Multiple-object tracking involves simultaneously tracking positions of a number of target-items as they move among distractors. The standard version of the task poses special challenges for children, demanding extended concentration and the ability to distinguish targets from identical-looking distractors, and may thus underestimate children’s tracking abilities. As a result, a modified version of the task called “Catch the Spies” was developed. Participants tracked one to four moving “spies” (targets) that had “disguised” themselves so that they could blend in with a crowd of 10 people (happy-faces). Tracking accuracy was measured in five age groups (6, 8, 10, 12, and 19 years old). All performed well above chance though there were age-related increases in the number that could be tracked at once. Overall, when the effects of age were statistically controlled, tracking performance was significantly better for action videogame players than non-players, and marginally better for action-sports participants than non-participants.

© 2005 Elsevier Inc. All rights reserved.

Keywords: Multiple-object tracking; Age differences; Attention; Working memory; Videogames
short period of time some of the items flash to indicate that they are targets. Once the targets stop flashing they become identical in appearance to distractors, and targets and distractors alike are sent into random, independent motion. After a period of time participants are asked to identify the items that are targets. The number of targets to be tracked at once is manipulated; target identification accuracy is measured.

Pylyshyn and Storm (1988) conducted the first of these studies, and their results suggest young adults can track four to five items simultaneously. These findings were interpreted as evidence of mental reference tokens called FINgers of INSTantiation (FINSTs) that are used to select a small number of items as future targets for the attentional focus. According to this account, each target is assigned a reference token, and these reference tokens allow the targets to be perceived as distinct from distractors and one another, permitting targets to maintain their identities even when they move and change properties. There are only a limited number of reference tokens because FINSTs serve to select some items for deeper (attentional) processing – it makes little sense to select everything at once. Consequently, there are corresponding limits to the number of items that can be individuated and tracked at once (Pylyshyn, 2001). Item individuation and tracking are thought to be fundamental to visual-motor coordination and the creation and maintenance of object-files (short-term episodic representations of all of the properties for selected individual objects within a visual scene, see Treisman, 1993). The theory is subject to debate but the basic finding that young adults can track four to five items at once has been replicated using a variety of techniques (e.g., Bahrami, 2003; Culham et al., 1998; Scholl & Pylyshyn, 1999; Sears & Pylyshyn, 2000; Yantis, 1992), though there may be expertise-related differences between different adult populations (Allen, McGeorge, Pearson, & Milne, 2004).

The multiple-object tracking task occupies an interesting position in the literature, spanning the boundaries between perception, attention, and working memory. It requires an ability to construct a representation of the world that is extended in time – one in which items’ identities are sometimes defined by their history. Studies with infants suggest the ability to track may develop early in life (e.g., Carey & Xu, 2001; Scholl & Leslie, 1999), and one recent study suggests the number of objects tracked may be stable by the age of 5 years old (Black & Pylyshyn, 2004).

However, there are reasons to suspect age-related changes in multiple-object tracking after the age of 5. Although there is controversy about whether multiple-object tracking is best considered an aspect of attention or working memory (e.g., Fougnie & Marois, 2004; Trick & Pylyshyn, 1994, respectively), there is reason to suspect age-differences because there are developmental improvements in both during the school years. For example, there are improvements in many indices of selective attention (see, Plude, Enns, & Brodeur, 1994, for a review), and it has been argued that the most notable developments occur in controlled-endogenous selection because the brain areas involved in this type of selection are among the last to mature (Enns & Trick, in press; Trick, Enns, Mills, & Vavrik, 2004). Controlled-endogenous selection involves the execution of specific attention-demanding process at the expense of other such processes and reflects goals that are determined by an individual’s specific knowledge, plans, and strategies for a certain situation. Multiple-object tracking clearly requires this type of selection because it requires deliberately selecting certain chosen targets and ignoring distractors. Although the attentional focus can only be directed to one object location at a time, Trick and Pylyshyn (1994) argue that in order to reposition the
attentional focus systematically from object to object when objects move, there needs to be a way to select and track more than one object at once so that items are not missed or visited twice. This ability is fundamental to deliberately controlling the position of the attentional focus.

There are also developmental improvements in short term/working memory during the school years. Baddeley’s tri-component theory dominates the working memory literature (Baddeley & Logie, 1999), and according to this account there are three limited capacity short-term stores: an articulatory-acoustic loop for linguistic information, a visual-spatial sketchpad for visual and spatial information, and an executive that controls and supervises the operation of the other two stores. These three components are thought to develop independently (Gathercole, 1998; Gathercole, Pickering, Ambridge, & Wearing, 2004). Of primary importance is the visual-spatial sketchpad, which is assessed using a variety of measures, including the Pattern Span and the Corsi Blocks task (Gathercole, 1998; Vecchi, Phillips, & Cornoldi, 2001). Recent evidence suggests that there may be separate stores for visual and spatial information (e.g., Logie & Pearson, 1997; Pickering, Gathercole, Hall, & Lloyd, 2001; Vicari, Belluci, & Carlesimo, 2003). Regardless, during the school years there are increases in the number of visual/spatial locations that can be correctly reported from memory, and given that tracking requires remembering item positions, the ability to store and correctly report the positions of tracked items should improve as well.

The goal of this study is to investigate the development of multiple-object tracking. However, there are special challenges associated with designing a multiple-object tracking task for children. The standard multiple-object tracking task requires extended concentration and the ability to distinguish targets from identical-looking distractors as well as the ability to track. If children fail to find the task engaging, or fail to see the importance of distinguishing between identical items (usually circles) on a screen, there is a danger that their performance on the task will underestimate their actual tracking abilities. Conversely, if the task is made too easy (e.g., with too short a tracking interval and too few distractors) it may not be demanding enough to reveal age differences. In this study we developed a version of the task called “Catch the Spies”, designed to be comprehensible and engaging to children, so that we could test them under the same stringent conditions normally employed with adults. The game involves tracking the positions of a number of sinister individuals in trench coats (spies) that have disguised themselves to blend in with other people in a crowd (happy-faces).

We used a difficult version of the task (10 item displays, 10 s tracking interval) and full report, which required participants to remember the positions of all targets. Though this technique requires that information be held in working memory, it has advantages when used for testing children. For one, it is easier to determine the number of items that are guessed. More important, full report requires fewer trials than partial report. In partial report participants track a certain number of targets and then are required to decide if one selected item is a target or distractor. This technique requires twice as many trials because for every number of items tracked, there needs to be equal numbers of trials where the selected item is a target and a distractor. Given that it is difficult to sustain children’s interest for extended numbers of trials, the full report technique was judged more appropriate for this first investigation. As it turns out, working memory demands may well be inevitable in
multiple-object tracking. There is evidence of memory scanning even with partial report. Pylyshyn and Storm (1988) found that response times to decide whether a single item was a target increased with the number of targets to be tracked, as if the participants were serially scanning through a list of target locations to make a decision about whether a single item was a target.

When introducing a new measure it is important to demonstrate that it is related to performance in everyday tasks. Given that multiple-object tracking should be important in situations where individuals have to coordinate their actions to multiple moving objects (items on a screen, projectiles, other players), experience with action-related videogames or sports should be associated with superior tracking accuracy. The prediction related to videogames has already found support in investigations of young adults. Green and Bavelier (2003a) observed experience with an action-videogame increased the number of items that could be tracked in a standard multiple-object tracking task. In fact, they found 10 h of practice at an action-videogame improved performance in a variety of attentional tasks; practice with other types of videogame did not (Green & Bavelier, 2003b). There are serious ethical and practical issues that preclude manipulating exposure to action-sports and videogames in school children in the present study. Nonetheless, Green and Bavelier’s findings suggest there should be an association between tracking accuracy in the Catch the Spies task and reported exposure to action-videogames and it seems reasonable that the same should hold for exposure to action-sports.

1. Method

1.1. Participants

The 97 participants were from five age groups: 6, 8, 10, 12, and 19 years (see Table 1). The children were from a local school whereas the 19-year-old participants were from the university participant-pool. Data from participants diagnosed with visual or attentional learning disabilities were excluded from these analyses, as were data from participants who had failed a grade.

1.2. Apparatus and materials

Testing was done on Macintosh G4 PowerBooks that had 21.5 cm × 32.5 cm viewing screens. The tracking field (the area in which items moved) on the computer screen was a

<table>
<thead>
<tr>
<th>Age group</th>
<th>Mean age (months)</th>
<th>SD age (months)</th>
<th>Grade</th>
<th>n</th>
<th>Number of females</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 years old</td>
<td>76</td>
<td>3.6</td>
<td>1</td>
<td>15</td>
<td>6</td>
</tr>
<tr>
<td>8 years old</td>
<td>99</td>
<td>3.4</td>
<td>3</td>
<td>20</td>
<td>13</td>
</tr>
<tr>
<td>10 years old</td>
<td>123.6</td>
<td>4.1</td>
<td>5</td>
<td>20</td>
<td>9</td>
</tr>
<tr>
<td>12 years old</td>
<td>146.3</td>
<td>3.4</td>
<td>7</td>
<td>19</td>
<td>8</td>
</tr>
<tr>
<td>19 years old</td>
<td>231.6</td>
<td>12.0</td>
<td>University</td>
<td>23</td>
<td>18</td>
</tr>
</tbody>
</table>
central black rectangle occupying $22.96^\circ \times 17.33^\circ$ visual angle when viewed from 45 cm. The remainder of the viewing screen was light grey. A fixation point stood at the centre of the tracking field, a $0.18^\circ$ white outline square that items could be seen to pass under during animation sequences. There were two types of item, $1.45^\circ$ blue happy-face figures and $1.53^\circ$ spies (see Fig. 1). Happy-face figures could move, and they moved randomly and independently of one another, avoiding each other and the edge of the tracking field. These items could touch but never occlude one another (they repulsed one another). Each item’s rate of movement varied from frame to frame, ranging between 0 and $9.35^\circ$/s (a random 0–3 pixels per frame in each of the horizontal and vertical directions, with each frame requiring 16.5 ms). Given that items could move 0–3 pixels in each of the vertical and horizontal directions every frame, that meant there were seven possible vertical movements (three up, two up, one up, zero up, one down, two down, three down) and seven possible horizontal movements (three left, two left, one left, zero left, one right, two right, and three right) for each item, for a total of 49 possible courses of action that each item could take. Item movement featured some degree of inertia, though there was also an element of randomness. Each frame of motion there was a $1/200$ probability that the item would change its horizontal direction (from left to right) and a $1/200$ probability it would change its vertical direction (from upwards to downwards).

Questionnaires were administered to find out which, if any, videogames and sports the participant engaged in. As well, information was gathered about medications and diagnosed visual or attentional learning disabilities.

1.3. Procedure

Questionnaires were administered before computerized testing. Parents filled out the questionnaire for the children whereas 19-year-old participants filled it out for themselves. Two female research assistants conducted the testing with computers, which occurred in small rooms under normal daylight lighting conditions. Participants were seated 45 cm from the computer screen. They were told that they were going to play a computer game in which they had to keep track of some spies that were trying to escape by disguising themselves to look like other people. Each “Catch the Spies” trial involved five phases:

1. Initialization. Participants hit the “OK” button to initiate the trial. Ten static happy-face figures were presented for 1105.5 ms on the screen. These figures appeared in random locations on the tracking field.
2. **Target acquisition.** For 1650 ms one to four of the happy-face figures alternated between happy-face and spy form (165 ms as happy-face, 165 ms as a spy, for the duration). This was done to signal that those items were spies – i.e., targets. (Pilot testing suggested that children as young as 5 years old could learn the positions of the targets within this time. Extending the target acquisition period increased the risk that children would become bored and look away from the screen.) Afterwards there was a 495 ms pause in which all 10 items resumed their happy-face form and remained static.

3. **Tracking.** The happy-faces (including spies disguised as happy faces) began moving randomly and independently of one another. Movement continued for 10 s.

4. **Report.** Items stopped moving and participants used the computer mouse to select the happy-face figures that were “really spies” (targets). Each happy-face item turned red when the mouse passed over it and green if the participant clicked on the item to select it. A second mouse click de-selected the item if participants changed their minds. Once participants were satisfied they had selected the spies, they hit the “OK” button to submit their response.

5. **Feedback.** The “spies” revealed themselves, alternating between happy-face and spy form for 1260 ms (165 ms as a happy-face, 165 ms as a spy, for the duration).

Children underwent pre-trial training to ensure they could use the computer mouse and distinguish targets from distractors even when the target (spy) was in disguise as a happy-face. Participants of all ages did the graded practice trials (eight trials, two at each target numerosity starting with the one target condition). There were 40 randomly ordered experimental trials, 10 at each target numerosity. Participants were encouraged to rest between trials if they felt the need.

### 2. Results

The dependent measure was the percentage of accurately identified targets. Thus, in a given trial, if a participant only identified three of the four targets that they were required to track, their accuracy would be 75%. Accuracy was measured as a function of age (6, 8, 10, 12 and 19 years) and number of targets (one to four). Because no main effects or interactions involving the sex of the participant emerged ($F < 1$ for all), data for males and females were combined.

Participants of all ages appeared to understand the task and find it engaging. The Catch the Spies performance for the 19-year-old participants was comparable with that of young adults in more standard versions of the multiple-object tracking task. The 19-year-old participants tracked one to four targets with great accuracy (96.5% on average) though there was an effect of target numerosity ($F(3,66) = 7.3, p < .001, \eta^2 = .25$). Participants of all ages performed far better than would be expected if they simply guessed the positions of all the targets. Nonetheless, as can be seen from Fig. 2, the age of the participant had a pronounced effect, especially at higher target numerosities. Because of the dramatic effects of age, there were unavoidable violations of the homogeneity of variance assumption. We used both the inverse sine transformation (Kirk, 1982, p. 83), and the Greenhouse-Geisser correction to compensate as much as possible. When analysis of variance was performed, all
Fig. 2. Mean tracking accuracy for 6-, 8-, 10-, 12-, and 19-year-old participants when tracking one to four spies in a display of 10 moving items (standard error bars included). Dotted lines indicate expected accuracies if participants were randomly guessing the positions of one or two targets.

Main effects and interactions were significant (age: $F(4,92) = 30.3, p < .001$, partial $\eta^2 = .57$; number of targets: $F(2.3,213.6) = 90.3, p < .001$, partial $\eta^2 = .50$; number of targets $\times$ age: $F(9.3,213.6) = 5.4, p < .001$, partial $\eta^2 = .19$).

Planned comparisons were performed on the transformed data to investigate the effects of age, compensating for inequalities in variance when necessary. The most meaningful comparisons were for the lowest and highest number of targets (one and four targets, respectively). When there was only one item to track, the 6-year-old children performed significantly worse than the 8-year-olds ($t(33) = -3.0, p < .01$) but the accuracy for the other four age groups approached 100%, and there were no significant differences ($F(3,78) = 2.18, p > .05$). For four items, age-differences were more apparent, and this was also mirrored in different amounts of variability within different age groups. The conservative Tamhane procedure was used to deal with unequal variances when performing tests of means (Kirk, 1982, pp. 120–121). This procedure revealed no significant differences between the 6 and 8-year-olds when tracking four items ($p > .05$). Nonetheless, the 6-year-olds differed significantly from those who were 10 or more years in age; the 8-year-olds differed significantly from those 12 years or more in age; and the 10-year-olds differed significantly from the 19-year-olds ($p < .05$ for all).

However, direct comparisons of tracking performance are not meaningful without a consideration of the expected outcome if participants were guessing. The most straightforward way to accomplish this was to compare observed performance when participants tracked two to four targets with what would be expected if participants were guessing the locations of some of the targets. Expected outcomes were derived for each number of targets by calculating the probability of randomly guessing the position of one or two of the targets (Freund, 1981, p. 181), and using these probabilities to calculate expected outcomes given that some items were truly tracked and some were guessed. Thus, for example, if a participant was given the task of tracking three items at once, and only tracked two and guessed one location from the remaining eight positions, their expected accuracy would be $(2 + 1/8)/3 = 70.83\%$. These expected accuracies are plotted in Fig. 2.
We then looked for the first place where the observed and expected accuracy functions crossed for a given age group. For example, from Fig. 2 it is apparent that the observed accuracy function for the 8-year-old participants approaches the expected accuracy function for guessing one location when tracking three items at once ($t(19) = 1.50, p > .05$). The 10-year-old children performed significantly better than they would have had they been guessing one location when tracking three targets ($t(19) = 7.13, p < .001$). Therefore, the 8-year-old children performed about as well as would be expected if they could track two items at once whereas the 10-year-old children performed about as well as would be expected if they could track three. Similarly, the observed function for the 12-year-old participants approaches the expected accuracy function for guessing one item at when the task is to track four items at once ($t(18) = 1.63, p > .05$). The 19-year-old participants performed significantly better than they would have had they been guessing one location when tracking four ($t(22) = 10.17, p < .001$). As a result, we concluded that the 12-year-old participants performed about as well as would be expected if they could track three at once whereas the 19-year-old participants performed as if they could track up to four.

Finally we determined the lower limits of tracking performance by looking at the data from the 6-year-old children. They had poorer accuracy than the older participants, and this difficulty was manifest even when there was only one target, which suggests that the 6-year-old children may have been having difficulties sustaining concentration over the time between the initial presentation of the items and the end of the tracking interval. Nonetheless, despite the demanding conditions of this study (10 s tracking intervals, 10 item displays, full report), the 6-year-old children seem to be capable of tracking more than one target at once because they performed significantly better than would be expected if they were guessing one target location when tracking two ($t(14) = 3.03, p < .01$).

To summarize, these results suggest that there are age-related improvements in multiple-object tracking performance. The 6- and 8-year-old children perform about as well as would be expected if they could track two items at once; 10- and 12-year-olds perform as if they could three at once; and 19-year-olds perform as if they could track four at once.

Experience with action-sports and videogames should be related to better tracking performance if multiple-object tracking is related to dynamic visual-motor coordination. For purposes of this study, action-videogames and sports were coded as those that involved rapidly coordinating movements to the actions of items on the screen, moving projectiles, or other players. Thus, videogames that involved shooting, chasing, or driving were considered action-games whereas puzzle and strategy games were not. Hockey, soccer, ultimate Frisbee, martial arts, etc. were coded as action-sports whereas swimming and golf were not.

The results of these analyses are shown in Fig. 3. Analysis of covariance (ANCOVA) was used to statistically control the powerful effects of age (age in months). Once the effects of age were controlled, no violations of the homogeneity of variance assumption occurred (Levene’s test, $p > .05$). Fifty-three participants played action-videogames (6, 9, 17, 15, and 6, respectively in 6–19 age groups: 83.7% of the males and 31.5% of the females). ANCOVA revealed an effect of action-videogames ($F(1,94) = 4.52, p < .05, \eta^2 = .05$), and a marginal action videogame $\times$ number of targets interaction ($F(2.2, 95.6) = 2.77, p < .06, \eta^2 = .03$, Greenhouse-Geisser correction). Fig. 3a presents adjusted mean percentages of correctly identified targets for those who were and were not reported to have experience with
action-videogames. Those with reported experience were more accurate but this difference was marginally more pronounced with larger numbers of targets. This finding is consistent with Green and Bavelier (2003) insofar as action-videogame players performed better than others especially with large numbers of targets.

Based on the questionnaire data, 67 individuals participated in action-sports (8, 17, 10, 15, and 17, respectively in the 6–19-year-old age groups; 69.8% of the males and 68.5% of the females). ANCOVA revealed a marginal effect of reported participation in action-sports once the effects of age were covaried out ($F(1,93) = 3.4$, $p < .07$, $\eta^2 = .04$) Fig. 3b shows the adjusted mean percentages of correctly identified targets for this analysis. These effects are weaker than those shown for videogames, but are novel insofar as they suggest that activities other than practice at computerized action-videogames may be related to superior performance on (computerized) multiple-object tracking tasks. The relationship between multiple-object tracking and participation in action-sports is consistent with other studies that have shown relationship between attentional indices and measures of motor coordination (Willson, Maruff, & McKenzie, 1997). However, these results are preliminary. The effects need to be replicated on larger samples, monitoring and controlling a host of factors (e.g., general physical fitness, familiarity with computers, parental involvement, socio-economic status, competitive orientation, general ability, working memory, standard measures of attention).
Because the performance of young adults was almost at ceiling in the first experiment, an additional 19 university students (mean age = 19 years: 0.3 months, SD = 9.7 months, 14 female) were tested on a more difficult version of the Catch the Spies task with three to six targets in 14 moving items (all else was the same). This study also proved useful in assessing the impact of the number of distractors on adult performance. The performance of a second group of 19-year-olds was compared with that of those in the initial study (with 10 item display size). Increasing the display size from 10 to 14 resulted in accuracy reductions for tracking three and four targets, especially at four (display size \( \times \) number of targets \( F(1,40) = 23.6, p < .001, \) partial \( \eta^2 = .37 \); number of targets \( F(1,40) = 27.1, p < .001, \) partial \( \eta^2 = .40 \); display size \( \times \) number of targets \( F(1,40) = 11.1, p < .005, \) partial \( \eta^2 = .22 \). In fact, as shown in Fig. 4 participants performed about as well as would be expected if they were tracking three positions and guessing the rest when there were 14 items in the display \( t(18) = 0.1, p > .05 \). This is consistent with a recent study on young adults that suggests that there may be differences in the number tracked based on display conditions (Alvarez & Franconeri, 2004).

3. Discussion

This study shows that age has significant effects on tracking performance in the 10 item “Catch the Spies” task, and the effects are particularly pronounced when there are three or four targets. This study also provides preliminary evidence that performance in the “Catch the Spies” task may be related to reported experience with real-life visual-motor tasks such as playing action-videogames and sports. However, before firm conclusions can be made this association needs to be replicated, controlling and monitoring the effects of a variety of factors such as socio-economic status, parental involvement, etc. Nonetheless, as a first investigation of the relationship between tracking performance and real-life activities in children, and given the limitations of third-person retrospective reports, these results were seen as promising.
The present article makes two principal contributions. The first is methodological. This study presents a new technique for measuring multiple-object tracking – one that reveals interesting age differences in performance that are also related to experience with certain real-life tasks. The Catch the Spies task is versatile in that it can be used with a wide variety of age groups (from children to senior adults) and it is easily modifiable, making it possible to look at a wide variety of issues by manipulating aspects of the task (the encoding conditions, the report procedures, activities during the tracking interval, rate of item motion, spatial distribution of items).

The second contribution is empirical. This study shows that there are improvements in tracking performance in late childhood into adolescence. What is it that is changing to produce superior performance in older participants? There are many possibilities. Perhaps children are acquiring knowledge or specific strategies that enable them to perform better. Perhaps they are becoming more efficient because of improved capacity or faster processing speed. Perhaps the improvements occur because children get better at sustaining their concentration for extended periods of time or engaging in operations that require high levels of spatial or temporal resolution. These improvements may be specific to multiple-object tracking or may reflect more general factors in cognitive development.

At this point it is impossible to determine which of these factors produced the age differences in performance. This is the first investigation of multiple-object tracking in this age range and the answers to this question are beyond the scope of one paper. However, it is possible to work through possibilities, eliminate some at the outset, and then consider the next line of research. What does multiple-object tracking require? For one, it requires a period of extended concentration on a visual display (over 13 s by the end of the tracking interval). This is true regardless of the number of items to be tracked at once. Consequently, one way to assess whether concentration span is important is to look at the performance for tracking one item. There is indeed evidence that differences in concentration span might explain some of the age-differences because the 6-year-old participants scored significantly worse than the other participants even when there was only one target. Beyond that age performance approached 100% when tracking one item and this suggests that concentration span may have only been a serious issue for the youngest participants.

Multiple-object tracking also requires storing and reporting the locations of multiple visual objects as well as selecting items and maintaining item selection in the face of item movement. In fact, limitations in visual working memory and selection are tightly coupled, because items have to be selected before they can be reported. However, working memory and selection can be differentiated insofar as they are susceptible to the influence of different variables. Per se, report of the final location of the item should be unaffected by item motion, though it might be affected by the presence of previously encoded visual or spatial information that has to be held in memory (e.g., Fougnie & Marois, 2004). Conversely, the ability to select and maintain selection in the face of item movement should be influenced by factors such as the presence of other moving items (targets or distractors) in the visual display.

Given that children showed decrements in performance even when there were only three or four targets, there is little to support the idea that these results were solely the result of limitations in working memory. There are a number of ways to measure visual/spatial working memory, but all show that even 6-year-old children can report the positions of
more than four items. For example, Pickering, Gathercole, Hall, and Lloyd (2001) presented children with two-dimensional matrices of squares in which some squares were black and others were filled. The matrix was removed and children were shown and empty matrix and required to point to the location of all the blackened squares. (Their choices were entered in via mouse click, as in the present study.) Even the 5-year-olds were capable of remembering 10 locations in this task. In fact, 5-year-old children correctly reported the locations of more than four items even when the squares were blackened one at a time, and the children were required to hold the positions in memory until all items were presented and then point to the locations in order. Our pilot study also showed that even the youngest children could encode and then report the positions of the maximum number targets used in this study. It is possible that there may be age-differences in terms of how long these spatial positions can be held in memory, but research on other types of working memory suggest that the important factor is not the duration of the interval between encoding and retrieval, but rather the type of processing that is going on during this interval (Gavens & Barrouillet, 2004). Working memory may indeed play an important role in performance in the multiple-object tracking task, but in order to investigate this a comparison would have to be made between simple delayed recall of item positions and recall when participants are performing secondary tasks that would demand different types of working memory (processes that tap the articulatory-acoustic loop, visual-spatial sketchpad, and the central executive).

Finally, it is possible that age-differences in tracking reflect differences in attentional selection: changes in the ability to select items and maintain selection (track them) even when items change properties (from spy to happy-face) and move randomly and independently through the display. However, tracking requires two distinct abilities. The first is item individuation, the ability to consider each of the targets as separate and distinct from other targets. More generally, this ability is necessary if an individual is to attend, foveate, or point to a specific target among others. However, it is also necessary to track individual targets when the items move independently of one another – even if the targets are used to define vertices in shapes created from the bounding contour for the targets (see Yantis, 1992). For purposes of simple individuation, the critical variable is the number of targets. As the number of targets increases there are increased chances for individuation failures: items missed or “fused” into one token. Perhaps there are age-differences in the number of items that can be individuated, but if this were true, this difference should emerge even in distractor-free paradigms such as enumeration. Enumeration requires item individuation because determining how many items there are requires that each item be considered once and only once (e.g., Trick & Pylyshyn, 1994; Wagner & Carey, 2003). The ability to individuate and maintain item identities as items move might thus be productively explored by investigating the ability to enumerate moving items (see Trick, Audet, & Dales, 2003).

However, the multiple-object tracking task has multiple targets AND multiple distractors. The second aspect of selection involves the ability to ignore distractors. He, Cavanagh and Intriligator (1997) demonstrated that there is a difference between sensory spatial resolution (the minimum distance at which a gap can be seen between contours) and attentional resolution (the minimum distance between items necessary so that a person can selectively attend one and ignore the other). Perhaps it is spatial resolution of attention that improves with age, and if this were true, the critical variable would be the number of distractors. Increasing the
number of distractors would reduce the average distance between a target and a distractor, thus requiring higher spatial resolution attentional analysis. There are already a number of studies that show that during the school years there are improvements in the ability to attend target items while ignoring distractors (see Plude et al., 1994; Trick, Trick, & Enns, 1998). In general, in developmental studies, when attention is needed to make a perceptual discrimination, the effect of the number of distractors diminishes with age. The problems that children have with distractors may be magnified in dynamic displays such as the ones used in the tracking study, because distractors have an opportunity to move closer to targets than is typical in static attentional tasks.

The present study shows clear evidence of the effect of the number of distractors on young adults. Performance was significantly worse when tracking three and four targets in 14-item than 10-item displays: young adults performed about as well as would be expected if they could track three at once in 14 items displays and four items at once in 10 item displays. However, when testing children we conformed to the usual practice in multiple-object tracking experiments, manipulating the number of targets while keeping the total number of items constant. Thus, as the number of targets increased the number of targets decreased (the number of targets + the number of distractors = 10). Therefore, put another way, the results of the present study show that children’s performance on the 10 item Catch the Spies task became worse as the number of distractors decreased. If the results of the present study conformed to previous studies of attention in children, performance would have improved as the number of distractors decreased.

Therefore, at this point it is important that the effects of attentional resolution (number of distractors) be disentangled from the requirements of target individuation (number of targets). The next step is to carry out a study in which the number of targets and the number of distractors are manipulated independently. It is possible that target individuation and distractor inhibition are accomplished by the same mechanism, as suggested by Pylyshyn (2001): FINSTs accomplish both tasks at once. However, it is also possible that these two operations require different abilities that develop at different rates.

Overall, multiple-object tracking is important to understanding how perceivers build a representation of the visual world that is extended in time, one that is adequate for action and thought. This article introduces a technique for investigating multiple-object tracking in children, one that reveals interesting patterns of age-related change and directions for future research. There is even preliminary evidence that performance in the Catch the Spies task is related to reported exposure to certain real-life activities: action-sports and videogames. These findings raise the intriguing possibility that certain types of sports and even the much-maligned videogame may play a role in honing certain attentional abilities.

Acknowledgements

This project was supported by a grant from the Natural Sciences and Engineering Research Council of Canada (238641-01). We would like to thank the following individuals: Jim Enns, Mary Ann Evans, and Zenon Pylyshyn for early discussions on the feasibility of this project (and ducks in disguise!); Brian Scholl for software consultation and helpful comments on an earlier draft of the paper; Ernest Dalrymple-Alford, Ian Newby-Clark and
Christopher Roney for statistical and editorial advice; and Tahlia Perl for help pilot-testing the young children. Some of the results for this study were presented at the annual meeting of the Canadian Society for Brain, Behaviour, and Cognitive Science (Trick, 2004).

References


