Age differences in enumerating things that move: Implications for the development of multiple-object tracking

LANA M. TRICK, DIANA AUDET, and LYNN DALES
University of Guelph, Guelph, Ontario, Canada

The attentional theory of spatial enumeration (Trick & Pylyshyn, 1994) predicts that subitizing, the rapid process (40–120 msec/item) used to enumerate 1–4 items, employs the same mechanism that permits individuals to track 4–5 moving items simultaneously, whereas enumerating more items requires moving attentional focus from area to area in the display. To test this theory, enumeration of static and moving items was investigated in 8-, 10-, 12-, and 20-year-old participants using a number discrimination task. As was predicted, random independent item motion did not substantially impede enumeration of 1–4 items regardless of age. However, even movement within a 1.14° square area slowed enumeration of 6–9 items, although on average the interference decreased with age from 788 msec for the 8-year-olds to 136 msec for the 20-year-olds. The relevance of this finding for theories of enumeration, multiple-object tracking, visual working memory, and object-based attention is discussed.

In this study, we investigated how the ability to enumerate moving things develops with age. This topic has ramifications for theories of enumeration and the development of enumeration. Although the research on enumeration has a long history (see, e.g., Jevons, 1871), to our knowledge there has never been an investigation of the enumeration of moving items. Because enumerating moving things requires keeping track of them, this research is also relevant to studies of multiple-object tracking (e.g., Pylyshyn & Storm, 1988).

Consider a visuospatial enumeration task: Viewers are required to enumerate dots in a visual display in which all of the dots are presented simultaneously, each dot occupying a different location in space. Response time (RT) and accuracy are measured. There are qualitative differences between enumerating small and large numbers of dots. When the numbers are small (fewer than 4 or 5 items), most adults report that enumeration is effortless and immediate. Within this range, there is only a modest increase in RT for each additional item (40–120 msec/item for most adults) and errors are rare. In contrast, when there are more than 5 items, most adults report that enumeration is effortful. Enumeration in this range is much more time-consuming (200–350 msec/item), and errors

become more frequent. When faced with 5 or more items, most adults report counting by groups (van Oeffelen & Vos, 1982)—that is, focusing on one part of the display, enumerating the items within that area, adding the result into a running total, focusing on the next part of the display, enumerating the items in that area, and adding that result into the running total until all items are enumerated. The differences between enumerating small and large numbers of items have been interpreted as evidence that there are two different enumeration processes: *subitizing*, a fast and accurate process specialized for small numbers of items, and *counting*, a slow and effortful process that can be used with larger numbers of items (see, e.g., Klahr & Wallace, 1976). The discontinuity in the enumeration slopes that distinguishes subitizing from counting has been observed in age groups ranging from children to the oldest adults (Chi & Klahr, 1975; Trick, Enns, & Brodeur, 1996), and some suggest that subitizing and counting involve different neural mechanisms (Sathian et al., 1999).

Several theories explain why small numbers of items are enumerated differently than large numbers in visuospatial displays. However, only one incorporates subitizing and counting into a general theory of vision and predicts specific cases in which observers cannot use their fast, accurate subitizing process to enumerate small numbers of items laid out in different locations in space (Trick & Pylyshyn, 1994). According to this theory, subitizing and counting differ because they have different attentional requirements. Subitizing makes use of a mechanism that assigns mental reference tokens (*FINgers of INSTantiation*, or *FINSTs*) to items in the display so that

This research was funded by a grant from the Natural Sciences and Engineering Research Council of Canada and was based on D.A.'s honors thesis and L.D.'s Advanced Special Topics research project. We thank Tahlia Perl, who tested the university students. Correspondence concerning this article should be addressed to L. Trick, Department of Psychology, University of Guelph, Guelph, ON, N1G 2W1 Canada (e-mail: ltrick@uoguelph.ca).

the attentional focus can later access the location of items even when they move, whereas counting requires the deployment of the attentional focus.

Full development of this idea will require working through a train of reasoning that originates from FINST theory (Pylyshyn, 2001). FINST theory was originally developed to explain how mechanisms of early vision could be used to maintain a stable representation of the world despite the changing patterns of retinal stimulation produced by eye movements and item motion (Pylyshyn, Elcock, Marmor, & Sander, 1978). A key feature of this theory is the idea of *reference tokens*. A reference token is basically a name tag used to refer to something, and FINST reference tokens are name tags that are assigned to specific items in a visual display. FINSTs are necessary because it is counterproductive to refer to an item by its location (e.g., “the item in the top left corner”) given that items change position when they move, and it is equally counterproductive to refer to an item by its properties (e.g., “the small black dot”) given that there are often many items with the same properties within one display (Trick & Pylyshyn, 1994). Moreover, item properties may change from moment to moment and may not even be fully known if the item has not yet been visited by the attentional focus. In the same way that children are named so that they do not have to be referred to by their positions or properties, each item within a small set is mentally named in the visual scene. FINST reference tokens act as pointer variables similar to those used in computer languages such as C. These variables carry information about where items are, although they do not refer to items by their locations or indicate what their properties are.

These reference tokens are necessary for both visually guided action and derivation of the spatial relations necessary for object recognition. (Spatial relations are what distinguish a random grouping of contours from the schematic drawing of a face, for example.) Thus, in order to perform actions such as first directing the attentional focus, then the eye, and then the hand, toward a specific object in a visual scene, there must be a way to refer to that specific object. Similarly, deriving spatial relations, such as determining whether two items are connected by a specific line, are inside or outside a specific contour, or are left or right of one another, requires a way to refer to specific items as separate entities. It is commonly thought that the attentional focus is used to derive the spatial relations between elements (see, e.g., Ullman, 1984), but for the attentional focus to accomplish this task there would need to be a way to refer to the constituent elements in the relation. Therefore, overall, FINSTs serve to preselect a small number of items as future targets for the attentional focus. The FINST mechanism can handle only 4–5 items at once because the number of reference tokens is limited. Although there may be individual differences in the exact number of reference tokens (Trick & Pylyshyn, 1994), in general the number is small because these reference tokens are used to select targets for the attentional focus. It makes little

sense to select everything at once. At the same time, if FINSTs are to be used to derive spatial relations, then there must be a way to select more than one item at once, and that would demand more than one reference token.

The best evidence for FINST theory is the multiple-object tracking task. Participants are shown a number of identical items, and a subset of the items is highlighted to indicate which items are targets. Then, targets and distractors become identical once again and all are sent into random, independent motion. Some time later, the participants are required to indicate which items are targets. Pylyshyn and Storm (1988; Sears & Pylyshyn, 2000) found that young adults could simultaneously track up to 4–5 items with accuracy under these conditions.

The attentional account of spatial enumeration is based on the assumption that subitizing exploits the same mechanism that permits multiple-object tracking: the ability to assign reference tokens to a small number of items. The difference is that, for subitizing, the reference tokens are checked off in order to access number names, whereas in multiple-object tracking they are used to distinguish targets from distractors. Thus, subitizing 3 items involves assigning 3 FINST reference tokens and then, one at a time, matching reference tokens to number names from long-term memory in the order of the number names (e.g., *one, two, three*). In the subitizing range, most of the increase in RT as number increases (i.e., the slope) is produced by the process that matches reference tokens to number names to derive the numeric response that corresponds to the cardinality of the display. Each time a number name is matched, it adds a certain amount of time to the response latency. That is why threshold studies of enumeration, which measure the minimum requisite stimulus exposure to attain a certain level of accuracy and do not include response choice as part of the latency, show a comparatively small increase in the requisite duration with numbers in the 1- to 4-item range (Folk, Egeth, & Kwak, 1988; Liss & Reeves, 1983; Oyama, Kikuchi, & Ichihara, 1981; Sagi & Julesz, 1984). The reason that number names must be checked off in sequence is because embedded within every 3-item display are 2-item and 1-item displays. If the overall numeric response is to be correct, then each and every item must be considered as an individual and paired with one and only one number name.

According to the attentional theory of enumeration, counting requires a more complex process than subitizing, because it occurs when items outnumber mental reference tokens and, thus, reference tokens have to be assigned and later reassigned in different parts of the display. This necessitates additional operations. First, items are grouped into clusters (van Oeffelen & Vos, 1982). The attentional focus is moved toward a given cluster and items within the cluster are subitized (i.e., FINST reference tokens are assigned; a number name is accessed). The resultant number is added into a running total held in working memory. Once enumerated, items are mentally “crossed off” as having been already attended.

Then, the attentional focus is moved to the position of a cluster of items that has not yet been “crossed off,” and the cycle repeats and continues repeating until all items are “crossed off.” The final value of the working memory total is then reported (see the group-and-add enumeration of Klahr & Wallace, 1976). The counting process is slow because it involves mentally adding items into a subtotal, storing subtotals in working memory, moving the attentional focus from location to location in the image, and “crossing off” items as counted so they are not reenumerated. Errors occur because people make mistakes in addition, lose track of the subtotal, or miss or recount items because they forget which items were attended (Tuholski, Engle, & Baylis, 2001). The attentional theory of enumeration requires FINSTs to explain subitizing, but goes beyond FINST theory insofar as it also incorporates spatial attention and both working and long-term memory, which are necessary to explain performance beyond the subitizing range.

This study has two goals. One is to provide an initial test of a critical assumption in the attentional theory of spatial enumeration: the idea that subitizing makes use of the same mechanism that permits individuals to track 4–5 moving items at once. The second is to examine developmental changes in the ability to deal with moving items. Children have exaggerated RT slopes in search tasks that require scanning attention through the display (see, e.g., Kaye & Ruskin, 1990). For this reason, it is especially interesting to look at the effects of motion because object motion challenges the mechanisms involved in systematically scanning multielement displays. When items move from location to location, it is possible that individuals might accidentally miss or reattend an item because an item never before attended moves into a previously attended location, or a previously attended item moves into a location not previously attended. FINSTs could be used to track 4 or 5 moving items and thus prevent this, but a different scheme would have to be used when there are more moving items.

This study constitutes a first step in bridging between the enumeration and the multiple-object tracking tasks. However, it is important to note that motion was carefully constrained in this study so that (1) items maintained their integrity in apparent motion, so they did not split, merge, or disappear; and (2) the motion was restricted to four frames so that all of the items would have traveled the maximum distance they could travel within 396 msec, before even the fastest participants could enumerate 1 item. In this way, we avoided confounding the distance that items traveled with the amount of time required to enumerate a particular number of items for participants of a particular age. Each item shifted randomly and independently of the others within a 1.14° area through four frames of motion. Individual items did not necessarily maintain a consistent trajectory through the four frames, but the sequence of frames repeated.

Although a variety of different paradigms have been used to study enumeration in the past (see Trick &

Pylyshyn, 1994, for a review), we chose to use a number discrimination task in this study. A key feature of this task is the reduction in response uncertainty: Whereas classic enumeration tasks can have 20 alternatives or more, number discrimination tasks have 2 possible responses: n and $n + 1$. There is a history of number discrimination studies in the enumeration literature (see, e.g., Folk et al., 1988; Trick et al., 1996). Generally, latencies are shorter in number discrimination tasks than in classic enumeration tasks, but subitizing and counting emerge as usual insofar as the RT slope is small within the 1- to 4-item range and large when there are 6 items or more. The general shape of the enumeration function, which plots RT as a function of number, remains the same (Trick, 2003). We used number discrimination in this study because we wanted to compare the results with two-alternative attentional tasks and were not interested in developmental changes in the effects of response uncertainty.

We predicted that the type of motion used in this study should not interfere substantially with number discrimination in the subitizing range (i.e., 1–4 items) because we propose that subitizing and multiple-object tracking use the same mechanism. Moreover, we predicted no developmental changes in school-aged children in this range. We believe this because the ability to discriminate between numerosities within the subitizing range emerges in infancy (see, e.g., Starkey & Cooper, 1995), and we propose that the ability to track small numbers of moving items must also develop before school age because it relies on the same mechanism.

In contrast, we predicted that motion would influence number discrimination latencies in the counting range (i.e., 6–9 items). We propose that counting demands some of the same abilities as attentional search, such as moving the attentional focus from area to area in the display, inhibiting or “crossing off” items once attended so they are not revisited. Children have exaggerated difficulties with attentional search, as is shown by their higher RT slopes in finding a target when discriminating targets from distractors requires attention (see, e.g., Kaye & Ruskin, 1990). There are a number of possible reasons this might occur, however; one contender is difficulty in inhibiting once attended items so they are not reattended. We predicted that motion would exaggerate this difficulty, because it would make it easier for children to accidentally reattend items since item positions change from moment to moment. This difficulty should produce not only exaggerated RTs, but also exaggerated RT slopes in the 6- to 9-item range because the task of keeping straight the positions of moving items would become more and more difficult as the number of items increases.

METHOD

Participants

One hundred thirty-eight individuals participated in the study. Of these, 105 were from a local elementary school: a group of 8-year-olds ($M = 7$ years, 6.9 mos., $SD = 2.9$ mos., Grade 2, $n = 30$, 16

female), 10-year-olds ($M = 9$ years, 9 mos., $SD = 4.2$ mos., Grade 4, $n = 43$, 17 female), and 12-year-olds ($M = 11$ years, 7.3 mos., $SD = 3.1$ mos., Grade 6, $n = 32$, 19 female). The remaining 33 were volunteers from the university psychology subject pool ($M = 20$ years, 5.5 mos., $SD = 16.9$ mos., 27 female). Only data from individuals with no perceptual, attentional, or learning disabilities were included. Information about disabilities was obtained either from the participants (in the case of the university students) or from the parents of the participants (in the case of the children) via questionnaires administered prior to the study. The participants within each age group were randomly assigned to either the static- or the moving-item condition. For each age group, the numbers of participants in each condition were as follows: 8-year-olds, 14 in the static condition and 16 in the moving condition; 10-year-olds, 22 in the static condition and 21 in the moving condition; 12-year-olds, 14 in the static condition and 18 in the moving condition; 20-year-olds, 17 in the static condition and 16 in the moving condition.

Apparatus and Stimuli

The experiment was conducted on a Macintosh G4 Powerbook. The participants were tested individually in a small, quiet room away from others and were seated 50 cm away from the computer display.

The task was to enumerate black dots, each having a diameter of 0.34° visual angle. The dots were presented within a 6×5 notional grid that occupied a $7.5^\circ \times 6.3^\circ$ area. Items on the grid were no closer than 1.51° . In the static-item condition, the position of the dots remained fixed until a response was made, whereas in the moving-item condition the dots were presented for 66 msec, disappeared for 33 msec, and reappeared somewhere within a 1.14° square area from where they had started. Dots changed position randomly and independently of one another through four frames of apparent motion. The sequence required 396 msec to complete, at which time it repeated and continued repeating until a response was made. Even with movement, the dots could come no closer to each other than 0.37° .

Procedure

The participants were required to discriminate between n and $n + 1$ in four different discriminations: 1 versus 2, 3 versus 4, 6 versus 7, and 8 versus 9. There was no number discrimination that involved the number 5 because we wanted to ensure that the elements in a given discrimination were either both in the subitizing range or both in the counting range. The response key for the low number in each discrimination was always on the left side (the "A" key), and the response key for the high number was on the right (the "L" key). Discrimination was blocked, and the order of the discrimination blocks was counterbalanced across subjects. Before each discrimination block, a female research assistant gave the participants comprehensive instruction and glued the relevant numbers to the computer keys (e.g., in the 3 vs. 4 discrimination block, a tag with the number 3 on it was glued onto the left key and a tag with the number 4 on it was glued onto the right key).

The participants began each of the four discrimination blocks with 2 demonstration trials and 18 trials of practice on the relevant discrimination. There were 36 trials in each of the four discrimination blocks, for a total of 144 experimental trials. Each trial began with a 747-msec fixation cross, and response timing began at the onset of the enumeration display.

RESULTS

There were three dependent measures: RT, RT slope (i.e., the increase in latency as a function of number), and error rates. RTs were trimmed in such a way that for each participant and for each number, only latencies that fell within two standard deviations of the individual's mean

at that number were included. Fewer than 4.9% of the trials were lost due to the trimming process.

Mean correct RTs and error rates were measured as a function of three independent variables: age group (8, 10, 12, and 20 years), condition (static item and moving item), and number of items (1, 2, 3, 4, 6, 7, 8, and 9). Because enumeration is thought to involve different processes in the 1- to 4- and 6- to 9-item ranges (subitizing range and counting range, respectively) and because these processes are related to specific predictions in the theory as they relate to motion, planned comparisons were performed and RTs and error rates in the 1- to 4- and 6- to 9-item ranges were analyzed separately.¹ The mean correct RTs are presented in Figure 1 and error rates in Figure 2.

Subitizing Range (1–4 Items)

Number and age had significant effects in the 1- to 4-item range [number: $F(3,390) = 68.9$, $MS_e = 11,817$, $p = .001$; age: $F(3,130) = 87.6$, $MS_e = 91,739$, $p < .001$]. Specifically, RTs were faster for small than for large numbers of items, and the older participants responded faster than the younger participants did. As was predicted, condition did not have a significant effect on latencies and did not interact with other factors in this range ($p > .1$ for all). Thus, although there were slight differences between conditions, the differences were not statistically reliable or systematically related with age.

Error rates were extremely low in the 1- to 4-item range (2.8% on average). There was little evidence of a speed–accuracy tradeoff because there were few systematic trends in the data. There was a marginal interaction between age and number of items in the 1- to 4-item range [$F(9,390) = 1.89$, $MS_e = 16.4$, $p = .051$] as a result of slightly exaggerated error rates for one item for the youngest participants, but otherwise no significant effects emerged ($p > .1$).

Counting Range (6–9 Items)

Motion slowed enumeration in the 6- to 9-item range [$F(1,130) = 8.0$, $MS_e = 2,665,840$, $p < .01$]. Number and age both had significant effects [number, $F(3,390) = 231.7$, $MS_e = 219,513$, $p < .001$; age, $F(3,130) = 70.7$, $MS_e = 2,665,840$, $p < .001$], and there was also a significant interaction between number and age in this range [$F(9,390) = 6.8$, $MS_e = 219,513$, $p < .001$]. Analyses also revealed a three-way interaction between age, condition, and number [$F(9,390) = 2.0$, $MS_e = 219,513$, $p < .05$]. For the 8- and 10-year-old children, movement slowed enumeration substantially, by an average of 788 and 486 msec, respectively, and this effect was consistent across the 6- to 9-item range. Thus, when data from these two age groups were analyzed separately, there was no evidence of a condition $\times$ number interaction in either group ($p > .1$ for both groups). Condition had an effect, although it was marginal [8-year-old group, $F(1,28) = 3.06$, $MS_e = 6,047,258$, $p = .091$; 10-year-old group, $F(1,41) = 3.57$, $MS_e = 2,843,619$, $p = .066$]. In contrast,

Figure 1. Mean response times for correct trials of enumerating static and moving items in 8-, 10-, 12-, and 20-year-old participants (standard error bars included).

for the 12- and 20-year-old groups, movement delayed the enumeration by an average of only 180 and 136 msec, respectively, in the 6- to 9-item range. In fact, overall, condition had no effect in the 12- and 20-year-old groups ($F < 1$ for both), although there were significant condition $\times$ number interactions for both age groups [10-year-old group, $F(3,90) = 2.83$, $MS_e = 135,482$, $p < .05$; 20-year-old group, $F(3,93) = 5.98$, $MS_e = 67,018$, $p < .005$]. Motion slowed the 6- versus 7-item discrimination more than the 8- versus 9-item discrimination in the 12-year-old group, whereas the reverse was true for the 20-year-old group.

Overall, the average error rate in the 6- to 9-item range was 7.8%. Number had a significant effect on the error rate, with higher error rates associated with higher numbers of items [$F(3,390) = 5.88$, $MS_e = 56$, $p < .001$]. There was also a significant three-way interaction between age, condition, and number [$F(9,390) = 1.95$, $MS_e = 56$, $p < .05$]. The effect did not seem to be systematic. When the data sets for the four age groups were analyzed separately, no significant effects emerged in the 8- and 12-year-old groups ($p > .1$). For the 10-year-old group, number had a significant effect on error rates [$F(3,123) = 2.8$, $MS_e = 58$, $p < .05$], with the highest error rate at 8 items. For the 20-year-old group, there was a main effect of number [$F(3,93) = 4.53$, $MS_e = 37$, $p < .01$] and an interaction between number and condition [$F(3,93) = 4.01$, $MS_e = 37$, $p < .05$], the result of an exaggerated

error rate at 9 items in the moving item condition. In general, there is little evidence of a speed-accuracy tradeoff, because conditions with the highest error rates also had the highest response latencies.

RT Slopes

In enumeration studies, RT slopes are typically compared in the 1- to 4-item and 6- to 9-item ranges in order to determine whether subitizing and counting occurred. If subitizing and counting have taken place as usual, then the slope in the 1- to 4-item range should be significantly smaller than the slope in the 6- to 9-item range. RT slopes were calculated separately for the subitizing and counting ranges using a technique employed in other number discrimination studies (see, e.g., Trick et al., 1996). Specifically, to calculate slope in the 1- to 4-item range, the mean RT for the 1 versus 2 discrimination was subtracted from the mean RT for the 3 versus 4 discrimination and divided by two. A parallel procedure was used in the 6- to 9-item range. RT slopes were then measured as a function of three independent variables: age group (8-, 10-, 12-, or 20-year-olds), condition (static item or moving item), and number range (1-4 or 6-9). The slopes are shown in Table 1. There was clear evidence of subitizing and counting. The typical disparity between the 1- to 4-item and the 6- to 9-item number range slopes emerged [$F(1,130) = 217.7$, $MS_e = 61,807$, $p < .001$]. There was also the usual age-related decrease

Figure 2. Mean percent errors for enumerating static and moving items in 8-, 10-, 12-, and 20-year-old participants (standard error bars included).

in slopes [$F(3,130) = 11.2, MS_e = 49,891, p < .001$], with the larger drop in the 6- to 9-item range [age $\times$ number range: $F(3,130) = 4.6, MS_e = 61,807, p < .01$]. Planned comparisons were performed, and subitizing (1- to 4-item) and counting (6- to 9-item) range slopes were analyzed separately. In neither case was there a significant main effect or interaction involving the movement condition ($p > .1$). Contrary to our prediction, there was no clear evidence that motion affects enumeration slopes in a systematic manner in the 6- to 9-item range. Perhaps

motion would have had a greater effect if item movements were not so constrained.

DISCUSSION

These results show that all the participants were capable of enumerating moving items—at least given the limited type of motion in this study. Apparent motion did not preclude subitizing insofar as the usual slope discontinuity between 1–4 and 6–9 occurred in all conditions. As would be predicted if subitizing made use of a mechanism normally employed for tracking moving items, motion did not slow enumeration to a great extent within the subitizing range (Trick & Pylyshyn, 1994). Moreover, as would be expected if the tracking mechanisms were fully developed by 8 years of age, the delay associated with movement was not systematically related to age.

Item movement substantially slowed enumeration in the counting range (i.e., 6–9 items), and that delay decreased systematically with age, at least between 8 and 12 years old. Given that a prior study (Trick, Dales, & Audet, 2002) found no differences in enumeration latencies for static and flickering items that flashed on and off at the same rate as the moving stimuli in this study, it seems probable that it is the apparent motion per se and not simply the onsets and offsets required by apparent motion that produced the effect in the 6- to 9-item range.

Table 1 Enumeration Reaction Time Slopes (in Milliseconds) as a Function of Age and Condition for Each Number Range					
Age Group	Condition	1- to 4-Item Range		6- to 9-Item Range	
		M	SE	M	SE
8 years	Static	97	19	679	87
	Moving	140	18	640	82
	Average	118	13	659	60
10 years	Static	56	15	506	70
	Moving	56	16	541	71
	Average	56	11	523	50
12 years	Static	48	19	662	87
	Moving	37	18	477	77
	Average	42	13	570	58
20 years	Static	36	17	213	79
	Moving	34	18	364	82
	Average	35	13	289	56

The results of the present study might be expected if enumerating moving items required a time-consuming process applied to the display as a whole to compensate for item motion when there were 6–9 items, and this process required considerably more time for the youngest groups of participants.

There are a number of possible sources of this age-related effect of motion in the 6- to 9-item range (i.e., the counting range). The attentional theory of enumeration assumes that enumerating 6–9 items requires attentional focus, and it is noteworthy that number discrimination slopes in the 6- to 9-item range seem to follow the same pattern of lifespan change as do RT slopes in attentional search. In previous studies, when performance was compared over five age groups (6, 8, 10, 22, and 72 years of age), in both attentional search slopes (Trick & Enns, 1998) and counting slopes for 6–9 items (Trick et al., 1996) there was a decrease up to 22 years of age and then a significant increase in the oldest age group. In contrast, in the 1- to 4-item range (i.e., the subitizing range), number discrimination slopes dropped with age to 22 years and then stabilized. If attention were shifted from location to location in the display, then item motion would make enumeration more difficult, because items that were already enumerated might be reattended if they shifted into a location that had not been checked.

At present, two mechanisms have been proposed to explain how items are not reattended in attentional search: marking (Watson & Humphreys, 1997) and inhibition of return (e.g., Klein & MacInnes, 1999; Snyder & Kingstone, 2000). Both proposals assume that previously attended elements are inhibited; both assume that the inhibition can be either object based, with inhibition staying with the object as it moves, or location based, with inhibition remaining fixed at a location (see, e.g., Tipper, Jordan, & Weaver, 1999; Tipper, Weaver, Jerreat, & Burak, 1994; Watson, 2001). However, at this point the evidence is incomplete and neither mechanism seems to have all that it would take to be useful in enumeration. Marking is thought to involve top-down attentional inhibition that competes for resources with other attention-demanding tasks (Watson & Humphreys, 2000), and although this mechanism can deal with many items at once (e.g., 24; Jiang, Chun, & Marks, 2002a), to date marking has been observed only using a paradigm in which one group of items precedes another (Jiang, Chun, & Marks, 2002b). In contrast, inhibition of return would be more useful in the general case in which all items are present at once, but at this point it has been shown only with smaller numbers of items in search tasks and seems to dissipate over time (Snyder & Kingstone, 2000). This would make inhibition of return less and less useful as number of items increases. Although counting is less accurate than subitizing, it is still reasonably accurate, even with large numbers of items.

Nonetheless, it seems possible that an object-based mechanism might be used in this study. Although each item's motion was limited to a 1.04° square area, moving

items could approach each other within 0.37° and thus invade each other's inhibitory zone if a simple location-based mechanism were used. Moreover, this study shows an age-related effect of motion, and there is evidence that suggests it is the object-based form of inhibition that is more subject to age-related change for both marking (Christ, McCrae, & Abrams, 2002) and inhibition of return (Watson & Maylor, 2002). Given that the cognitive processes most vulnerable to the effects of aging are often the last to develop, from either perspective it seems plausible to expect developmental improvements in the ability to deal with moving items in the 6- to 9-item range. For example, it has been argued that object-based inhibition of return relies on cortical mechanisms, whereas location-based inhibition of return involves subcortical mechanisms (e.g., McCrae & Abrams, 2001; Tipper et al., 1997). Cortical development lags behind subcortical development, and it seems possible that this might explain why counting 6–9 moving items is so much harder for young children than counting 6–9 static items.

Another interpretation of the age-related effect of motion in the 6- to 9-item range is related to a form of working memory. Perhaps it is in object-based processing that the demands of attention and working memory meet. When large numbers of items are to be enumerated, item movement may tax a certain form of working memory—that is, visuospatial memory; the viewer needs to be aware not only of where an item is but of where it used to be. Developmental increases in working memory span are well documented, and Tuholski et al. (2001) found that adults with small working memory spans had higher counting slopes than did those with larger spans. (There were no differences between these groups in subitizing slope; there were also no differences in the maximum number of items that could be subitized.) People with small memory spans were also more adversely affected than those with large memory spans by the presence of distractors that had to be distinguished from targets using the attentional focus. Tuholski et al. concluded that working memory capacity differences reflect differences in capability for controlled attentional processing. At this point, more investigation is required. Nonetheless, it is in counting that the interplay between basic visual, attentional, and working memory processes can best be seen.

In general, the results of the present study support the attention-based theory of spatial enumeration in that they show that motion has different effects in the subitizing and counting ranges. This is the only theory of enumeration according to which motion assumes such a critical role. However, there are several other theories, and it is relevant to consider the effect of motion in light of them.

Some theories explain patterns of performance in visuospatial enumeration by appealing to visual and spatial factors that are correlated with number but are not number per se. For example, one theory relies on a correla-

tion between number and pattern: One item forms a point, 2 items form a line, and 3 items often fall into a triangular configuration. According to pattern theory (Mandler & Shebo, 1982), enumeration of 1–3 items is simply pattern recognition. This theory offers no account of how the enumeration of 4 or more items occurs or how the enumeration process develops, and consequently it cannot explain the age-related change in the counting range. (There are other problems as well: This theory does not explain why people do not enumerate any linear configuration of items as 2 items or any triangular configuration as 3 items. The fact that people can avoid saying “two” when they see lines of 3, 4, 5, or more items suggests that they must already know how many items there are before they decide whether they can use pattern as a cue.)

Other theories explain visuospatial enumeration by appealing to principles so general that they could explain any kind of enumeration regardless of sensory modality (e.g., vision, hearing, touch) or dimension (i.e., spatial enumeration, in which units are objects laid out at different points in space at the same time, vs. temporal enumeration, in which units are events defined by existence at different points in time which may even occur in the same location in space). For example, Gallistel and Gelman (1992) proposed that the difference between subitizing and counting originates in the nature of the number representation. According to this account, the representation of 1–3 items is analogue and permits rapid preverbal enumeration, whereas the representation of higher numbers is different and requires a different (verbal) enumeration process. Because this account is so general, it has difficulty explaining how a spatial factor such as motion has different effects on visual enumeration depending on the number of items and the age of the participants. Although there is merit in studying factors common to all types of enumeration, some issues are unique to visuospatial enumeration, and for this type of enumeration variables related to spatial attention seem to determine whether or not people can enumerate small numbers of items using the fast and accurate subitizing process (Trick & Pylyshyn, 1994).

Overall, the present study establishes a link between the enumeration and multiple-object tracking tasks. Nonetheless, there are still important differences between tasks: (1) Multiple-object tracking involves distinguishing between targets and (identical-looking) distractors as well as tracking individual objects, and (2) multiple-object tracking requires extended concentration as items move over extended distances over longer periods of time. Given that both of these requirements might serve to make the classic multiple-object tracking task especially challenging for children, we thought it advantageous to start with a simpler task. In this study, we have shown that young children have at least some of the abilities needed for multiple-object tracking, although only a limited form of motion was tested. It may not always be possible to enumerate moving things, but learning when

enumeration is possible and when it is not may contribute to a broader understanding of vision and visual attention.

REFERENCES

- CHI, M. T., & KLAHR, D. (1975). Span and rate of apprehension in children and adults. *Journal of Experimental Child Psychology*, **19**, 434–439.
- CHRIST, S. E., MCCRAE, C. S., & ABRAMS, R. A. (2002). Inhibition of return in static and dynamic displays. *Psychonomic Bulletin & Review*, **9**, 80–85.
- FOLK, C. L., EGETH, H., & KWAK, H.-W. (1988). Subitizing: Direct apprehension or serial processing? *Perception & Psychophysics*, **44**, 313–320.
- GALLISTEL, C. R., & GELMAN, R. (1992). Preverbal and verbal counting and computation. *Cognition*, **44**, 43–74.
- JEVONS, W. (1871). The power of numerical discrimination. *Nature*, **3**, 281–282.
- JIANG, Y., CHUN, M., & MARKS, L. (2002a). Visual marking: Dissociating effects of new and old set size. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **28**, 293–302.
- JIANG, Y., CHUN, M., & MARKS, L. (2002b). Visual marking: Selective attention to asynchronous temporal groups. *Journal of Experimental Psychology: Human Perception & Performance*, **28**, 717–730.
- KAYE, B., & RUSKIN, E. (1990). The development of attentional control mechanisms. In J. T. Enns (Ed.), *The development of attention: Research and theory* (pp. 227–244). Amsterdam: Elsevier.
- KLAHR, D., & WALLACE, J. G. (1976). *Cognitive development: An information processing view*. Hillsdale, NJ: Erlbaum.
- KLEIN, R. M., & MACINNES, W. J. (1999). Inhibition of return is a foraging facilitator in visual search. *Psychological Science*, **10**, 346–352.
- LISS, P., & REEVES, A. (1983). Interruption of dot processing by a backward mask. *Perception*, **12**, 513–529.
- MANDLER, G., & SHEBO, B. J. (1982). Subitizing: An analysis of its component processes. *Journal of Experimental Psychology: General*, **111**, 1–22.
- MCCRAE, C. S., & ABRAMS, R. A. (2001). Age-related differences in object- and location-based inhibition of return of attention. *Psychology & Aging*, **16**, 437–449.
- OYAMA, T., KIKUCHI, T., & ICHIHARA, S. (1981). Span of attention, backward masking, and reaction time. *Perception & Psychophysics*, **29**, 106–112.
- PYLYSHYN, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition*, **80**, 127–158.
- PYLYSHYN, Z. W., ELCOCK, E. W., MARMOR, M., & SANDER, P. (1978). Explorations in visual-motor spaces. In *Proceedings of the Second International Conference of the Canadian Society for the Computational Studies of Intelligence*. University of Toronto.
- PYLYSHYN, Z. W., & STORM, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, **3**, 179–197.
- SAGI, D., & JULESZ, B. (1984). Detection versus discrimination of visual orientation. *Perception*, **13**, 619–628.
- SATHIAN, K., SIMON, T. J., PETERSON, S., PATEL, G. A., HOFFMAN, J. M., & GRAFTON, T. S. (1999). Neural evidence linking visual object enumeration and attention. *Journal of Cognitive Neuroscience*, **11**, 36–51.
- SEARS, C. R., & PYLYSHYN, Z. W. (2000). Multiple-object tracking and attentional processing. *Canadian Journal of Experimental Psychology*, **54**, 1–14.
- SNYDER, J. J., & KINGSTONE, A. (2000). Inhibition of return and visual search: How many separate loci are inhibited? *Perception & Psychophysics*, **62**, 452–458.
- STARKEY, P., & COOPER, R. (1995). The development of subitizing in young children. *British Journal of Developmental Psychology*, **13**, 399–420.
- TIPPER, S. P., JORDAN, H., & WEAVER, B. (1999). Scene-based and object-centered inhibition of return: Evidence for dual orienting mechanisms. *Perception & Psychophysics*, **61**, 50–60.

- TIPPER, S. P., RAFAL, R., REUTER-LORENZ, P. A., STARRVELDT, Y., RO, T., EGLY, R., DANZIGER, S., & WEAVER, B. (1997). Object-based facilitation and inhibition from visual orienting in the human split-brain. *Journal of Experimental Psychology: Human Perception & Performance*, **23**, 1522-1532.
- TIPPER, S. P., WEAVER, B., JERRETT, L., & BURAK, A. (1994). Object- and environment-based inhibition of return in visual attention. *Journal of Experimental Psychology: Human Perception & Performance*, **20**, 478-499.
- TRICK, L. M. (2003, June). *Age-related changes in the effect of response uncertainty*. Poster presented at the annual meeting of the Canadian Society of Brain, Behaviour, and Cognitive Science, Hamilton, Ontario.
- TRICK, L. M., DALES, L., & AUDET, D. (2002, June). *Enumerating items as they move: Item individuation and tracking*. Poster presented at the annual meeting of the Canadian Society of Brain, Behaviour, and Cognitive Science, Vancouver.
- TRICK, L. M., & ENNS, J. T. (1998). Life-span changes in attention: The visual search task. *Cognitive Development*, **13**, 369-386.
- TRICK, L. M., ENNS, J. T., & BRODEUR, D. A. (1996). Lifespan changes in visual enumeration: The number discrimination task. *Developmental Psychology*, **32**, 925-932.
- TRICK, L. M., & PYLYSHYN, Z. P. (1994). Why are small and large numbers enumerated differently? A limited capacity preattentive stage in vision. *Psychological Review*, **101**, 80-102.
- TUHOLSKI, S. W., ENGLE, R. W., & BAYLIS, G. C. (2001). Individual differences in working memory capacity and enumeration. *Memory & Cognition*, **29**, 484-492.
- ULLMAN, S. (1984). Visual routines. *Cognition*, **18**(1-3), 97-159.
- VAN OEFFELEN, M. P., & VOS, P. G. (1982). Configurational effects on the enumeration of dots: Counting by groups. *Memory & Cognition*, **10**, 396-404.
- WATSON, D. G. (2001). Visual marking in moving displays: Feature-based inhibition is not necessary. *Perception & Psychophysics*, **63**, 74-84.
- WATSON, D. G., & HUMPHREYS, G. W. (1997). Visual marking: Prioritizing selection for new objects by top-down attentional inhibition of old objects. *Psychological Review*, **104**, 90-122.
- WATSON, D. G., & HUMPHREYS, G. W. (2000). Visual marking: Evidence for inhibition using a probe-dot detection paradigm. *Perception & Psychophysics*, **62**, 471-481.
- WATSON, D. G., & MAYLOR, E. A. (2002). Aging and visual marking: Selective deficits for moving stimuli. *Psychology & Aging*, **17**, 321-339.

NOTE

1. If RT analyses were performed with age, condition, and the full range of numbers (1, 2, 3, 4, 6, 7, 8, and 9), the results were as follows. All main effects were significant at $p < .01$ [age: $F(3,130) = 83.4$, $MS_e = 1,642,123$; condition: $F(1,130) = 7.9$, $MS_e = 1,642,123$; number: $F(7,910) = 687.8$, $MS_e = 258,492$]. There were also significant interactions between number and age [$F(21,910) = 35.5$, $MS_e = 25,849$, $p < .001$] and number and condition [$F(7,910) = 5.2$, $MS_e = 25,849$, $p < .001$]. A marginal three-way interaction between age, number, and condition emerged [$F(21,910) = 1.56$, $MS_e = 258,492$, $p = .054$].

(Manuscript received December 11, 2002;
revision accepted for publication July 14, 2003.)