Available online at www.sciencedirect.com
SCIENCE DIRECT®

Cognition 101 (2006) 217–245

COGNITION

www.elsevier.com/locate/COGNIT

Enumeration versus multiple object tracking: the case of action video game players

C.S. Green *, D. Bavelier

Department of Brain and Cognitive Sciences, University of Rochester,
RC 270268, Meliora Hall, Rochester, NY 14627-0268, USA

Received 9 August 2005; received in revised form 12 October 2005; accepted 14 October 2005

Abstract

Here, we demonstrate that action video game play enhances subjects' ability in two tasks thought to indicate the number of items that can be apprehended. Using an enumeration task, in which participants have to determine the number of quickly flashed squares, accuracy measures showed a near ceiling performance for low numerosities and a sharp drop in performance once a critical number of squares was reached. Importantly, this critical number was higher by about two items in video game players (VGPs) than in non-video game players (NVGPs). A following control study indicated that this improvement was not due to an enhanced ability to instantly apprehend the numerosity of the display, a process known as subitizing, but rather due to an enhancement in the slower more serial process of counting. To confirm that video game play facilitates the processing of multiple objects at once, we compared VGPs and NVGPs on the multiple object tracking task (MOT), which requires the allocation of attention to several items over time. VGPs were able to successfully track approximately two more items than NVGPs. Furthermore, NVGPs trained on an action video game established the causal effect of game playing in the enhanced performance on the two tasks. Together, these studies confirm the view that playing action video games enhances the number of objects that can be apprehended and suggest that this enhancement is mediated by changes in visual short-term memory skills.

Keywords: Video games; Subitizing; Enumeration; Multiple object tracking

---

* Corresponding author. Tel.: +1 585 275 0695.
E-mail address: csgreen@bcs.rochester.edu (C.S. Green).

0022-2860/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2005.10.004

1. Introduction

1.1. Video games and visual skills

Video game play leads to a number of alterations in visual attention, visuo-motor coordination, and other perceptual/cognitive processes (Dorval & Pepin, 1986; Gopher, 1992; Gopher, Weil, & Bareket, 1994; Greenfield, DeWinstanley, Kilpatrick, & Kaye, 1994; Green & Bavelier, 2003; Li & Atkins, 2004; McClurg & Chaille, 1987; Orosy-Fildes & Allan, 1989; Subrahmanyam & Greenfield, 1994; Yuji, 1996). For example, several researchers have noted that video game play decreases the reaction time (RT) of subjects asked to respond to visual stimuli (Orosy-Fildes & Allan, 1989; Yuji, 1996). In the case of visual attention, Greenfield and colleagues (Greenfield et al., 1994) demonstrated that in a simple target detection task, expert video game players showed a diminished attentional cost (measured by RT) when presented with low probability targets compared to that observed in non-players, indicating enhancements in the ability to divide attention. Gopher and colleagues, working in collaboration with the Israeli military, demonstrated that cadets trained on a video game performed significantly better than their untrained peers on measures of flight performance (Gopher et al., 1994). Finally, our own work (Green & Bavelier, 2003) has demonstrated that video game players outperform non-players on different aspects of visual attention, in particular in the flexibility and efficiency with which they distribute attention over space and time. Such increases in visual attention in video game players may have real-world practical implications.

The goal of the present study is to assess the claim that, in addition to its effect on the spatial and temporal aspects of visual attention, action video game playing also modifies the number of objects that can be apprehended. In this paper, we compare the performance of video game players (VGPs) and non-video game players (NVGPs) on both an enumeration task and the multiple object tracking (MOT) task. These two particular tasks were chosen because they allow the separate assessment of the number of items that can be enumerated in parallel (also termed immediate apprehension), that can be enumerated serially (also termed counting), and that can be successfully tracked over time. Our results suggest a relationship between the number of items that can be accurately counted and the number of items that can be tracked. In contrast, the number of items that can be immediately apprehended seems to be under the control of distinct constraints. In addition, our results reveal for the first time a dissociation between reaction time measures and accuracy measures during the enumeration task. The consequences of these findings for our understanding of the mechanisms at play in enumeration and MOT studies are considered in Section 7.

1.2. The enumeration task

The enumeration task has classically been used to study the number of items that can be readily attended. This task requires participants to report the number of briefly flashed items in a display as quickly and as accurately as possible. Participant performance on this task appears best captured by two distinct processes, easily seen when RT is plotted against the number of items presented. When viewing enumeration performance in this manner a clear discontinuity, or elbow, is seen, giving the appearance of a bilinear function. For low numbers of items, usually in the range of one to four items (Atkinson, Campbell, & Francis, 1976; Oyama, Kikuchi, & Ichihara, 1981; Trick & Pylyshyn, 1993; Trick & Pylyshyn, 1994), subject performance is extremely fast. The RT slopes are near zero over this range-also termed the subitizing range. As numerosity increases above this range, each additional item has a substantially greater cost in terms of RT. The cost to performance is evident in the steep slope observed beyond about four items. The change in reaction time slopes is mirrored by a similar change in accuracy, with accuracy remaining stable and high within the subitizing range and exhibiting a linear decrease with increasing numerosity beyond the subitizing range. Although reaction times and accuracy are typically correlated in enumeration studies, the subitizing range has been defined in terms of reaction time as the range of numerosities that can be apprehended without a significant increase in reaction time as numerosity is increased.

The discontinuity in the enumeration curve has been the subject of much debate and has been posited to have various explanations. One proposal holds that the discontinuity marks a fundamental difference in the perceptual quality of the display when few versus many items are presented. Following this line of thought, some have proposed that performance when few items are present is exceptionally efficient because it relies on canonical pattern recognition or configurational effects, which are not available when many items are present (Mandler & Shebo, 1982; van Oeffelen & Vos, 1982). Along the same line, density differences have also been proposed to lead to the observed discontinuity (Atkinson et al., 1976). A second view holds that the discrepancy can be accounted for by a single cognitive process, counting, measured at different points of a continuum (Gallistel & Gelman, 1992). In this model, the elbow in the curve is thought to reflect a switch from a fast non-verbal mode of counting, to a slower verbal mode of counting (Gallistel & Gelman, 1992). This view is supported by the fact that many studies have demonstrated that the RT curve is not truly flat, even for low numerosity; instead there is often a small slope (between five and ten times shallower than what is seen for greater numbers of items) (Akin & Chase, 1978; Oyama et al., 1981). It has also been speculated that increases in memory load, light for few items but increasingly heavy for many items, may play a role. An alternative view posits that the discontinuity reflects two fundamentally different cognitive processes. From this viewpoint, the nearly flat region has been taken to reflect an automatic, parallel, and immediate process, which has been dubbed ‘subitizing’ (Kaufman, Lord, Reese, & Volkman, 1949). For simplicity, throughout the paper, the region over which performance is immediate will, for ease of exposition, be referred to as the subitizing range. According to this model, the mechanism(s) that underlie the subitizing process are hypothesized to be severely capacity limited (on the order of three to four items). When the number of items to enumerate exceeds the capacity of this automatic system, subjects must use a separate process wherein serial attention is employed to ‘count’ the items from visual short-term memory. As it is well known that information in STM decays over time, the process is slower and more error-prone. This is clearly reflected in the steep region of the performance graph, which will be referred to as the counting range throughout the remainder of the paper.

While there is much dissent as to whether the subitizing and counting range reflect two separate cognitive processes or a more continuous process with added constraints from short term memory and indexing as the numerosity increases, all parties agree that the subitizing range provides an estimate of the number of items that can be concurrently apprehended. We therefore decided to compare video game players (VGPs) and non-video game players (NVGPs) on the enumeration task.

2. Experiment 1

In Section 2, we asked whether playing action video games would alter enumeration ability. Section 2 also tests the view that video game play primarily affects peripheral vision. Although our own work (Green & Bavelier, 2003), as well as that of others (Orosy-Fildes & Allan, 1989), has shown improvements occurring in both central and peripheral vision, it is still commonly thought that video game play predominantly modifies peripheral vision. This view partly finds its support in the fact that many video games require subjects to distribute attention peripherally, as 'enemies' can appear at any location. To test the effect of video game play on both the central and peripheral visual field, VGPs and NVGPs underwent two different enumeration experiments, one with a field of view restricted around fixation (5° square; henceforth referred to as the narrow field of view condition) and the other with a much wider field of view (20° square; the wide field of view condition). If video game play disproportionately affects far peripheral vision, any VGP advantage over NVGPs should be magnified in the wide field of view condition. However, if the effects of video game play are consistent across retinal eccentricities, at least for the task under study, then performance should be somewhat equivalent between the two populations at both fields of view.

Using two fields of view also allows us to test the effect of field of view on the two components of the enumeration curve. The subitizing hypothesis suggests that the flat region of the performance graph reflects a preattentive spatially parallel mechanism and therefore predicts a similar subitizing span under both conditions. In contrast, models with canonical patterns or density as the primary causative mechanism predict that performance should be affected by this change in sparseness. In addition, the counting portion of the enumeration task could be adversely affected by the wider field of view as attending to a larger field of view reduces the amount of resources available at any particular location in space and increases the probability that the items will be inaccurately enumerated. The comparison of the two conditions may therefore help in distinguishing between the various accounts of enumeration performance. Sections 2 and 3 were partially reported in Green & Bavelier (2003).

2.1. Method

2.1.1. Subjects

Twenty-six males with normal or corrected-to-normal vision were placed into one of two groups, VGP or NVGP, based upon their responses to a questionnaire given prior to the experiment. Only males underwent testing because of the relative paucity of females with sufficient video game experience.

The criterion to be considered a VGP was a minimum of 3–4 days a week of action video game usage for the previous six months. Thirteen right-handed males with a mean age of 19.4 years fell into this category. Ten of the thirteen subjects reported daily video game usage for at least the previous six months, while the remaining three reported playing several times a week for the same time span. A highly abridged list of the games reported as played includes: Grand Theft Auto, Half-life, Counter-Strike, Marvel versus Capcom, Rogue Speare, and Super Mario Kart.

The criterion to be considered a NVGP was little, although preferably no, video game usage in the past six months. Thirteen males (eleven right- and two left-handed) with a mean age of 19.3 years fell into this category. Eleven of the thirteen members reported no video game experience whatsoever in the past year. The remaining two quantified their video game experience as once per month or less.

Written informed consent was obtained from each subject and each subject was paid $7.50 for each hour of participation.

2.1.2. Stimuli and procedure

Subjects viewed the display binocularly with their head positioned in a chin rest at a test distance of 60 cm. Each trial began with the presentation of a small white fixation cross in the center of a dark screen. After 500 ms, a stimulus consisting of a random number of white squares was presented for 50 ms (between 1 and 12 squares each subtending 0.5×0.5 degrees). Subjects were then allowed to respond. This study did not use the typical vocal response of enumeration studies; instead subjects were asked to press the number on the keyboard corresponding to the number of squares they believed were presented. Subjects were instructed to respond by a keyboard press as quickly as possible while maintaining accuracy.

For the response, the numbers 1–12 were placed on the keyboard above their respective keys (10 on the 0, 11 on the −, and 12 on the +). Subjects were instructed to use whatever key pressing strategy was most comfortable. Most used the four fingers of the left hand on 1–4, the four fingers of the right hand for 5–8 and moved the right hand for numbers above 8. However, some subjects used different strategies.

Each subject underwent two experimental blocks of 360 trials each (1–12 items presented, 30 repetitions of each number of items, pseudorandom presentation). The only difference between the two experimental blocks was the extent of the visual field over which the white squares could be presented. In the narrow field of view condition, the white squares were presented in a random location within an invisible boundary square in the center of the screen measuring $5^\circ \times 5^\circ$. It was also constrained such that there was at least a 0.5° separation between adjacent squares. In the wide field of view condition, the squares could be presented over a much wider field of view ($20^\circ \times 20^\circ$). It should be noted that because the squares were kept at the same size in the two field of view conditions, the square density was greater in the narrow than in the wide field of view condition.

The order of the two conditions was counterbalanced. No effect of order was found ($P > 0.7$) and thus will not be discussed further.

2.2. Results

Four measures of performance will be discussed for each experiment-accuracy breakpoint, percent correct, average response and RT.

Enumeration studies rely typically on reaction time analyses. In our case, however, the use of a keyboard response, rather than a vocal response, makes the interpretation of RT difficult. Although subjects were trained in advance in typing the key that corresponded to a given number, it is likely that this method of response nevertheless introduced additional variability in the measurement of RTs that may not be consistent across the two groups. Video game players are known to have faster key press RTs (Orosy-Fildes & Allan, 1989) and because there were 12 possible response keys to choose from, issues such as working memory (the ability of a subject to remember which key each finger was resting upon) and strategy (where the fingers were centered, use of both hands versus only one hand, etc.) could have played a role. With these caveats in mind, the RT data will be presented last.

The accuracy breakpoint refers to the point at which the discontinuity in accuracy occurs (the point where the elbow forms in the function). As previously discussed, accuracy performance during enumeration paradigms appears well approximated by a bilinear model. To determine an individual subject's accuracy breakpoint, their percent correct data was fit to a bilinear model using the least squares method. As our goal was to gain a quantitative measure of the breakpoint, as well as a measure of the slope of the two linear components that intersect at the breakpoint, only the data points for 1–8 items were entered into the model. This range was chosen based on previous evidence that the breakpoint lies approximately in the middle of this span, as well as to minimize the contribution of processes such as guessing, estimation biases, and strategy that may come into play for greater numbers of items. Each subject's curve was modeled as an intersection between two linear components; the first component was constrained to have a slope very near zero (maximum slope of 3% per item) while the second component was modeled as linearly increasing (as the data is plotted in terms of error rate, not percent correct) with the slope allowed full room to vary in order to best fit the data. The output of the model was therefore the slope of the two lines as well as the accuracy breakpoint-the point at which the two lines intersected. Although the accuracy breakpoint has classically been well mirrored by the subitizing span, or the shallow near zero slope of the reaction time data curve, the use of the term subitizing will be reserved for RT analyses in this paper, as Section 4 will reveal for the first time a dissociation between accuracy and RT measures of breakpoints.

The dependent variable percent correct is self-explanatory.

The last dependent variable is average response, as previous models have hypothesized a prominent role of estimation ability in enumeration performance (Krueger, 1982). Average response refers to the average response the subject made when presented with a given number of squares. Perfect performance corresponds to a line starting at the origin with a slope of one. Overestimation leads to a deviation above the perfect line, whereas underestimation leads to a digression below the perfect line.

Accuracy breakpoint was analyzed in a 2(VGP status: VGP/NVGP)×2(field of view: narrow/wide) ANOVA while percent correct, RT, and average response were each initially analyzed in a 2(VGP status: VGP/NVGP)×2(field of view: narrow/wide)× 12(number of squares) ANOVA.

2.2.1. Overall analyses

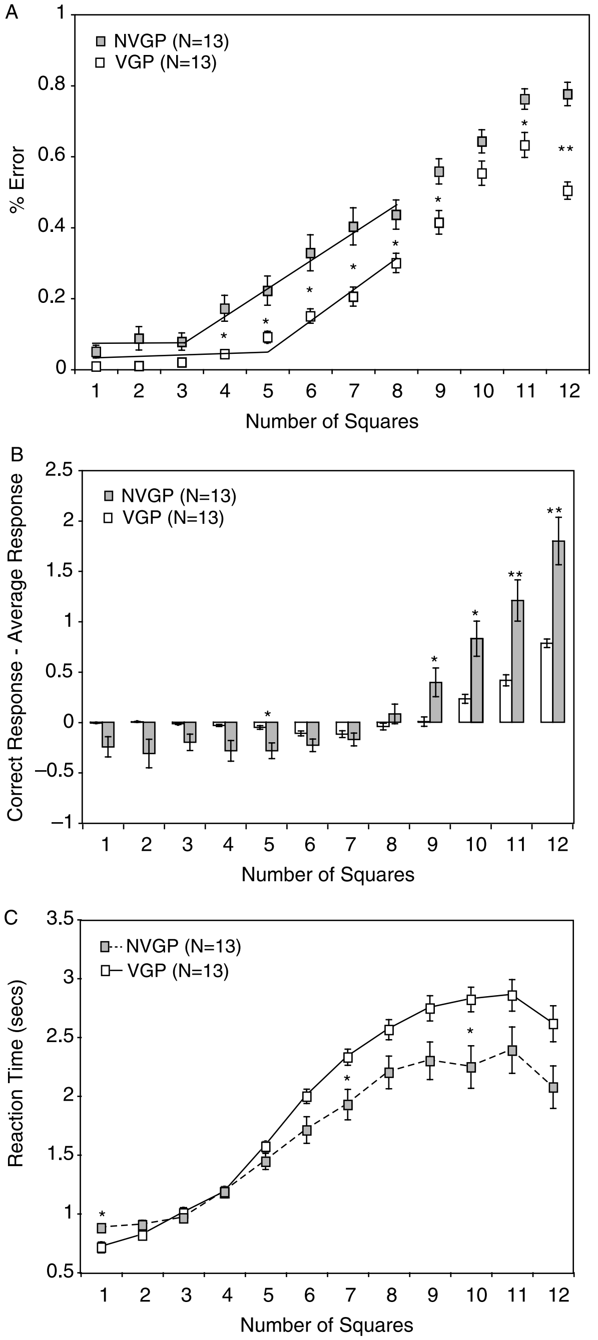
2.2.1.1. Accuracy breakpoint. The analysis of accuracy breakpoint revealed a clear main effect of VGP status with VGPs switching about 2 items beyond the accuracy breakpoint of NVGPs, VGP: 5.0+/− .2 squares, NVGP: 3.0+/− .3 squares, $F(1,24)=27.7$, $P<0.001$ (Fig. 1(A)). No interactions with field of view were observed. Separate analysis of the slope of each component yielded no main effect of VGP status, in particular, the lack of effect for the second slope indicates that both groups fall-off at a similar rate beyond their inflection point in both field of view conditions. While there was no effect of field of view in the analysis of the slope of the first component, a main effect of field of view was observed in the analysis of the second component, with the wide field of view having a significantly greater slope than the narrow field of view, wide: 13.7+/− 0.7%, narrow: 11.1+/− 0.7%, $F(1,24)=7.3$, $P=0.013$. Since the accuracy breakpoint was equivalent in the two field of view conditions, it appears that the ability to apprehend low numerosity is unperturbed by the size and density of the display, whereas the greater slope of the second component in the wide field of view condition suggests that the ability to count higher numbers of squares falls off more quickly in the wide field of view condition.

2.2.1.2. Percent correct. The expected strong main effect of number of squares was observed, with performance decreasing with increasing numerosity, $F(11,264)=202.8$, $P<0.001$. Importantly, a main effect of VGP status was observed with VGPs outperforming NVGPs, VGP: 75.5%+/−1.4% correct, NVGP: 62.4%+/−1.8% correct, $F(1,24)=9.1$, $P=0.006$ (Fig. 1(A)). Furthermore, the predicted VGP status×number of squares interaction was also found with both groups performing similarly well for low numbers (1–3), but VGPs outperforming NVGPs as the number of squares exceeded 3, $F(11,264)=3.3$, $P<0.001$.

Also, performance was better in the narrow field of view condition than in the wide field of view condition, narrow: 71.5+/−1.5% correct, wide: 66.4+/− 1.8% correct, $F(1, 24)=9.1$, $P=0.006$. This main effect can largely be attributed to differences in performance for the very large numbers (10–12) between the two fields of view which was further reflected in the observed interaction of field of view and number of squares, $F(11,264)=2.5$, $P=0.005$.

Finally, a VGP status×field of view×number of squares interaction, $F(11,264)=2.02$, $P=0.03$ will be further explored in analyses separated by field of view, but appeared to be rooted in disproportionately poor performance by NVGPs on even relatively low numbers of squares during the wide field of view condition.

2.2.1.3. Average response. The expected main effect of number of squares, with average response increasing with increasing numerosity was observed, $F(11,264)=1398.3$, $P<0.001$. VGP status and number of squares interacted with both groups having similar average responses for low number of items, but NVGPs beginning to consistently underestimate the number of items before VGPs did, $F(11,264)=7.1$, $P<0.001$


A
▨ NVGP (N=13)
☐ VGP (N=13)
% Error
0
0.2
0.4
0.6
0.8
1
1 2 3 4 5 6 7 8 9 10 11 12
Number of Squares
*
**
B
▨ NVGP (N=13)
☐ VGP (N=13)
Correct Response - Average Response
−1
−0.5
0
0.5
1
1.5
2
2.5
1 2 3 4 5 6 7 8 9 10 11 12
Number of Squares
C
▨-- NVGP (N=13)
☐— VGP (N=13)
Reaction Time (secs)
0.5
1
1.5
2
2.5
3
3.5
1 2 3 4 5 6 7 8 9 10 11 12
Number of Squares

(Fig. 1(B)). Field of view also interacted with number of squares, as subjects began to consistently underestimate the number of items earlier in the wide field of view (around 8 items) than in the narrow field of view (around 10 items) $F(11,264) = 5.3$, $P < 0.001$. Finally, a VGP status × field of view × number of squares interaction prompted analyses separated by field of view discussed subsequently, $F(11,264) = 2.4$, $P = 0.007$.

2.2.1.4. RT. Despite the previously mentioned potential pitfalls with the RT data, the statistics were in harmony with the percent correct data. A main effect of number of squares was observed with subjects taking longer for greater numbers of squares, $F(11, 264) = 101.8$, $P < 0.001$. The main effect of VGP status on RT was non-significant ($P > 0.05$). However, a VGP status × number of squares interaction was observed, rooted in the fact that VGPs were faster than NVGPs for low number of items (1–2) but then became much slower for greater number of items, $F(11,264) = 3.3$, $P < 0.001$ (Fig. 1(C)).

A main effect of field of view was also observed with subjects taking longer in the narrow field of view condition than the wide field of view condition, narrow: 1.91 +/− .05 s, wide: 1.7 +/− .05 s, $F(1,24) = 14.7$, $P = 0.001$. The effect of field of view further interacted with number of squares, as RTs were relatively equivalent for low numbers of squares, but much longer for larger numbers of squares in the narrow field of view condition, $F(11,264) = 6.2$, $P < 0.001$. Finally, a VGP status × field of view × number of squares interaction $F(11,264) = 1.9$, $P < 0.05$, was caused by VGPs taking a disproportionately long time to respond in the narrow field of view condition when many squares were presented.

Because of the observed interactions outlined above, the two field of view conditions were further analyzed separately by field of view along the three components revealing field of view by video game status interactions (percent correct, average response, RT).

2.2.2. Narrow field of view analyses

2.2.2.1. Percent correct. Main effects of number of squares and VGP status were again seen, number of squares: $F(11,264) = 148.7$, $P < 0.001$; VGP status: $F(1,24) = 11.9$, $P = 0.002$, with accuracy decreasing with increasing numerosity and VGPs again outperforming NVGPs by a large margin, VGP: 78.2 +/− 1.9% correct, NVGP: 64.8 +/− 2.3% correct. A VGP status × number of squares interaction, $F(11,264) = 2.6$, $P = 0.004$, outlined identical effects as seen in the main ANOVA and is captured by the difference in

---

Fig. 1. (A) Enumeration accuracy—% correct and breakpoint: VGPs clearly outperform NVGPs on the enumeration task. The VGP breakpoint (the elbow in the regression line) is approximately two items beyond that seen in the NVGPs. Furthermore, the VGP advantage continues to hold even for high numerosity; (B) average response: NVGPs begin to underestimate the true number of items (bar above zero) about two items before VGPs. The relative inaccuracy of the NVGPs compared to VGPs can also be seen for all numbers of squares as the VGP bar does not begin to deviate from zero (on average perfectly correct) until around 10 items; (C) RTs: NVGPs and VGPs have very similar RTs for low numerosity, VGPs being slightly faster, but as the number of items to enumerate increases, the VGP RT becomes much greater than NVGPs. This is unlikely to represent a speed/accuracy trade-off in the conventional sense, in that longer RTs allow for more information to decay from visual memory. Instead, these results may indicate a more stable visual memory representation in the VGP population (Error bars denote SEM, $* = P < 0.05$, $** = P < 0.001$).

accuracy breakpoint, with both groups performing similarly well for low numbers of items (1–3 items $P > 0.07$) but VGPs significantly ($P < 0.05$) outperforming NVGPs for all subsequent numbers of items except for 10 squares ($P = 0.055$).

2.2.2.2. Average response. A main effect of number of squares was again observed, $F(11, 264) = 1358.2$, $P < 0.001$. A VGP status × number interaction closely followed the main ANOVA with both groups giving very similar average responses for low numbers of items, but with VGPs making better estimates for higher numbers, $F(11,264) = 4.2$, $P < 0.001$.

2.2.2.3. RT. A main effect of number of squares was observed $F(11,264) = 86.9$, $P < 0.001$. A VGP status × number interaction with VGPs responding faster for small numbers of squares, but becoming slower than NVGPs when greater numbers of squares were presented, $F(11, 264) = 2.8$, $P < 0.001$.

2.2.3. Wide field of view analyses

2.2.3.1. Percent correct. A main effect of number of squares was observed with performance as always decreasing with increasing numbers of squares, $F(11,264) = 135.3$, $P < 0.001$. A main effect of VGP status was again observed, VGP: 72.9 +/− 2.1% correct, NVGP: 60.0 +/− 2.7% correct, $F(1,24) = 5.9$, $P = 0.02$. And a VGP status × number of squares interaction again highlights the differences seen in accuracy breakpoint with the VGP advantage in accuracy only becoming evident for items above 3, $F(11,264) = 3.3$, $P < 0.001$.

2.2.3.2. Average response. Main effects of number of squares, $F(11,264) = 1069.4$, $P < 0.001$ and a VGP status × number of squares, $F(11,264) = 8.3$, $P < 0.001$ were observed.

2.2.3.3. RT. A main effect of number of squares was observed, $F(11,264) = 76.8$, $P < 0.001$. Also, as in the narrow field of view analysis, a VGP status × number of squares interaction, $F(11,264) = 3.4$, $P < 0.001$, reflected the fact that VGPs were faster than NVGPs for low numbers, but slower for higher numbers.

2.3. Discussion

The main finding of Section 2 is that VGPs enumerate more accurately than NVGPs. VGPs are able to enumerate with high accuracy about two items more than NVGPs, as exemplified by the reliable difference in accuracy breakpoint between the two populations. Beyond the accuracy breakpoint, the estimate of the number of objects presented remained more accurate in VGPs than for NVGPs. Both groups underestimated the true value at high numerosities, an expected result since subjects were limited in their response to up to 12 items (although see Mandler & Shebo, 1982, who also report systematic underestimation of numerosity beginning at set size of nine when the maximum number of items was 20). However, the results reveal an additional pattern of underestimation that differs between the two groups. At high numerosities (nine and above), NVGPs systematically underestimated the number of squares presented to a greater extent than VGPs, and were consistently faster than VGPs. While this result takes the form of a simple speed/accuracy trade-off, there are other possible mechanistic explanations for it, such as faster loss of information in memory in NVGPs than in VGPs. If the ability to count accurately from short term memory decays faster in NVGPs than in VGPs, NVGPs' responses would show greater underestimation of numerosity as well as faster RTs as the number of items they could enumerate is smaller. This point will be discussed further in Section 7.

Section 2 also reveals a systematic effect of the size of the field of view on performance at high numerosity. The wide field of view condition led to more errors, systematic underestimation and faster RTs than the narrow field of view condition suggesting an important role for the spread of attention in the enumeration of high numerosities. Importantly, display size did not affect performance at low numerosity, as exemplified by the comparable accuracy breakpoints found in the two field of view conditions. The different sensitivity of low and high numerosity displays to the field of view condition is further documented by the interactions between field of view and number of squares observed in percent correct, average response, and RT analyses. This finding lends support to the view that two distinct sets of constraints are underlying performance for low and high numerosity displays in the enumeration task. In this paradigm, field of view co-varies with density, with the wide field of view condition corresponding to lower density displays than the narrow field of view condition. The data pattern, however, runs counter to that predicted by models that suppose that canonical patterns or density may be at the source of the low/high numerosities split. Indeed, such models predict better performance for low density than high-density displays, a prediction which runs contrary to the finding of greater error in the wide condition and equal accuracy breakpoint across conditions.

Accuracy breakpoints and RTs for low numerosities were similar for VGPs and NVGPs across fields of view, indicating that the advantage conferred by video game playing applies equally to narrow and wide displays. Thus, the effects of game playing are not confined to peripheral locations, but are also visible more centrally. This is consistent with the fact that video game play, in addition to necessitating constant peripheral monitoring, also commonly requires the subject to actively attend to the center of the visual field, which normally contains the primary object of interest.

Section 2 demonstrates that expert video game players outperform non-players in the enumeration task by having both an extended accuracy breakpoint and greater accuracy in the counting range. It is unclear, however, whether the very act of playing is a causative factor in the improved enumeration performance of VGPs or whether selecting for video game players inherently biased us in selecting individuals with better visual skills. After all, individuals with good visual abilities probably have an advantage when it comes to playing video games, and thus may be more prone to become video game players than individuals with poor visual skills. This issue is addressed in Section 3.

3. Experiment 2

Results from Section 2 indicate enhanced enumeration performance in VGPs. While our hypothesis predicts that extensive video game playing leads to this enhanced skill, it could also be the case that VGPs have inherently better perceptual skills and/or were somehow genetically endowed with greater attentional abilities. Another explanation is that what is learned during video game play is not necessarily perceptual in nature, but is instead a perceptual-motor skill. Although the use of percent correct as our primary dependent measure in Section 2 was chosen to minimize the effect of visuo-motor facilitation in our measures, it is possible that by alleviating the demands of the motor response, video game playing allows VGPs to have more ‘left-over’ resources available to process the sensory stimulus.

To control for these two possibilities, NVGPs underwent video game training in Section 3. Some underwent video game training using an action video game, whereas others played a game that placed heavy demands on visuo-motor coordination but did not tap aspects of visual attention of interest in Section 2. If the differences observed in Section 2 are due to the demands of action video game playing and not due to better visuo-motor control or genetically endowed traits, a notable improvement in the enumeration task should be observed following training in the action game trainees, but not in the control game trainees.

Finally, while Section 2 employed only males as subjects, by using an equal number of males and females in Section 3, we can test whether the effects of video game play are similar across gender.

3.1. Method

3.1.1. Subjects

The study initially enrolled 20 NVGPs that were equally and randomly divided between the experimental and the control group (five males/five females in each group). The criteria for NVGP remained the same as in Section 2. All subjects underwent training as described below. One male from the experimental group and one male and one female from the control group did not finish training. Thus, five females and four males (mean age = 20.4, 8 right-handed) made up the final experimental group, while the final control group consisted of four females and four males (mean age = 19.7, all right handed).

3.1.2. Pre-test

As differences between the two fields of view were minimal with respect to the effect of video game play, and because of two additional experimental tasks, not relevant to this paper, subjects completed only the narrow field of view condition.

3.1.3. Apparatus

3.1.3.1. Testing. The apparatus was identical to that described in Section 2.

3.1.3.2. Training. The control group played on the same experimental setup (computer, monitor, refresh, and resolution) the experiment itself was conducted on. The action group played on one of two Dell computers each equipped with 20″ flat-panel LCD monitors.

3.1.4. Training stimuli and procedure

For both groups, training consisted of playing the pre-determined video game for one hour per day for ten out of fifteen days. The nine members of the experimental group played the game Medal of Honor: Allied Assault™ (henceforth referred to as the action video game). This game was chosen to be similar to those played by our VGPs from Section 2. It has a relatively simple interface, uses first-person point of view and requires effective monitoring of the entire monitor display (extent from fixation about 13°-height × 16°-width). Subjects played the game straight through for the first eight days, beginning each day at the point where they had finished the previous day. On days nine and ten they returned to the beginning of the game in order to quantitatively measure their improvement by comparing performance over mission one during their first (days 1–2) and last (days 9–10) playing. The eight members of the control group played the game Tetris™, which was displayed to cover the entire extent of the screen. However, because Tetris™ adds graphics on the side of the screen, the effective game area extended 13°-height × 9°-width from fixation. This game was selected to control for the effect of improved visuo-motor coordination, while placing little demand on the simultaneous processing of multiple items. Accordingly, the version of Tetris™ on which subjects were trained had the preview block option turned off. Furthermore, this group serves as a control for any possible effects due to familiarity with the task (test–retest). In a manner analogous to the action-trained group, improvement was quantitatively measured by comparing performance on day 1 versus that on day 10.

3.1.5. Post-test

After video game training, subjects were re-tested on the same experiment as in the pre-test, as well as the other two aforementioned unrelated tasks (Green & Bavelier, 2003).

3.2. Results

3.2.1. Game play

In order to assess game improvement, two different measures were taken for each group during the first and last playing. Improvement measures were determined by computing the difference between post-training performance and pre-training performance divided by pre-training performance.

For the action game, the two measures were shooting accuracy (number of targets hit/total number of shots fired) and the number of deaths before completing the first mission. In both measures, subjects showed improved performance following training. For shooting accuracy a 68% improvement was seen. In terms of the number of deaths to complete mission one, a 42% improvement was observed.

For the control game, the two measures were high score (highest score achieved in one level) and highest level reached. Again, all subjects improved following training (high score = 71% improvement, highest level reached = 67% improvement).

These results establish that both groups were engaged in their training and showed improvement on the training task.

3.2.2. Enumeration results

Similar analyses as in Section 2 were carried out. In the case of percent correct, average response, and RT each of the factors was initially analyzed in a 2(trained game: action/control)×2(test: pre/post)×12(number of squares) ANOVA and in the case of accuracy breakpoint 2(trained game: action/control)×2(test: pre/post) ANOVA. Gender was not included as a factor, as a preliminary analysis indicated there was no main effect of gender, nor were there any interactions with gender in any of the effects observed.

3.2.2.1. Accuracy breakpoint. A main effect of test was observed indicating a slight increase in accuracy breakpoint at post-test, pre: 3.0+/−0.15 squares, post: 3.8+/−0.31 squares, $F(1,15)=6.4$, $P=0.02$. Importantly for our hypothesis, a trained game×test interaction was also observed, caused by a greater post-test improvement in the action game than in the control game, $F(1,15)=6.1$, $P=0.03$. In analyses separated by trained game, a main effect of test was observed in only the action group, $F(1,8)=10.6$, $P=0.01$, not in the control group ($P>0.9$), indicating that training had a significant effect on accuracy breakpoint in the action game alone. No effect was observed on the slope of either component (all P's>0.2) (Fig. 2(A) and (B)).

3.2.2.2. Percent correct. A main effect of number of squares was observed as expected, $F(11,165)=111.6$, $P<0.001$. A main effect of test was also observed with subjects being more accurate post-test than pre-test, pre: 66.9+/−2.1% correct, post: 72.2+/−1.9% correct, $F(1,15)=23.0$, $P<0.001$. Importantly, an interaction was observed between trained game and test caused by the action group showing greater improvement than the control group, $F(1,15)=15.4$, $P<0.001$. Furthermore, significant interactions between test and number of squares, $F(11,165)=2.1$, $P=0.02$, and between trained game, test and number of squares, $F(11,165)=2.4$, $P=0.008$ reveal a similar pattern as in Section 2 (Fig. 2(A) and (B)). Following training, the action group was as accurate as the control group for low numbers of squares, but outperformed the control group on higher numbers.

3.2.2.3. Average response. A main effect of number of squares was observed, $F(11,165)=5473.0$, $P<0.001$, as well as the expected interaction between test and number of squares, $F(11,165)=4.2$, $P<0.001$, indicating more accurate average responses following training. Importantly, a trained game×test×number of squares, $F(11,165)=2.4$, $P=0.008$ revealed that the bulk of the changes in accuracy were in the action group. As was seen with the VGPs in Section 2, the trained group became better estimators of large numbers of items following training.

3.2.2.4. RT. The only effect was a main effect of number of squares, $F(11,165)=48.4$, $P<0.001$. This is unsurprising however because of the way we required subjects to respond.

3.3. Discussion

Section 2 establishes that even relatively little action video game play (10 h) is sufficient to alter enumeration performance. In fact, similar changes to those described in

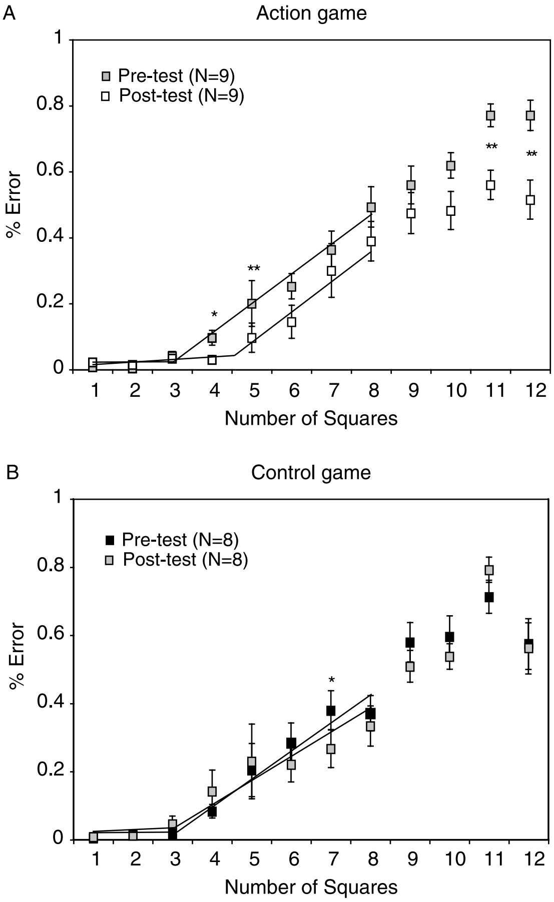


Fig. 2. (A) Enumeration accuracy—% correct and breakpoint-Action game: the action group demonstrates a clear increase both in breakpoint (by around 1.5 squares) and in overall accuracy; (B) enumeration accuracy—% correct and breakpoint-Control game: the control group showed no sign of improvement after training, thus ruling out test-retest improvements or improvements in visuo-motor control as possible hypotheses to explain the improvement in the action group (Error bars denote SEM, $* = P < 0.05$, $** = P < 0.001$).

Section 2, albeit of lesser amplitude, were observed in action game trainees for the main aspects of the task. These include a shift of accuracy breakpoint as well as greater accuracy and better estimation for numerosities above the accuracy breakpoint. This conclusively demonstrates that action video game play is at the source of the improved performance on the enumeration task. Furthermore, the fact that no effect of gender was observed, indicates that the consequences of video game play are not sex dependent.

Taken together, Sections 2 and 3 indicate that action video game play induces two main changes in performance on the enumeration task. First, measures of accuracy breakpoint show that gamers switch from the shallow to the steep component of the enumeration accuracy curve beyond the point where control subjects switch. Second, beyond the accuracy breakpoint, measures of percent correct and average response indicate greater accuracy in gamers than their controls.

When considering the underlying mechanism for these changes, it is important to recognize that the estimation of the subitizing range in enumeration studies has typically relied on RT measures rather than accuracy. In fact, the number of items that can be apprehended in a fast and parallel manner classically defines the subitizing range. Although accuracy measures have typically resulted in a similar pattern of results to RT measures in previous studies, we cannot rule out the possibility that the greater accuracy breakpoint in gamers may be due to more accurate counting rather than an increase in the number of objects that can be immediately apprehended. Whether gaming does modify the ability to apprehend more objects at once remains an open question that will be addressed in Section 4.

4. Experiment 3

A few caveats in the interpretation of Section 2 were addressed in Section 4. First, subjects gave vocal responses thus allowing precise measurement of RT. Second, afterimages may have played a role in the effects of gaming reported in Sections 2 and 3, as the stimuli display was not masked. Although Simon & Vaishnavi (1996) reported no substantial increase in the subitizing range when the items were presented as afterimages (thus allowing up to 60 s of viewing), because their primary dependent measure was accuracy rather than reaction time, caution dictated that in Section 4, a backwards pattern mask be employed, specifically designed to eliminate afterimages as a viable source of information that could be used unequally by the two groups.

4.1. Method

4.1.1. Subjects

Twenty-two males (none of which had participated in Section 2 or 3) with normal or corrected vision were again placed into one of two groups, VGP or NVGP, based upon their responses to a questionnaire, slightly modified from that used in Section 2. The questionnaire and criteria to be considered a VGP were altered slightly to allow a more accurate measure of the amount of time each subject spent playing specific types of video games. The questionnaire asked the number of hours (0, 0–1, 1–2, 3–5, 5–10, 10+) spent playing each of several types of video games (action, sports, fantasy, role playing, other) per week. To be considered a video game player, a subject needed a minimum of 5 h a week of action video game usage for the previous six months. Eleven right-handed males with a mean age of 19.1 years fell into this category.

The criterion to be considered a non-video game player was zero hours per week of action games. Eleven right-handed males with a mean age of 20.3 years fell into this category.

Written informed consent was obtained from each subject and each subject was paid $7.50 for each hour of participation.

4.1.2. Stimuli and procedure

The stimulus/procedure was identical to that described in the narrow field of view portion of Section 2 except for four changes. First, to rule out any role of afterimages, a backwards pattern mask was presented for 500 ms following the presentation of the squares. The mask consisted of a black and white checkerboard pattern of the same contrast/luminance as the stimulus, designed to eliminate afterimages as a potential source of information. The second change was that the presentation time was increased to 100 ms—a change necessitated by the addition of the mask. Third, the task was shortened by including only 1–10 squares. Finally, the most important change was the method of response. The voice onset time was used to measure RT. The trials were recorded and scored for accuracy offline. If the subject hesitated or changed their response (“thre…no…four”), the trial was omitted from RT and accuracy results (this circumstance occurred on less than 1% of all trials).

4.2. Results

The results were analyzed using the same four dependent measures (accuracy breakpoint, percent correct, average response, RT) as well as a new component-RT breakpoint, which was computed in the same manner as the accuracy breakpoint.

4.2.1. Accuracy breakpoint

There was a main effect of VGP status on the breakpoint for accuracy with VGPs again having a 2.5 item advantage over NVGPs, VGP: 5.0+/−0.3 squares, NVGP: 2.5+/−0.32 squares, $F(1,20)=33.2$, $P<0.001$ (Fig. 3(A)). Again, there was no significant difference in the slope of either component of the model (P’s >0.08). This result replicates those of Sections 2 and 3.

4.2.2. Percent correct

A main effect of increased errors with increasing number of squares was observed, $F(9, 180)=119.1$, $P<0.001$. More importantly, there was a main effect of VGP status, VGP: 77.4+/−2.5% correct, NVGP: 60.5+/−3.2% correct, $F(1,20)=27.8$, $P<0.001$ and a VGP status × number of squares interaction, $F(9,180)=5.0$, $P<0.001$ (Fig. 3(A)). As seen in Section 2, this is due to the two groups being roughly equivalent at small numbers of squares ($P>0.05$ for two and three squares), but diverging for larger numbers with VGPs retaining high accuracy for even large numbers of squares (VGPs more accurate than NVGPs, $P<0.05$ for 4–9 squares).

4.2.3. Average response

There was only a main effect of number of squares, $F(9,180)=1792.6$, $P<0.001$. There was no main effect of VGP status or any interactions with VGP status, indicating that on average the two groups performed quite similarly. This difference with the previous experiments is not surprising as Section 4 included up to 10 squares while the

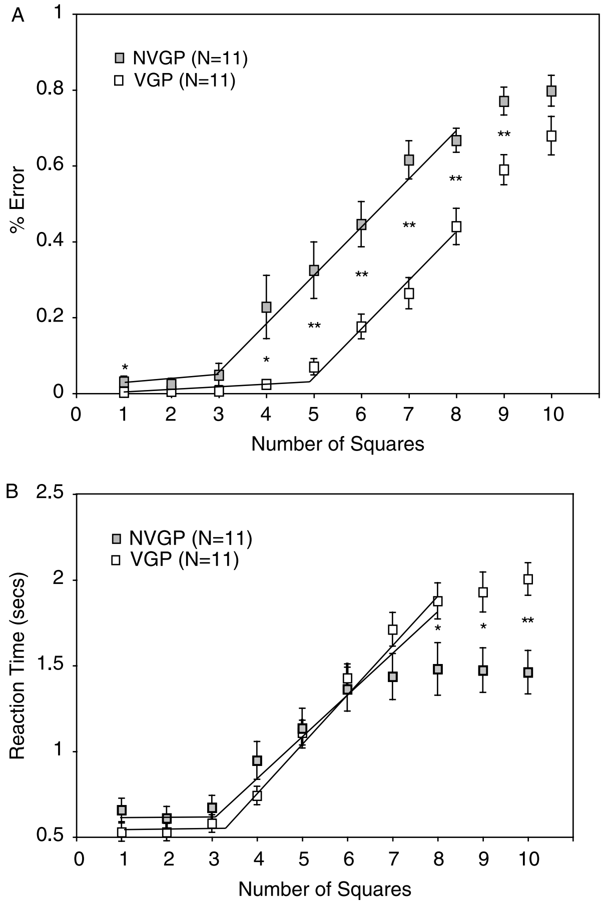


Fig. 3. (A) Enumeration accuracy—% correct and breakpoint: As was seen in Section 2, VGPs show a clear advantage both in breakpoint as well as overall accuracy; (B) RT: also as was seen in Section 2, VGPs and NVGPs have extremely similar RTs for low numbers of items, but VGPs begin to take significantly longer than NVGPs as the number of items to be enumerated increases. The RT breakpoint, which offers an index of the number of items that can be instantaneously apprehended, is similar in both groups (Error bars denote SEM, $* = P < 0.05$, $** = P < 0.001$).

significant difference in average response were most marked at very high numerosity (11–12 squares) in the two previous experiments.

4.2.4. RT

An ANOVA on the simple RT scores revealed a main effect of number of squares, $F(9, 180) = 155.5$, $P < 0.001$ and a VGP status × number of squares interaction, $F(9,180) = 12.0$, $P < 0.001$ with the VGPs again starting slightly (although not significantly) faster for low numbers of squares but becoming significantly slower ($P < 0.05$) than NVGPs for high numbers of squares (8–10) (Fig. 3(B)).

4.2.5. RT breakpoint

In contrast to the accuracy breakpoint, no main effect of VGP status was found for RT breakpoint, VGPs: 3.4+/− .1 items, NVGPs: 3.0+/−.2 items, $F(1,20)=2.8$, $P=0.1$. Although the trend appears in the same direction as the accuracy breakpoint, the size in terms of number of squares is certainly much smaller than has been observed in accuracy (Fig. 3(B)). No effect of VGP status was seen in the slope of either component (P's$\,>0.5$).

4.3. Discussion

The main finding of Section 4 is that there is a difference in accuracy breakpoint between gamers and non-gamers despite similar RT breakpoints. Clearly the accuracy breakpoint of the VGP population does not reflect the 'true' number of items that can be immediately apprehended. Instead, it is apparent that despite near perfect accuracy for up to five items in the VGP group, the RTs for four and five squares are slower than that for one to three items. This finding provides the first large dissociation of RT and accuracy on an enumeration task. It has generally been taken for granted that when subjects begin to 'count', they begin to lose information (Sperling, 1960). The more items that are presented, the longer subjects need to count, and the less accurate they become. However, VGPs are apparently counting for four and five dots (as demonstrated by the RT data), while their accuracy continues to be nearly perfect.

This enhancement in counting ability is also seen in the large differences in accuracy and RT between VGPs and NVGPs for greater numbers of items. These results suggest that the VGPs continue to successfully 'read-off' the items from their visual memory after the NVGPs have made their best attempt. The correlation between enumeration performance and working memory ability has been previously demonstrated in the literature (Tuholski, Engle, & Baylis, 2001). Subjects that performed poorly on a working memory task also performed poorly on the counting portion of the enumeration paradigm while no differences were observed in the subitizing span. It therefore seems most likely that what is truly being altered by video game play is at the level of counting, rather than at the level that mediates immediate apprehension of numerosity.

While it is clear that the capacity of the immediate apprehension mechanism (or in other words the subitizing range defined by the RT breakpoint) in the enumeration task is similar in VGPs and NVGPs, what is left open is whether VGPs can simultaneously track more items than NVGPs. While some researchers have suggested that there may be a connection between the number of items that can be automatically enumerated and the number of items that can be tracked, it is not necessary that they be rooted in the same mechanism. To more directly address this question we made use of the multiple object tracking (MOT) paradigm (Pylyshyn, 1989) which requires subjects to dynamically allocate attention to multiple objects and sustain that attention for several seconds.

5. Experiment 4

In our version of the MOT paradigm, subjects view a number of randomly moving circles. At the beginning of the trial, some subset of the circles is cued. The cues then disappear and subjects are required to keep track of the circles that were cued (now visually indistinguishable from uncued circles) as they continue to move randomly about the screen. After several seconds of tracking, one of the circles is highlighted and the subject must make a yes (was cued)/no (was not cued) decision. This method of response, rather than the more typical method of asking the subject to indicate each of the initially cued objects, was employed to minimize the role of working memory in the response process, and in doing so gain a cleaner measure of the number of items that can be successfully tracked.

While previous theories have suggested a preattentive link between subitizing and MOT performance (Pylyshyn, 1989), it is generally accepted that there is a large dynamic attentional component to the MOT task as well (Scholl, Pylyshyn, & Feldman, 2001). The task requires active allocation of visual attention in order to successfully track targets embedded in a field of competing, and visually identical, distracting elements. Several studies have demonstrated that attention is actually split between the items during tracking (Sears & Pylyshyn, 2000). Furthermore, neuroimaging has revealed activation in what are thought of as attentional areas-parietal and frontal regions-when subjects perform a MOT task (Culham, Brandt, Cavanagh, Kanwisher, Dale and Tootell, 1998; Culham, Cavanagh, & Kanwisher, 2001).

While the subitizing span offers a glimpse at the number of items that can be immediately apprehended, the MOT paradigm offers a good measure of the number of items than can be simultaneously tracked, and some have suggested, simultaneously attended, over a period of time. We therefore decided to use the MOT task to clarify whether VGPs can actually track more items at once as well as examine what, if any, relationship exists between multiple object tracking and enumeration performance.

5.1. Method

5.1.1. Participants

Twenty males (none of whom had participated in previous experiments reported in this paper) with normal or corrected vision were again placed into one of two groups, VGP or NVGP, in a manner identical to that used in Section 4. Ten right-handed males with a mean age of 19.4 years were placed into the VGP category, while ten right-handed males with a mean age of 20.6 years were placed into the NVGP category. None of the participants were red/green colorblind.

Written informed consent was obtained from each subject and each subject was paid $7.50 for each hour of participation.

5.1.2. Stimuli and procedure

Each observer viewed the display binocularly with his head positioned in a chin rest at a test distance of 57 cm. Subjects were instructed to fixate within a center ring (radius = 0.25°). Subjects pressed a key to begin each trial. Each trial began with 16 circles (radius 0.5°) moving randomly on a circular gray background (radius of circular background = 10°) at a rate of 5°/s. The circles repelled one another before contact (0.5° minimum separation), and were repelled by the outer edges of the background and by the center fixation circle. At the start of the trial, some number (1–7) of these circles were cued (colored red) while the rest were green. During this time the subject was instructed to attend to the red circles, as they would shortly change to green after which time the subject had to track the circles that were previously cued. After 2 s the cued circles changed to green, leaving all 16 circles visually indistinguishable. The subject had to track the cued circles for 5 s after which one of the 16 circles turned white (probe). The subject had to press either a yes or no key in response to whether the probe circle was one of the originally cued circles. The probe circle was one of the originally cued circles 50% of the time. Each number of cued circles (1–7) was presented 20 times (10 yes, 10 no) for a total of 140 trials. In many other instantiations of this paradigm subjects were not eye-tracked (Pylyshyn, 2004), or eye-movements were found to have few implications (Pylyshyn & Storm, 1988). However, because of the use of more objects than most MOT paradigms (up to 7) and the possibility that the two groups could differ in eye movement strategies, subjects were eye-tracked and trials where they made an eye-movement greater than 1° from center were omitted from later analyses. This happened fairly rarely (around 6% of trials) and did not differ between groups nor did the occurrence of eye-movements appear to affect accuracy or affect accuracy differently between the two groups ($P > 0.4$).

5.2. Results

The results were analyzed in a 2(VGP status: VGP/NVGP)×7 (number of circles to track) ANOVA as earlier analyses had indicated there was no effect ($P > 0.7$) of whether the answer was yes or no (i.e. no response bias). As expected, a main effect of number of circles to track was found with accuracy decreasing with increasing number of circles, $F(6, 108) = 60.6$, $P < 0.001$. Importantly, a main effect of VGP status, VGP: 84.3+/−1.6% correct, NVGP: 78.2+/−1.8% correct, $F(1,18) = 9.2$, $P = 0.007$ was observed indicating better performance in VGPs (Fig. 4). As VGP status and number of circles to track did not

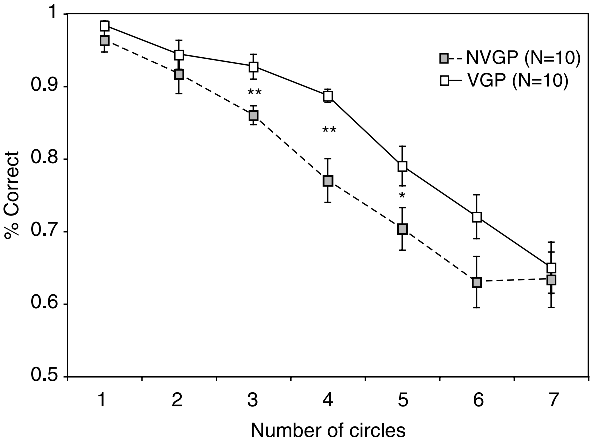


Fig. 4. Multiple object tracking performance: VGPs demonstrate a substantial increase in the accuracy with which multiple items can be tracked compared to NVGPs. The effect is most pronounced for 3–5 items to track (Error bars denote SEM, $* = P < 0.05$, $** = P < 0.001$).

interact, we can assume the VGP advantage was relatively equal across number of circles. Although there was no interaction of VGP status with the number of circles to track, individual analyses separated by number of circles to track were performed for comparison with following analyses and indicated a significant advantage ($P < 0.05$) for VGPs only for three to five circles to track and a marginally significant advantage at 6 circles to track ($P = 0.07$). At the ends of the spectrum (one circle to track or seven circles to track) the two groups were equivalent (P's > 0.1).

5.3. Discussion

The results indicate that VGPs outperform NVGPs when it comes to tracking several objects over time. Unlike for enumeration accuracy and RT, there was no significant interaction between VGP status and number of items on MOT accuracy. However, the results of the two paradigms do appear qualitatively similar, with VGPs and NVGPs having comparable performance with relatively few items, the differences only emerging after some critical threshold in load is exceeded. The data is furthermore quite consistent with the recent findings of Trick and colleagues (Trick, Jaspers-Fayer, & Sethi, 2005) who reported that children (6–19 years old) who played action video games performed significantly better than non-action game playing children on a version of the multiple object tracking task. Before drawing further conclusions, Section 6 investigates causation through a controlled training study.

6. Experiment 5

As in Section 3, it is critical to show a causative effect of video game play on MOT performance. In this experiment, a larger sample of NVGPs than in Section 3 was trained for 30 h, three times as long as the training in Section 3. The choice of training time was determined by other tasks not reported in this paper.

6.1. Method

6.1.1. Subjects

The study enrolled 32 NVGPs that were equally and randomly divided between the experimental and the control group. The criteria for NVGP remained the same as in Section 5. All subjects underwent training as described below. In all eight females and eight males (mean age = 21.3, all right-handed) made up the final experimental group, while the final control group consisted of nine females and seven males (mean age = 21.0, 15 right handed).

6.1.2. Apparatus

6.1.2.1. Testing. The apparatus was identical to that described in Section 5.

6.1.2.2. Training. Both groups played their respective games on 20″ monitors. The action game group played on Dell FlatPanel displays, whereas the control group played on CRT monitors.

6.1.3. Training stimuli and procedure

For both groups, training consisted of playing the pre-determined video game for a total of 30 h (maximum of 2 h per day, minimum of 5 h per week, maximum of 8 h per week). The sixteen members of the experimental group played the game Unreal Tournament 2004 (henceforth referred to as the action video game), a different action game than previously used. This game was chosen to be similar to those played by our VGPs; it has a relatively simple interface, uses first-person point of view and requires effective monitoring of the entire visual field (extent from fixation about 13°-height × 16°-width). One source of confounding in the previous game used was the fact that players could learn the development of the story and develop efficient 'wait and ambush' strategies. Unreal Tournament 2004 was chosen because there is no 'script'. Instead, the game is controlled by the action of 32 AI agents rather than linear story development. Each hour session of the action game was divided into three 20-minute blocks. The difficulty of each block was adjusted based upon the kill/death ratio. If in a block the player scored more than twice as many kills than they had deaths, the difficulty level was increased one level. Also, players were periodically re-tested on lower difficulty levels to quantitatively assess improvement.

The sixteen members of the control group played the game Tetris™, which was displayed to cover the entire extent of the screen. As such, the field of view of the Tetris™ game was actually slightly larger than that of the action game (which was the same as in Section 3-13° × 16°). The effective control game area extended 18°-height × 13°-width from fixation. This game was selected to control for the effect of improved visuo-motor coordination, while putting little demands on the processing of multiple objects at once. Accordingly, the version of Tetris™ on which subjects were trained had the preview block option turned off. In a manner analogous to the action-trained group, improvement was quantitatively measured by comparing performance on Day 1 versus that on Day 30.

6.1.4. Post-test

After video game training, subjects were re-tested on the same experiment as in the pre-test, as well as the other aforementioned unrelated tasks.

6.2. Results

6.2.1. Game play

In order to assess game improvement, several measures (slightly different than those collected in Section 3) were used. However, as in Section 3, a percent change score was calculated for each of the measures.

For the action game, the two measures used were kills and deaths. For each of five levels of game difficulty (level five being the highest level that all players attained) the measure taken on their first playing of the level (which because of the way in which difficulty was progressed was not necessarily on the first day of training) was compared with their final playing of the level on Days 29–30. A substantial increase in number of kills, and decrease in number of deaths was seen at each difficulty level (Table 1).

For the control game, the average and median scores from Day 1 were compared with the same values on Day 30. As in the action game, the control players showed substantial

Table 1
Action game improvement-Section 6: at each difficultly level, the action game trainees in Section 6 improved in both measures of performance from their first experience on a given difficult level to their final day

% Change		
Difficulty	Kills	Deaths
1	226.3	−64.1
2	147.6	−38.0
3	160.3	−27.4
4	79.8	−32.5
5	52.0	−32.3

improvements after training, the mean score improving by 323% and the median score by 359%.

As in Section 3, these results demonstrate that both groups were engaged in their training and showed improvement on the training task.

6.2.2. MOT performance

As in Section 3, no effect of gender was found in preliminary analyses, and as such, a 2(trained game: action/control) × 2(test: pre/post) × 7(number of circles to track) ANOVA was run on accuracy. Again, only the trials where subjects did not break fixation were used (as in Section 5, eye-movements occurred in only approximately 5% of trials). Also, as previously observed, there was no effect of trained game or test on the number of eye movements ($P > 0.6$) and again eye-movements did not appear to affect accuracy.

The expected main effect of number of circles was found, $F(6,180) = 155.6$, $P < 0.001$ with accuracy decreasing with increasing numbers of circles to track. Also a main effect of test was found, pre: 75.3 +/− 1.2% correct, post: 78.1 +/− 1.1% correct, $F(1,30) = 5.1$, $P = 0.03$. Most importantly for our hypothesis, however, an interaction between trained game and test was found, $F(1,30) = 13.4$, $P = 0.001$, indicating an unequal effect of training with the action group improving approximately 7.5% whereas the control group remained stable (Fig. 5(A) and (B)).

Separating the groups, and running a 2(test:pre/post) × 7(number of circles) ANOVA revealed that only the action group showed a main effect of test, $F(1,15) = 15.5$, $P = 0.001$ as well as a test × number of circles interaction, $F(6,90) = 2.3$, $P = 0.04$. No effect of test or any interactions with test were found in the control group.

6.3. Discussion

Section 4 demonstrates that relatively little video game play leads to substantial differences between groups and further demonstrates that the effects observed in Section 5 were not due to an inherent population bias.

7. General discussion

The five experiments presented demonstrate that action video game play increases the number items that can be enumerated and tracked simultaneously over time. In Section 2,

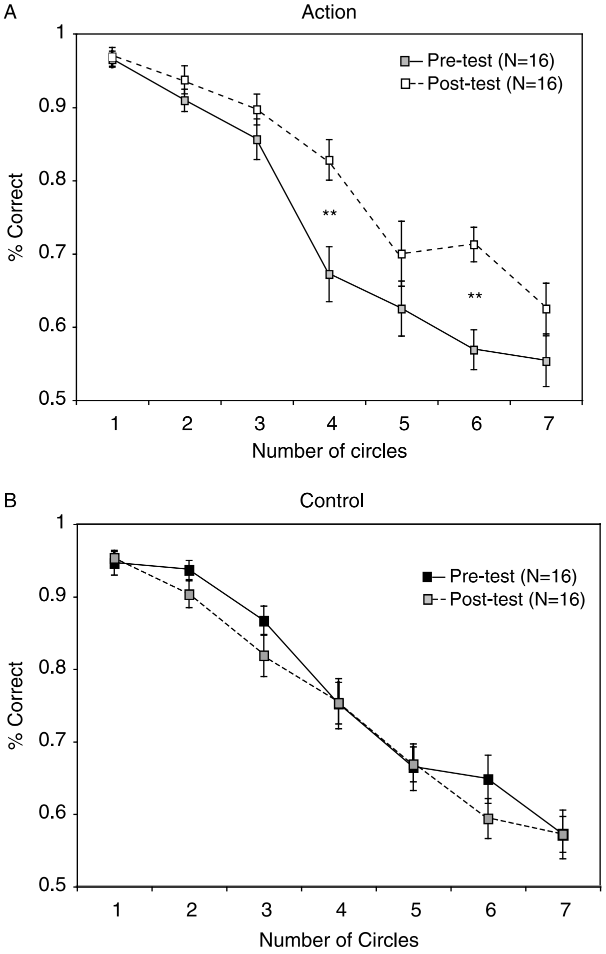


Fig. 5. (A) Multiple object tracking performance-Action game: those subjects trained for 30 h on an action video game show a marked improvement in their MOT performance as compared to their pre-test scores; (B) multiple object tracking performance-control game: performance was identical before and after training in the control group. These results rule out test-retest as a source of confounding as well as increases in visuo-motor coordination (Error bars denote SEM, $* = P < 0.05$, $** = P < 0.001$).

habitual action video game players display enhanced enumeration accuracy as compared to non-players. Section 3 establishes a causal role for action video game play, as NVGPs specifically trained on an action video game show similar enhancements. Section 4 demonstrates for the first time a dissociation between accuracy and reaction time measures of the subitizing range and establishes that action game playing does not modify the number of items that can be immediately apprehended, but rather enhances accurate counting. By making use of the multiple object tracking paradigm, Section 5 demonstrates an effect of VGP status on the ability to simultaneously track multiple objects over an extended period of time. The significant improvement in MOT performance seen in NVGPs after action game training in Section 6 demonstrates that action video game playing has a causal role in the measured effects.

Taken together, these five experiments suggest that action video game play may enhance some aspects of visual working memory. Several lines of evidence point to this conclusion. First, VGPs demonstrate enhanced enumeration accuracy even at very high numerosities. Second, this enhanced accuracy is accompanied by an increase in RTs. Although this pattern would be the expected speed/accuracy trade-off in a system in which evidence accumulates over time but does not decay, the behavior under study relies on visual short-term memory in which representations are known to decrease in fidelity over time (Lee & Harris, 1996; Nilsson & Nelson, 1981; Sperling, 1960; Vogels & Orban, 1986). Delaying responses in studies of short-term memory does not lead to increased accuracy, but rather decreased accuracy as the memory representations have more time to fade. Thus, an alternative explanation seems warranted in which video game experience leads to enhancements in some aspect(s) of visual short-term memory. At least two alternatives are possible, one based on the durability of the memory trace and another on the speed of cycling through the memory trace. In the first case, action video game experience may lead to a more durable visual memory trace. This view would be consistent with the accuracy and reaction time data as well as the average response data where NVGPs begin to underestimate the number of squares well before the VGPs. One may speculate that after a certain period of time, NVGPs begin to 'drop' items from visual memory, and at this point they simply make their best guess (from viewing Figs. 1(C) and 3(B), one can surmise that the NVGPs RTs appear to plateau at around eight items). Conversely, if it were the case that VGPs possess a more durable memory trace, they would be able to continue counting beyond the point where the NVGPs have stopped, which would account for both the greater accuracy and longer RTs. In addition, this process may also sustain better tracking ability in the MOT by allowing more durable indexing of the dynamics of the objects to be tracked. However, a change in the fidelity of working memory representations in gamers is only one possible explanation for the observed results. A possible alternative hypothesizes that items in working memory are not necessarily kept simultaneously active, but instead one or a few items are constantly refreshed by a visit from a single focus of attention that moves from item to item in a cyclical fashion. As the speed of cycling through the items increases, the number of items that could be successfully maintained in short term memory would correspondingly increase. It is therefore possible to capture the present findings by assuming that the speed of cycling through memory traces is faster in VGPs than NVGPs, thus accounting for both the better counting and multiple object tracking performance. It should be further noted that factors unrelated to visual short-term memory, such as estimation ability and response bias, may also be at work in the enumeration paradigm, particularly for high numerosities. For instance, VGPs may be better able to judge when 'the most' number of squares were presented, without necessarily being able to explicitly count each item. Also, as previously mentioned, because the maximum response was capped at some maximum value, a bias toward underestimation for the larger numerosities is created that may not be exactly equal in the two populations. A role for these differences in estimation/bias cannot be ruled out in interpreting some of the current results, especially at high numerosities, but they remain an unlikely explanation at lower numerosities where the accuracy breakpoint is seen to shift between NVGPs and VGPs.

Beyond the effect of action video game play, these findings also lend strong support to models of enumeration performance that propose relatively distinct constraints for the two components of the enumeration performance curve. The dissociation between the accuracy and RT breakpoints in gamers is probably the most robust indication to-date of separate mechanisms, one that is sensitive to gaming (counting) and one that is not (subitizing). Similarly, the comparison of two different fields of view in Section 2 indicates that the mechanism(s) behind subitizing are less malleable that those behind counting. Indeed, in Section 2, performance over the subitizing range was quite similar across visual field conditions. Only in the counting range did performance differ with more accurate performance for the smaller field of view and denser displays. Models that suppose fundamental differences in the characteristics of the display (density, patterns, etc.) between low and high numerosity stimuli cannot readily account for the overall pattern of results reported here, be it the effect of gaming or that of visual field size. Models of enumeration studies which posit two separate mechanisms—a fast and parallel one for subitizing and a more serial process for counting-more naturally capture the main findings. Under this view, the mechanism underlying subitizing would show little to no sensitivity to gaming or visual field size, and be highly specific to the enumeration of low numerosities. In contrast, the mechanism underlying counting would be much more plastic, showing enhancement with gaming and be facilitated by the use of a small visual field.

Although some have suggested a link between immediate apprehension in the enumeration task and performance on the MOT paradigm (for instance, that they may both utilize preattentive mechanisms, or FINSTs—Pylyshyn, 1989; Trick & Pylyshyn, 1994), our results suggest that the subitizing range does not index the same process as the MOT. VGPs demonstrate no enhancement in subitizing range as measured by RT, but do demonstrate an enhancement in MOT ability. Also, while there is virtually no cost in terms of speed or accuracy moving from one to three items in the enumeration paradigm, a clear decrease in accuracy is observed with each additional item in the MOT paradigm (even moving from one to two items). Thus, it appears that the number of items that can be immediately apprehended as measured by RT measures in enumeration studies is not necessarily a good predictor of the number of items than can be simultaneously tracked. Although our data do not allow us to draw strong conclusions, our findings suggest that the number of items that can be accurately counted may be a better correlate of tracking capabilities, as both of these measures are found to improve with gaming.

This study establishes that when it comes to the number of objects that can be attended, a distinction should be drawn between a fast, parallel behavior that displays little plasticity and a more serial behavior that displays a range of plastic behaviors. As such these studies make several contributions, both to our understanding of the processes indexed by the enumeration and MOT paradigms, as well as to our understanding of the nature of the changes that occur as a result of action video game play. It will be, however, for future experiments to fully characterize the consequences of these results for models of attention and working memory.

Acknowledgements

We thank J. Cohen, N. Fernandez, D. McColgin, and K. Schneider for help with subjects, data analysis, and software support, M. Dye for comments on earlier drafts, and S. Dehaene for discussion. This research was supported by NIH grants EY O16880 and DC 04418 to D.B.

References

Akin, O., & Chase, W. (1978). Quantification of three-dimensional structures. *Journal of Experimental Psychology: Human Perception and Performance*, *4*, 397–410.

Atkinson, J., Campbell, F., & Francis, M. (1976). The magic number 4+/−0: A new look at visual numerosity judgments. *Perception*, *5*, 327–334.

Culham, J. C., Brandt, S., Cavanagh, P., Kanwisher, N. G., Dale, A. M., & Tootell, R. B. (1998). Cortical fMRI activation produced by attentive tracking of moving targets. *Journal of Neurophysiology*, *80*, 2657–2670.

Culham, J. C., Cavanagh, P., & Kanwisher, N. G. (2001). Attention response functions: Characterizing brain areas using fMRI activation during parametric variations of attentional load. *Neuron*, *32*(4), 737–745.

Dorval, M., & Pepin, M. (1986). Effect of playing a video game on a measure of spatial visualization. *Perceptual and Motor Skills*, *62*, 159–162.

Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition*, *44*, 43–74.

Gopher, D. (1992). The skill of attentional control: Acquisition and execution of attention strategies. In D. E. Meyer, & S. Kornblum (Eds.), *Attention and performance XIV* (pp. 299–322). Cambridge, MA: MIT Press.

Gopher, D., Weil, M., & Bareket, T. (1994). Transfer of skill from a computer game trainer to flight. *Human Factors*, *36*(3), 387–405.

Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature*, *423*, 534–537.

Greenfield, P. M., DeWinstanley, P., Kilpatrick, H., & Kaye, D. (1994). Action video games and informal education: Effects on strategies for dividing visual attention. *Journal of Applied Developmental Psychology*, *15*, 105–123.

Kaufman, E., Lord, M., Reese, T., & Volkman, J. (1949). The discrimination of visual number. *American Journal of Psychology*, *62*, 498–525.

Krueger, L. E. (1982). Single judgements of numerosity. *Perception and Psychophysics*, *31*, 175–182.

Lee, B., & Harris, J. (1996). Contrast transfer characteristics of visual short-term memory. *Vision Research*, *36*, 2159–2166.

Li, X., & Atkins, M. S. (2004). Early childhood computer experience and cognitive and motor development. *Pediatrics*, *113*(6), 1715–1722.

Mandler, G., & Shebo, B. J. (1982). Subitizing: An analysis of its component processes. *Journal of Experimental Psychology: General*, *111*(1), 1–22.

McClurg, P. A., & Chaille, C. (1987). Computer games: Environments for developing spatial cognition. *Journal of Educational Computing Research*, *3*(1), 95–111.

Nilsson, T. H., & Nelson, T. M. (1981). Delayed monochromatic hue matches indicate characteristics of visual memory. *Journal of Experimental Psychology: Human Perception and Performance*, *7*, 141–150.

Orosy-Fildes, C., & Allan, R. W. (1989). Psychology of computer use: XII. Videogame play: Human RT to visual stimuli. *Perceptual and Motor Skills*, *69*, 243–247.

Oyama, T., Kikuchi, T., & Ichihara, S. (1981). Span of attention, backward masking and RT. *Perception and Psychophysics*, *29*, 106–112.

Pylyshyn, Z. W. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition*, *32*, 65–97.

Pylyshyn, Z. W. (2004). Some puzzling findings in multiple object tracking (MOT): I. Tracking without keeping track of object identities. *Visual Cognition*, *11*(7), 801–822.

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, *3*(3), 179–197.

Scholl, B. J., Pylyshyn, Z. W., & Feldman, J. (2001). What is a visual object? Evidence from target merging in multiple object tracking. *Cognition*, *80*, 159–177.

Sears, C. R., & Pylyshyn, Z. W. (2000). Multiple object tracking and attentional processing. *Canadian Journal of Experimental Psychology*, *54*, 1–14.

Simon, T. J., & Vaishnavi, S. (1996). Subitizing and counting depend on different attentional mechanisms: Evidence from visual enumeration in afterimages. *Perception and Psychophysics*, *58*(6), 915–926.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs*, *74*, 1–29.

Subrahmanyam, K., & Greenfield, P. M. (1994). Effect of video game practice on spatial skills in girls and boys. *Journal of Applied Developmental Psychology*, *15*, 13–32.

Trick, L. M., Jaspers-Fayer, F., & Sethi, N. (2005). Multiple-object tracking in children: The 'Catch the Spies' task. *Cognitive Development*, *20*, 373–387.

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 and Performance*, *19*(2), 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*(1), 80–102.

Tuholski, S. W., Engle, R. W., & Baylis, G. C. (2001). Individual differences in working memory capacity and enumeration. *Memory and Cognition*, *29*(3), 484–492.

van Oeffelen, M. P., & Vos, P. G. (1982). Configurational effects on the enumeration of dots: Counting by groups. *Memory and Cognition*, *10*(4), 396–404.

Vogels, R., & Orban, G. A. (1986). Decision processes in visual discrimination of line orientation. *Journal of Experimental Psychology: Human Perception and Performance*, *12*, 115–132.

Yuji, H. (1996). Computer games and information processing skills. *Perceptual and Motor Skills*, *83*, 643–647.