defining visual regions-of-interest: Comparison with eye fixations. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, **22**, 970-982.
- RAYNER, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, **124**, 372-422.
- RYAN, J. D., HANNULA, D. E., & COHEN, N. J. (2007). The obligatory effects of memory on eye movements. *Memory*, **15**, 508-525. doi:10.1080/09658210701391022
- SALTZMAN, I. J., & GARNER, W. R. (1948). Reaction time as a measure of span of attention. *Journal of Psychology*, **25**, 227-241.
- SIMON, T. J., & VAISHNAVI, S. (1996). Subitizing and counting depend on different attentional mechanisms: Evidence from visual enumeration in afterimages. *Perception & Psychophysics*, **58**, 915-926.
- TRICK, L. M., ENNS, L. T., & BRODEUR, D. A. (1996). Life span changes in visual enumeration: The number discrimination task. *Developmental Psychology*, **32**, 925-932.
- TRICK, L. M., & PYLYSHYN, Z. W. (1993). What enumeration studies can show us about spatial attention: Evidence for limited capacity preattentive processing. *Journal of Experimental Psychology: Human Perception & Performance*, **19**, 331-351.
- 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.
- VAN OEFFELEN, M. P., & VOS, P. G. (1984). Enumeration of dots: An eye movement analysis. *Memory & Cognition*, **12**, 607-612.
- WAGNER, R. A., & FISCHER, M. J. (1974). The string-to-string correction problem. *Journal of the ACM*, **21**, 168-173.
- WATSON, D. G., MAYLOR, E. A., & BRUCE, L. A. M. (2007). The role of eye movements in subitizing and counting. *Journal of Experimental Psychology: Human Perception & Performance*, **33**, 1389-1399. doi:10.1037/0096-1523.33.6.1389
- WOLTERS, G., VAN KEMPEN, H., & WIJLHUIZEN, G. J. (1987). Quantification of small numbers of dots: Subitizing or pattern recognition? *American Journal of Psychology*, **100**, 225-237.
- ZELINSKY, G. J., RAO, R. P. N., HAYHOE, M. M., & BALLARD, D. H. (1997). Eye movements reveal the spatio-temporal dynamics of visual search. *Psychological Science*, **8**, 448-453.

(Manuscript received February 7, 2009;  
revision accepted for publication October 10, 2009.)