Visuospatial experience modulates attentional capture: Evidence from action video game players

Greg L. West
Department of Psychology, University of Toronto, Canada
Sara A. Stevens
Department of Psychology, University of Toronto, Canada
Carson Pun
Department of Psychology, University of Toronto, Canada
Jay Pratt
Department of Psychology, University of Toronto, Canada

Visuospatial experience, the prolonged engagement in a demanding visual task, obtained through action video game play enhances several visual and cognitive processes. The underlying mechanisms involved in these processes, however, remain unclear. Here we demonstrate that experience with action video games modulates early sensory processing, resulting in increases sensitivity to salient visual events that capture attention. In two experiments, we show using a Temporal Order Judgment (TOJ) and a Signal Detection Paradigm (d') that action video game players show greater sensitivity to exogenous sensory events in the visual array. These results suggest that visuospatial experience modulates the earliest sensory aspects of visual processing.

Keywords: plasticity, spatial vision, attention


Introduction

Our ability to efficiently detect change in the environment is critical to our survival. Because of this, it is thought that attentional mechanisms have evolved in such a way that resources are automatically oriented to abrupt and/or salient changes in the visual field (Irwin, Colcombe, Kramer, & Hahn, 2000; Jonides & Yantis, 1988; Theeuwes, 1995; Yantis, 1993); a process often referred to as attentional capture. Interestingly, little is known about how long term experience might affect this well studied process. Some forms of experience have been shown to affect attentional capture. For example, differences in magnitude of attentional capture have been observed between young and elderly individuals (Kramer, Hahn, Irwin, & Theeuwes, 2000; Pratt & Bellomo, 1999), and previous experience with specific stimulus properties (e.g., color) have been shown to contingently capture attention (Folk, Remington, & Johnston, 1992; Folk, Remington, & Wright, 1994). The effect of long term visuospatial experience on attentional capture, however, remains unclear at this point. Individuals that habitually play action video games engage in prolonged and demanding visuospatial activity, and therefore are an excellent population to use in the investigation of long term experience on attentional capture.

Recently, prolonged experience playing action video games have been shown to affect several domains of attentional performance. For example, action video game players (VGPs) have been show to have faster response times (RTs) when detecting visual targets (Castel, Pratt, & Drummond, 2005; Yuji, 1996), and have a larger useful field of view (UFOV) suitable for accurate target detection (Feng, Spence, & Pratt, 2007; Green & Bavelier, 2003), when compared to non-video game players (NVGPs). More recently, Green and Bavelier (2007) demonstrated that VGPs could accurately identify a target surrounded by distractors at a lower target-distractor distance threshold; suggesting that VGPs have enhanced spatial resolution of visual processing and highlighting the potential for action video game use in a rehabilitation setting. Importantly, it has been consistently demonstrated that observed differences between VGPs and NVGPs are due to experience with action games and not due to self-selection (e.g., Feng et al., 2007; Green & Bavelier, 2003, 2006, 2007).

It is still unclear at this point, however, what underlying sensory and attentional mechanisms are affected through this obtained experience. This is because differential attentional effects are not always observed between VGPs and NVGPs. For example, no difference in the magnitude of inhibition of return (IOR), a later attentional process where target detection is impaired at spatial locations that were recently attended (Posner & Cohen, 1984), were found between VGPs and NVGPs (Castel et al., 2005). Notably, VGPs still had faster RTs when detecting targets that were not recently attended, suggesting that VGPs...
might more efficiently process information in the visual environment through faster stimulus-response mapping. The lack of a differential effect between VGPs and NVGPs in the observed magnitude of IOR also suggests, however, that action video game playing does not necessarily affect later stages of attentional deployment, but instead might primarily affect earlier sensory and attentional processes; possibly resulting in increased sensitivity to salient changes in the visual array that capture attention. This raises the question of whether early sensory and attentional processes are modulated through long term experience with visuospatially demanding tasks, such as an action video game.

The present study seeks to examine how habitual action video game playing affects initial attentional deployment in situations where attentional capture occurs. If salient sensory events show a greater ability to capture the attention of VGPs, it is expected that increased sensitivity to these changes in the visual array will be observed when compared to NVGPs. We employed two different paradigms to measure the magnitude of attentional capture in VGPs and NVGPs. In Experiment 1, we used a Temporal Order Judgment (TOJ) paradigm that measured participants’ sensitivity to an exogenous peripheral cue. In Experiment 2, we used a demanding signal detection paradigm to measure participant’s sensitivity to the presence of a presented target with properties known to capture attention (i.e., an abrupt change in motion; Abrams & Christ, 2003). Importantly, tasks were unspeeded to avoid conflating known speed-accuracy trade-off effects with initial signal detection that would otherwise be present in a speeded RT based task.

### Experiment 1

In Experiment 1 we used a TOJ paradigm to assess whether VGPs show greater attentional capture to an exogenous peripheral cue, as compared to NVGPs. Abrupt visual onsets, such as a peripheral cue, have been shown to involuntarily capture attention (e.g., Remington, Johnston, & Yantis, 1992; Yantis, 1993). In the classic TOJ task an exogenous cue is used to shift spatial attention to a particular location before the onset of two target items separated by some variable stimulus onset asynchrony (SOA). Participants then report which target item was perceived to have appeared first. Typically, stimuli that are attended, due to the spatial orienting of the cue, are perceived earlier than stimuli that are unattended, a phenomenon known as visual prior entry. Visual prior entry occurs because attention accelerates the processing of sensory stimuli, thereby decreasing the time between the physical onset of a stimulus and its further processing (Shore, Spence, & Klein, 2001; Stelmach, & Herdman, 1991; Ulrich, 1987). Visual prior entry, as measured through TOJs, can assess the magnitude of capture produced by different forms of cues. For example, Shore et al. (2001) observed that an exogenous cue (i.e., a peripheral flash) produced a greater degree of prior entry than an endogenous cue (i.e., an arrow). Through this same logic, TOJs can be used to assess how different populations might respond differently to the same attentional cue.

Visual prior entry is measured by calculating a point of subjective simultaneity (PSS), which indicates the SOA needed for the participant to perceive both the cued target item and the uncued target item as arriving simultaneously. Thus, the PSS shows how much time the uncued stimulus must occur before the cued stimulus in order for both to be perceived as occurring simultaneously. This provides an accurate temporal measure of attentional capture produced by an exogenous cue. Importantly, various control tasks such as ternary response tasks (Ulrich, 1987), simultaneity judgments (Stelmach & Herdman, 1991), and altering task instructions (Shore et al., 2001; West, Anderson, & Pratt, in press) have shown that prior entry effects are not driven by response bias. In other words, accelerated stimulus perception associated with prior entry has been reliably shown to result from allocated spatial attention and not some other non-attentional account.

We hypothesized that if VGPs do in fact show greater degrees of attentional capture as compared to NVGPs, then a greater degree of visual prior entry would be observed in VGPs. In other words, in VGPs, a greater amount of time between the uncued and cued stimulus would have to occur for both stimuli to be perceived as arriving simultaneously as increased capture in the VGPs speed the sensory processing of the item at the attended location.

### Method

#### Subjects

Twenty-four 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. 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 with a minimum of two hours of playing time on each day. An abridged list of the games reported as played included: Grand Theft Auto 3, Half-Life 2, Counter-Strike, Need for Speed, Halo 3, Rainbow Six: Las Vegas, Mario Kart, and Call of Duty 3. The criterion to be considered a NVGP was little or no video game usage in the past six months. Twelve males (eleven right- and one left-handed) with a mean age of 19.4 years fell into the VGP category, while twelve right-handed males with a mean age of 19.6 years fell into the NVGP category. Nine of the twelve members in the NVGP category reported no video game experience whatsoever in the past six months. The remaining three described their video game experience as once per month or less and with games that were more
puzzle or strategy based (e.g., Red Alert 2, Civilization IV, Age of Empires III). No reliable difference in age between the two groups was found, $t < 1$. Written informed consent was obtained from each subject and each subject was paid $10.00 for each hour of participation. Subjects participated in both Experiments 1 and 2, with the order of experimental presentation counterbalanced.

**Stimuli and procedure**

Stimuli were presented on a 19-in ViewSonic Graphic Series G90fb monitor (1024 × 768 resolution; 85-Hz refresh rate). Viewing distance was held constant at 44 cm by chinrest. The display (see Figure 1) consisted of two boxes and a central fixation cross that remained on the screen throughout the trial. After 1000 ms, an exogenous cue was presented, which was produced by the line making up the cued box thickening to 8 pixels for 45 ms and then returning to its original size (2 pixels), each location was equally likely to be cued. After a cue-target interval of 60 ms, the first stimulus (either a horizontal or vertical line) was presented in one location followed by the presentation of the remaining stimulus in the other location. Participants indicated which stimulus was perceived as having the first onset by pressing “z” for horizontal first and “/” for vertical first. Both stimuli had an equal chance of being validly cued (cue and target appearing in the same location), and an equal chance of appearing in the left or right location.

The PSS was calculated by adjusting the SOA between the onset of the vertical and horizontal bar stimuli using an adaptation of Stelmach and Herdman’s (1991) step-function procedure. Participants began the experiment with an SOA of 192 ms. On invalidly cued trials, the SOA was reduced by 12 ms when participants made a correct judgment of temporal order. Conversely, when an incorrect judgment was made the SOA increased by 12 ms. Participants went through this procedure until maximal uncertainty was reached, which was indicated by the SOA reversing direction consecutively six times (Suzuki & Cavanagh, 1997). We obtained the value of the PSS by averaging the SOA of the last two reversals. The PSS was calculated to be the SOA where the participant was maximally uncertain as to which stimulus had first onset, making an equal amount of “horizontal first” and “vertical first” responses. In other words, the PSS was the calculated amount of time the uncued item at one spatial location needs to arrive before the cued item at the other spatial location in order for both items to be perceived as arriving simultaneously.

**Results**

The average PSS for VGPs and NVGPs is shown in Figure 2, with larger values indicating greater degrees of attentional capture in response to the exogenous cue. The PSS was calculated for each individual subject as the average SOA value between the last two reversals, where participants had reached maximal uncertainty regarding which stimulus had the first onset. Overall, it was found that VGPs showed a higher PSS (84.9 ms) as compared to NVGPs (55.2 ms; $t(22) = 2.17, p < .04$). In other words, in VGPs the uncued stimulus had to arrive 84.9 ms before the cued stimulus to be perceived as arriving simultaneously, while in NVGPs this was reduced to 55.2 ms.

![Figure 1](https://jov.arvojournals.org/pdfaccess.ashx?url=/data/journals/jov/932852/) The stimuli and procedure for **Experiment 1** is outlined. On each trial participants were presented with two placeholder boxes on either side of fixation. After 1000 ms one randomly chosen placeholder was cued for 45 ms followed by the onset of either a horizontal or vertical line in one of the placeholders after 60 ms. The onset time of the second stimulus was determined by the step-function which monitored temporal order judgment performance.
Thus, VGPs were observed to be more sensitive to the capture effects of the exogenous cue, resulting in longer SOA periods between both target stimuli needed to correctly perceive their temporal order.

**Experiment 2**

In Experiment 2 we used signal detection methods (Green & Swets, 1966; Macmillan & Creelman, 1991) to further assess whether VGPs are more sensitive to the detection of salient sensory events that capture attention, in this case an abrupt change in motion (Abrams & Christ, 2003). The signal detection model uses the relationship between hit and false alarm rates to compute an independent measure of discriminator ability. Through this, sensitivity ($d'$) is computed as a measure of how well a target event is detected amongst some background noise. A larger measure of $d'$ is indicative of an increased ability of the observer to detect signal from noise. In addition, estimates of $d'$ are thought to be independent of response bias when heterogeneity of variance between both samples is assumed; in other words, the sensitivity measure is relatively unaffected by the response strategy an observer adopts (Rotello, Masson, & Verde, 2008). It is also important to note that measures of $d'$ reflect early perceptual processing, such as the sensory encoding of a stimulus event, and is not sensitive to later response-related processes (Handy, Kingstone, & Mangun, 1996; McDonald, Teder-Sälejärv, & Hillyard, 2000). Thus, observed differences between VGPs and NVGPs in the current experiment would reflect differences in sensory processing and not some later visual selection process.

In the present experiment, participants were presented with a dynamic display of moving schematic ‘swimmer’ stimuli, and told to covertly monitor the array for the presence of an change in movement in one randomly chosen swimmer. Thus, this experiment aimed to test differences in attentional capture in VGPs and NVGPs as would be encountered more closely in the natural world; a display of salient visual activity relative to a backdrop of ongoing visual activity, and not simply a static visual background. Indeed, the inspiration for our task came from lifeguards at a swimming pool who must monitor busy visual scenes full of normal swimming motion for instances of abnormal swimming motion that would correspond to a person in trouble. We hypothesized that visuospatial experience obtained by VGPs would result in their increased sensitivity to detect salient changes in motion, and thus have a greater ability to capture attention from the level of baseline motion in the background when compared to NVGPs.

**Method**

**Stimuli and procedure**

Stimuli were presented on the same type of monitor as in Experiment 1, with the viewing distance adjusted to 20 cm in order to increase the range of visual angle.
available on the display screen (approximately 94° × 70°). The display for Experiment 2 (see Figure 3 and Auxiliary Video 1) consisted of a central fixation cross which remained on screen for 1000 ms before the onset of either 15 (low perceptual load condition) or 30 (high perceptual load condition) schematic swimmers which each consisted of a circle 3° in diameter with arms 2 pixels in thickness. Each swimmer moved around the visual array in a straight pathway chosen at random, reversing its trajectory if it came in contact with another swimmer or with the invisible 3° boundary surrounding the perimeter of the screen. The arms of each swimmer oscillated at a randomly jittered rate of 1–3 Hz. The target stimuli were Non-Swimmers Targets (NSTs) which occurred on half the trials. On NST trials, one of the swimmers stopped in its location 10°, 20°, or 30° from fixation and increased its arm oscillations at a randomly jittered rate of 15–18 Hz. The NST had an equal chance of occurring at each eccentricity, and in each quadrant of the display array. To reduce anticipatory responses, the length of each trial was chosen at random between 1500 and 3500 ms, with non-target trials timing out at this chosen time value, and NSTs having their onset at this time point. As an additional control, catch trials were also intermixed with experimental trials. On these trials, a circular ‘buoy’ object subtending 4° in diameter appeared on screen in a random location, at which point participants had to respond by pressing the ‘3’ key with their right hand.

Participants completed 20 practice trials followed by a total of 440 experimental trials (including 20 catch trials), with an equal amount of trials presented in each perceptual load and target eccentricity condition. Participants were instructed to hold their gaze centrally while covertly monitoring the visual array for NSTs and to press the “1” key on the numeric keypad with their right hand when they detected the onset of an NST at which point a hit was recorded. Conversely, if participants indicated the presence of an NST when one was not displayed, a false alarm (FA) was recorded. Trials timed out after 2000 ms of the onset of the NST. Feedback was given. A 200 Hz tone was presented when a target was missed. Breaks were given every 55 trials.

Results

The overall raw data for Experiment 2 can be seen in Table 1. To measure group differences between VGPs and NVGPs in detection sensitivity to salient motion changes in the visual field, \( d' \) measures for both the between-subjects factor of video game playing and the within-subjects factors of perceptual load and target eccentricity were calculated. When either the hit rate is one (100%) or the FA rate is zero, \( d' \) has an unlimited numeric value and hence cannot be included in the analysis. To account for this, the estimates of \( d' \) were converted as proposed by Macmillan and Creelman (1991), and proportions of one

![Figure 4](https://jov.arvojournals.org/pdfaccess.ashx?url=/data/journals/jov/932852/ on 10/17/2018)

Figure 4. The average \( d' \) score at each eccentricity in each perceptual load condition is shown for video game and non-video game players. Error bars represent one SEM.
were converted to \([1-1/2N]\)) and proportions of zero were converted to \([1/2N]\), with \(N\) representing the number of trials within that proportion.

Estimates of \(d'\) were submitted to a 2 (VGP Status; VGP vs. NVGP) \(\times\) 2 (Perceptual Load; Low vs. Load) \(\times\) 3 (NST Eccentricity; 10\(^{\circ}\), 20\(^{\circ}\), or 30\(^{\circ}\)) mixed factorial design repeated measures ANOVA. This revealed a main effect of VGP status, with VGPs showing an overall higher \(d'\) detection estimate (2.41) than NVGPs (1.59) \(F(1,22) = 12.76, p < .01,\) partial \(\eta^2 = .48\). Main effects of perceptual load \(F(1,22) = 10.10, p < .01,\) partial \(\eta^2 = .32,\) and NST eccentricity \(F(2,44) = 58.06, p < .01,\) partial \(\eta^2 = .73\) were also observed. A significant Perceptual Load \(\times\) NST Eccentricity interaction was also observed \(F(2,44) = 6.64, p < .01,\) partial \(\eta^2 = .24.\) Planned comparisons between \(d'\) estimates of VGPs and NVGPs at each NST eccentricity within each perceptual load condition were also conducted. Planned comparisons using one-tailed \(t\)-tests, due to the a-priori nature of our hypothesis, were conducted with a corrected threshold of \(p < .05,\) using the Student–Newman–Keul’s procedure. As shown in Figure 4, all but the 10\(^{\circ}\) eccentricity in the low perceptual load, and the 30\(^{\circ}\) eccentricity in the high perceptual load condition showed a significant difference in estimates of \(d'\) between VGPs and NVGPs.

### Discussion

Taken together, the two experiments in this study demonstrate that habitual action video game play modulates early sensory processing, resulting in increases in sensitivity to salient visual events that capture attention. Using the TOJ paradigm, Experiment 1 demonstrates, for the first time, a differential cueing effect between VGPs and NVGPs. Just as TOJs have been demonstrated to be sensitive enough to detect a difference between the effect of an exogenous and endogenous cue on attentional orienting in normal individuals (Shore et al., 2001), our current work demonstrates that this paradigm can also detect differences in sensory functioning between different populations. Specifically, it was found that VGPs show heightened sensitivity to a peripheral uninformative cue when compared to NVGPs, demonstrating the cue’s increased ability to capture their attention. Experiment 2 extends this finding using a dynamic visual display, demonstrating VGPs’ increased sensitivity to salient changes in motion relative to an array of background motion. In addition, it is noteworthy that VGPs continued to demonstrate increased sensitivity to targets occurring at larger eccentricities (Feng et al., 2007; Green & Bavelier, 2003), and across both high and low perceptual load conditions. Overall, both experiments suggest the degree of visuospatial experience possessed by an individual modulates the initial stages of sensory processing where attentional capture occurs. Nevertheless we need to point out that the present study compared individuals whose self-identifications allowed us to place participants into the two respective groups. Thus there is the possibility that our results are due to some form of population bias, however, previous studies of this nature have ruled out this possibility with the inclusion of a separate training experiment (Feng et al., 2007; Green & Bavelier, 2003, 2006, 2007). With this said, the current study does fit well with existing evidence suggesting that such experience obtained through action video game play can affect cognitive processes only thought to be partially related to a game’s task. For example, short term memory skills (Green & Bavelier, 2006), and visual search time (Castel et al., 2005); have also been found to be modulated through action video game play. Our current findings suggest that these online cognitive processes may be in part driven by this enhanced sensitivity to salience in the visual field. It should also be noted that it is possible that these findings do not reflect attentional capture, but some earlier enhancement of sensory processing. Recently proposed frameworks of attention and perception, however, suggest that both these processes may occur in unison (e.g., Serences & Yantis, 2006), and thus an enhancement in sensory processing would also reflect an enhancement in the process of attentional orienting.

The results of Experiment 1 report for the first time a difference in attentional orienting produced by a cue between VGPs and NVGPs. A significant difference in attentional orienting using a standard cueing paradigm was not observed in a previous study (Castel et al., 2005). This is possibly due to the fact that TOJs only measure attentional engagement whereas typical cueing tasks also involve disengagement processes. Because of this, the TOJ paradigm is a more sensitive measure of initial attentional deployment where capture occurs.

It is also important to note that both experiments did not focus on RT, but were based on reports of the contents of the visual display. Therefore, it is unlikely that increased stimulus-response mapping ability in VGPs was responsible for our observed effects, as is found in expertise effects involving more specific cognitive domains (Hambrick & Engle, 2002; Kramer & Willis, 2003; McCarley, Kramer, Wickens, Vidoni, & Boot, 2004). This raises the question of what neural mechanisms are modulated from habitual action video game play. It is possible that several systems experience changes during habitual play including the subcortical striatal dopaminergic system and the cortical posterior parietal attentional network. There is now substantive evidence that implicates the striatal dopaminergic system with error monitoring (Schultz, Dayan, & Montague, 1997), and with responses to salient, unpredicted events (Hurvitz, 2000; Jensen et al., 2003), making it a candidate system of possible modulation in VGPs. Also, some evidence indicates that this system can be modulated during video game play, with activity levels showing a direct relationship with game task performance (Koeppe...
et al., 1998). The posterior parietal attentional network, a group of structures including the dorsal frontal cortex, posterior parietal lobe, lateral pulvinar nucleus, and the superior colliculus, have long been thought to be responsible for the engaging and shifting of both voluntary and involuntary attention (Hopfinger, Buonocore, & Mangun, 2000; Posner & Peterson, 1990). Parts of this system are recruited along with limbic structures during attentional orienting to salient and/or novel stimuli in the visual array (Downar, Crawley, Mikulis, & Davis, 2002; Knight & Scabini, 1998), thereby also making it a possible group of structures that experience long term modulation through action video game play.

Finally, it is worth mentioning that increased visuospatial abilities are sometimes associated with increased aptitude in higher levels of cognitive function associated with science and mathematics (Delgado & Prieto, 2004), and performance on standardized tests (e.g., the SAT; Casey, Nuttall, Pezaris, & Benbow, 1995). This is thought to be due to the fact that lower level cognitive functioning supports higher level processes, thereby suggesting that training in these lower level domains might facilitate performance in higher level domains. This suggests, from a pedagogical standpoint, that focusing on the training of visuospatial ability at an early age, similar to that experienced through action video games, might prove to be an important part of an effective educational system. Indeed, some educational institutes that specialize in cognitive rehabilitation for individuals who have suffered from brain trauma or learning disability have adopted this philosophy, with reports of high efficacy (e.g., Doidge, 2007). Further investigation into the possibility of such visuospatial training enhancing higher cognitive function in healthy younger individuals is warranted.

The present study has shown that action video game experience can modulate the degree to which visual stimuli automatically engage attention over other items in the visual field. This finding has important practical and theoretical implications relating to cognitive rehabilitation, education, and theories of neural plasticity. The fact that game experience can modulate the earliest of sensory and attentional processes, as well as other stages of information processing, suggests that action video game experience can have positive and profound effects in a variety of settings.

Acknowledgments

This research was supported by the Lifesaving Society of Canada and a grant from the Natural Sciences and Engineering Research Council of Canada to Jay Pratt. Greg West and Sara Stevens were supported by Canada Graduate Scholarships from the Natural Sciences and Engineering Research Council of Canada. We thank James Honeyman from the Lifesaving Society of Canada, and Hrag Pailian, Ruo Guo, Bry Kotyk, and Brad Gerard for assistance with this project.

Commercial relationships: none.
Corresponding author: Greg West.
Email: greg.west@utoronto.ca.
Address: Department of Psychology, University of Toronto, 100 St. George Street, Toronto, ON M5S 3G3, Canada.

References


West, G. L., Anderson, A. K., & Pratt, J. (in press). Motivationally significant stimuli show visual prior