Attention to Low & High Prevalence Events in Action Video Game Players & Non-Action Video Game Players Using Sustained Multiple Object Tracking & Change Detection Tasks

Keith R. Racioppo
State University of New York at New Paltz
May 14th, 2020

Advisor: Dr. Elizabeth Hirshorn
Readers: Dr. Giordana Grossi & Dr. Doug Maynard
Abstract

Surveillance is an important real-world skill involving several cognitive abilities over a prolonged period. Multiple object tracking (MOT) and change detection research have begun to conceptualize the cognitive processes associated with surveillance in a laboratory setting. The current study incorporated change detection into a more sustained MOT task than what had been studied previously. This experimental design may better represent real-world situations in which identification of changes in items occurs in the real world due to the often-infrequent rate in which it is necessary to recognize changes. Additionally, long-term action video game experience and short-term experiences, such as exposure to rates of prevalence, are examined to help identify potential trainings to improve performance. After four 10-minute MOT trials, a short change detection task was conducted to assess a possible relationship of gaming and recent prevalence experience on later tasks. A main effect of items tracked was the only significant effect found throughout the research, indicating individuals can effectively track 2 items for changes and not 4. A main effect of prevalence was found in the MOT task, giving merit to the inclusion of prevalence in change detection MOT tasks put forth by this study. However, neither gaming experience in the MOT task nor prevalence experience in the follow-up task led to improved performance in the task. This study was conducted during the COVID-19 pandemic, which resulted in a small sample size and low power. Further research is necessary to examine potential mechanisms for surveillance training, but the current design can serve as guidance for future studies.
Table of Contents

I. How to Study Surveillance: Multiple Object Tracking (MOT) & Change Detection……5

II. The Role of Experience in Effective Surveillance……………………………………...7

III. Current Study…………………………………………………………………………10

IV. Methods

Participants………………………………………………………………………………11

Materials………………………………………………………………………………12

Procedure………………………………………………………………………………13

V. Results

MOT Task………………………………………………………………………………16

Change Detection Task……………………………………………………………..17

VI. Discussion……………………………………………………………………………18

VII. References…………………………………………………………………………23

VIII. Figures………………………………………………………………………………25

IX. Appendixes…………………………………………………………………………32
Attention to Low & High Prevalence Events in Action Video Game Players & Non-Action Video Game Players Using Sustained Multiple Object Tracking & Change Detection Tasks

The ability to survey, scan, or track objects effectively in order to identify changes that require further action is an important skill. Surveillance is a necessary process in several occupations such as lifeguards, TSA employees, and air traffic controllers, as well as in simple everyday tasks such as driving. Effective surveillance over an extended period requires maintaining a level of alertness and vigilance in order to both keep track of moving items and additionally recognize salient out of the ordinary changes. For example, keeping track of swimmers in a pool to notice a change from swimming to struggling or watching for a person that may be submerged underwater, represents a situation in which vigilance would be required for an extended period. Factors that may contribute to the efficiency of surveillance are the number of items being tracked (e.g., number of people in the pool) (Wu & Wolfe, 2016) and how obvious a change might be (e.g., yelling for help or quietly submerging underwater) (Wolfe, Place, & Horowitz, 2007). Other contributions to effective surveillance may be the rate in which important changes occur (Wolfe et al., 2007) or individual experiences that may affect skill level (Green & Bavelier, 2003; Green & Bavelier, 2006; Clark, Fleck, & Mitroff, 2011).

Understanding the psychological processes associated with sustained surveillance is important to determine the cognitive limits of surveillance capacities and how these skills are performed in individuals. Once the processes that are involved in effective surveillance are better understood, this area of research has the potential to help better train the individuals who are expected to perform effective surveillance.
How to Study Surveillance: Multiple Object Tracking (MOT) & Change Detection

Recent research investigating surveillance has focused primarily on the capacity or number of objects someone can effectively keep track of using Multiple Object Tracking (MOT). MOT studies ask participants to track multiple items around a fixed screen with similar distractors present. At the end of each 10-20 second trial participants are asked to identify the items they were supposed to be attending to (Green & Bavelier, 2006; Wolfe, Place, & Horowitz, 2007; Wu & Wolfe, 2016). Studies examining MOT have shown that participants have a maximum attentional limit of tracking four items effectively (Cowan, 2001; Wolfe et al., 2007).

With surveillance, it is important to not only effectively keep track of items, but to also identify changes in those items when they occur. Variations on the MOT task have included the need to identify changes in items (e.g., a T changing to an L). This may simulate a scenario where an individual is the only active lifeguard at a pool, not only watching all swimmers, but also making sure to attend to changes in the swimmers, such as when someone may need help. When a change detection task is required in addition to tracking multiple items, the number of items able to be tracked efficiently falls to 2-3 (Wu & Wolfe, 2016). Although MOT tasks with change detection in a lab setting cannot perfectly replicate real-world scenarios, it is easy to see how these types of MOT tasks can help build a foundation for the understanding of real-world surveillance tasks.

Wu and Wolfe (2016) conducted studies where all 2, 4, 6 and 8 items on the screen had to be attended to for changes using various stimuli, such as letters (T’s to L’s and vice versa) or objects (open/closed book) changing states. In addition to change events, transfer events were also used in some trials. Cartoon animals or shapes would transfer an item from one to the other, such as an apple or a dot, from one stimulus to another. When a change was detected,
participants were asked to press a predetermined key on the keyboard to freeze the scenario, which would then conceal items as neutral shapes to prevent counting or further analysis, and asked them to click on the item that changed or items that were involved in an exchange. The findings indicate that, regardless of the task or object being studied, only two items could be effectively tracked and identified for a change or exchange event. However, up to four items could be effectively tracked for changes and not exchanges, while 6 and 8 items could not be tracked effectively in either case (Wu & Wolfe, 2016). Regardless of the stimuli (letters, objects, and animals), the number of items able to be correctly monitored was the same. This result supports the idea that the same attentional resources are used to track different types of objects, as well as claims by Cowan (2001) that the maximum number of items that can be effectively tracked is four.

Wolfe, Place, and Horowitz (2007) utilized a dynamic MOT task, where tracked items were removed and added throughout each trial, keeping a total of 2-3 items tracked at any point in each of the 20 second trials. Feedback was given in one study, in which at the end of each trial participants were asked if the indicated item was being tracked with a tone indicating correct or incorrect after their selection. A follow up study used a neutral feedback tone regardless of whether the selection was correct or incorrect. The dynamic MOT task studies found that compared to the fixed groups, where the same items remained tracked throughout each trial, items were better attended to with neutral feedback, compared to the fixed conditions without any feedback, by over 15% (Wolfe et al., 2007). Participants seem to have performed better due to a neutral feedback from an item changing during the task, compared to the lack of information provided during the fixed scenario. The neutral feedback being given by both selection tone and the removal or addition of items in the dynamic tasks may have effects on identifying changes
more similar to real world situations where these skills are used, as opposed to not getting any feedback throughout these tasks. In real life scenarios, visual information is being conveyed and updated continuously on the state of the item being tracked. The addition and subtraction of items may help convey similar information in a research setting by giving participants some updates on tracked items throughout the trial, compared to having no updates given at all during each of the fixed trials. Returning to the lifeguarding example, a person will not remain in the same state throughout their time in the pool area. They may get out and rest, requiring little to no attention at all, or they may jump in the shallow end and require more attention. For these reasons, the current study will use a dynamic task, as well as provide neutral feedback to ensure that more information is being conveyed.

Lastly, a typical MOT trial is very short. Current MOT research is usually conducted in short 10-20 second repetitive trials asking participants to identify the tracked items at the conclusion of each trial (Green & Bavelier, 2006; Wolfe, Place, & Horowitz, 2007; Wu & Wolfe, 2016). However, these short trials are not representative of the real-world situations in which surveillance is conducted. It is more common for surveillance to be conducted over an extended period rather than in short increments (e.g., lifeguard shifts are minutes or hours, not seconds). Therefore, the current study aims to research surveillance that better reflects the environment in which it is performed by using longer 10-minute MOT trials.

The Role of Experience in Effective Surveillance

There is evidence to support that various types of experience may influence performance in surveillance related tasks. Short-term experience, such as the rate of exposure to items of interest in static images, has demonstrated increased rates in identifying items in later trials
There is also research demonstrating that long-term experience may also improve surveillance performance. Experience that is gained over an extended period, such as action video game playing experience, has demonstrated a relationship between experience and related cognitive functions that may influence improved performance on surveillance tasks (Green & Bavelier, 2006; Clark et al., 2011). Gaining a better understanding of how these types of experiences affect performance on surveillance tasks, as well as how short-term and long-term experiences may influence one another in improving performance is a focus of this study.

Short-term experiences may be developed over an hour or a day, compared to other experiences that are gained over months, years, or a lifetime. Understanding how experiences of a relatively brief duration may influence performance on a task can help develop possible trainings or interventions for individuals performing surveillance tasks. Beginning to understand how short-term experience can improve detecting changes in real-world surveillance tasks, it is important to better examine how often or infrequently the changes occur. Short-term experience has been shown to improve identification of Low Prevalence Events (LPEs) in studies involving change detection in static images (Wolfe et al., 2013). LPEs and High Prevalence Events (HPEs) have both been studied in tasks, such as luggage in TSA screenings (Wolfe et al., 2013), mammograms looking for cancerous cells (Evans, Birdwell, & Wolfe, 2013), and in tasks identifying arbitrary items among distractors (Hout et al., 2015). In previous static prevalence studies, it has been shown that, unless presented with examples of what is being looked for frequently, items of interest are less likely to be recognized (Wolfe et al., 2013; Evans, Birdwell, & Wolfe, 2013; Hout et al., 2015). In Wolfe et al. (2013), five trials of 100 TSA luggage photos were presented to newly trained TSA employees. These trials were presented in low-low-low-
high-low prevalence order, with low prevalence trials having a 5% rate of images requiring further attention (e.g. 5 images of interest out of 100), and high prevalence having a 50% rate. Prevalence effects were found in these newly trained TSA employees, demonstrating an increased detection rate in recognizing items of interest following high prevalence trials, compared to trials following low prevalence trials. Low prevalence trials were attended to less effectively each trial until the high prevalence trial was presented. When the LPE trial was presented following the HPE trial, the items of interest were attended to more effectively than all other LPE trials. More false alarms were also reported at significantly higher rates in LPE trials following HPE. This is not necessarily negative in these types of situations because it is better to think you see an issue that is nothing, rather than overlooking something that is an issue.

These findings lend support to the ability of short-term experiences and training aiding effective identification of target items, but research on how performance as a function of change prevalence in MOT tasks has not yet been explicitly studied. The current study will examine prevalence of changes and whether HPEs can be considered a form of ‘short-term’ training mechanism in a more realistic dynamic surveillance task, and further how exposure to LPE and HPE scenarios can affect how one identifies changes in subsequent tasks. Another type of experience that has been studied in MOT and other attentional tasks is action video game playing. What separates an Action Video Game Player (AVGP) from a Non-Action Video Game Player (NVGP) are the many hours played over a sustained period. Previous research has required participants to have played 5 or more hours of action video games per week for past 12 months in order to be determined an AVGP (Green & Bavelier, 2006; Clark et al., 2011). Thus, action video game experience can be considered ‘long-term’ experience by the standards of previous research. Action video game experience has been shown to improve certain cognitive
functions, such as change detection (Clark, 2011), which may improve performance in sustained surveillance MOT tasks. Additionally, action video game experience has demonstrated improved performance in other related areas of research such as multiple object tracking (Green & Bavelier, 2006) and visual working memory (Blacker, Curby, Klobusicky & Chein, 2014). However, these previous studies have been conducted in a series of short successive trials, rather than in a prolonged continuous scenario which may better represent real-life situations in which these tasks are conducted. The current study will explore whether these reported benefits of action video game playing extend to a more sustained task, rather than just the previous studies that found these benefits using 10-20 second trials.

**Current Study**

The proposed research is investigating possible ways to improve the understanding of MOT and change detection tasks, in order to be able to inform how those processes are conducted in real world scenarios. More specifically, this research will examine how action video game experience, prevalence of change events, and number of items tracked, influence performance in a sustained surveillance MOT task. Accuracy in a subsequent change detection task will also be assessed to better examine more realistic change detection, as well as the possible effects of prevalence on later tasks to explore aspects of surveillance training. It is predicted that there will be main effects for Tracked Items, Group, and Prevalence (See Figure 1), in which AVGP and HPE will lead to better identification of the objects being tracked at the conclusion of each trial (tracking accuracy) and when changes occur (change accuracy). Based on previous research, all groups should be best at tracking 2 items and the worst at tracking 4. A three-way interaction is predicted between Group, Prevalence, and Tracked Items, in which the
interaction between Prevalence and Tracked Items will be influenced by Group, such that NVGPs will benefit the most from prevalence when tracking 4-items.

For the follow up task, we will be looking to see if there is a main effect of Group or recent prevalence experience on performance in a change detection task. It is predicted that there will be main effects for both Group and Recent Prevalence Type, such that AVGPs will perform better than NVGPs, and that performance will be better when the most recent trial type was an HPE compared to an LPE (See Figure 2). It is also predicted that there will be an interaction between Group and Recent Prevalence Type, such that the difference between AVGP and NVGP will be larger in those individuals that most recently had LPE trials, compared to HPE trials, in the MOT task. If these interactions are found, it would lend support to recent experience and video game training being able to best improve ability to identify changes in stimuli.

**Methods**

**Participants**

The study consisted of 18 participants (12 female, 6 male) that were SUNY New Paltz psychology undergraduates, over 18 years of age, had normal or corrected-to-normal vision, and passed the Ishihara color screening (Ishihara, 1989). Recruitment of participants was halted due to the COVID-19 pandemic which resulted in fewer participants than was originally anticipated. The initial design of the study sought 40 total participants evenly distributed by video game experience and gender. This would have created groups of 10 for each video game experience and prevalence level. Each group of 10 would have then been evenly split by gender to prevent any limitations of the study which is similar to how previous research has been designed. Participants were recruited through the SUNY New Paltz subject pool, as well as emails.
Participants were compensated with psychology course credits for a single 1.5-hour testing session.

Participants were screened using the “Video Game Questionnaire” (see Appendix A) and criteria set by Green and Bavelier (2003), revised March 2018, to meet the AVGP and NVGP requirements. Those that reported playing action video games over the past 12 months and met the necessary criteria set by Green and Bavelier (2003), revised March 2018 (see Appendix B), were considered AVGPs. Participants that played less than 1 hour of action or other types of video games per week, or no video games at all were considered NVGPs. All other participants that failed to qualify for either AVGP or NVGP were removed from analysis. A total of 6 participants (1 male, 5 female) data that were not analyzed because they did not meet the qualifications for either AVGP or NVGP. The final data set consisted of 7 AVGPs (2 female, 5 male), and 5 NVGPs (5 female, 0 male).

Materials

The study was run in a well-lit room on a Dell Optiplex 7040 with a 22-inch Dell P2214H monitor with a 1920x1080 pixel resolution and a 60Hz screen refresh rate. Participants were placed at an approximately 57cm viewing distance from the monitor without head restraint, at this distance 1cm of screen is roughly equivalent to a visual angle of 1°. All stimuli were roughly 0.83°x0.83° in size and were presented within a gray display nearly 20°x20° as in previous MOT tasks (Wu & Wolfe, 2016), with a black border filing the screen. The MOT task was presented, and responses collected with MATLAB R2016a software with the Psychophysics Toolbox Version 3 (Brainard, 1997).

The follow up change detection task was presented using the same computer and monitor as the MOT Task. The stimuli consisted of 8 photographs of natural scenes that were presented
filling roughly 19ºx13º of their visual field. The stimuli were presented in the center of the screen with a white background filling the remaining portion of the screen. These photographs were retrieved from the Cengage CogLab website (Francis & Neath, 2019) and have been previously used and verified in change detection tasks. There were 2 images provided for each photo, one of the images was the original and the other photoshopped to make a change in the photo while making the photo seem not doctored. There was a total of 4 images with changes used, with changes being equally distributed among the four quadrants of the visual field. The stimuli for the change detection task were presented, and responses collected electronically using E-Prime 3.0 (Version 3.0.3.60) (Psychology Software Tools, 2016).

**Procedure**

Participants were separated by AVGPs and NVGPs. The study consisted of four 10-minute MOT trials using 4 and 8 letters (2 trials of each), with equal amounts of T’s and L’s presented as stimuli on the screen to start each 10-minute trial. During each trial, half of the letters were tracked, never exceeding 2 or 4, for 4 and 8 item trials respectively. Participants were given a 5-minute practice trial with 1 tracked item and 1 distractor, prior to conducting the study to ensure the process and instructions were as clear as possible. Trials were counterbalanced by prevalence and ordered so that all participants received the 4 letter trials (2 tracking) first, followed by the 8 letter trials (4 tracking). This was done to prevent participants from being overwhelmed with an initial trial of 8 items. Half of the participants received LPEs first (5 change events in 10-minutes), followed by HPEs (30 change events in 10-minutes), and the other half of participants received HPEs prior to LPEs. Participants were asked to track items as they were added and subtracted, ~30s (+/- 3s) for each addition/subtraction, by briefly being
highlighted red for 2s to remove and highlighted green for 2 seconds to add items, timed to not occur or interfere with the change event by 3s.

Each trial began when the participant was ready with a click of the mouse. The letters for the trials would then appear with half of the letters on the screen (with at least 1 of each T & L) being green to be tracked, with the remaining letters being black as distractors. All letters moved on the screen with a velocity of 4º/s, as previous studies have used (Wu & Wolfe, 2016). After 5s the tracked letters would turn black and remain tracked. During the change events, one of the letters being attended to would change to the other (e.g., T → L) every ~20s (+/- 3s), a total of 30 change events per trial for HPEs, and ~120s (+/- 3s), a total of 5 change events for LPEs. If/when a change was detected, participants were asked to press the spacebar as fast as possible. Then the stimuli changed from letters to neutral circles with a “?” in the center (Figure 3). This was done to prevent counting or other strategies, while allowing participants identify the location of the letter that changed. Participants had 5s to recognize a change with the space bar, and 5s to make their selection by mouse click. The success rate of correctly identifying these changes through each MOT trial will be recorded as change accuracy. A correct response was recorded if participants hit the spacebar within the 5s of a change event and correctly identifying the item that changed. A neutral tone was given after each mouse click selection, as well as when a change was not detected within the 5 seconds. This was done to give participants feedback throughout the trial about the rate of change.

At the conclusion of each trial participants were asked to identify the 2 or 4 items they were tracking to record tracking accuracy. They were asked to click until all items that were supposed to be tracked were identified, meaning if it took 3 clicks to identify 2 items, a ⅔ or 67% rate of success would be recorded, as previous research has used for a dependent measure.
A 2-minute break was given after each 10-minute trial to allow participants a moment to relax before the following trial.

Following the 4 MOT trials and breaks, a change detection task replicating that of Clark, Fleck and Mitroff (2011) was performed. All participants received the same exact task in the same order, which consisted of 2 practice images (1 with a change) and 8 study images (4 with changes). Each trial began with a fixation cross being presented for 500ms. A photo was then presented for 250ms, then a 250ms screen mask, followed by the same photo for 250ms with or without a change, ending with a neutral screen. Participants were asked to click with the mouse where they were last looking on the image. After their location selection, participants would be prompted with 3 options, “I guessed”, “I need to verify”, or “I saw it”. If participants had indeed just guessed and had not seen any change they were to select “I guessed”. If they believed that they saw a change, but needed another representation to be certain, they should select “I need to verify”. Participants were informed only to select “Saw” if a change is detected, which would then begin the next trial. If “I guessed” or “I need to verify” was selected the process would repeat for up to 15 presentations for each of the 8 photos (4 with change, with 1 in each quadrant and 4 without change). Properly identifying the change and selecting “Saw” would indicate properly identifying a change. The area for each image with a change that was recorded as correct was defined by clicking with a rectangular box that encompassed the change by roughly 30 pixels on each side of the change (Clark et al, 2011). This was done to accommodate participants clicking on a masked screen and may not have clicked in exactly the right location. If a change was not detected within the 15 presentations using “Guess” and “verify” selections the trial ended, and participants were prompted that the next trial would begin with their click. In order to correctly identify a no change trial, participants would have to view all 15 presentations.
selecting “Guess” or Verify” for each presentation, unless incorrectly selecting “Saw” during the trial (Clark et al., 2011).

Results

MOT Task

For the purposes of this study the most important data in the MOT task is the change accuracy and tracking accuracy. A three-way mixed ANOVA was used to examine the effects of Experience (AVGP & NVGP), Prevalence (Low & High), and Tracked Items (2 & 4) for the MOT task. Normality checks were carried out on the residuals for each dependent variable using Shapiro-Wilk tests. The normality checks were approximately normally distributed for most dependent variable, except for change accuracy in 2-item tracked high prevalence trials and tracked accuracy in both low prevalence trials for 2 and 4 tracked items. The lack of normality in these data may be due to the small overall N for the study. Due to the failure for some data to meet normalcy requirements, there is in increased likelihood of Type I error in following results. Additional analyses were run using action video game experience as a continuous variable in several ways to help alleviate some of the issues with the data and include all participants, but the findings remained relatively unchanged. There were no main effects found for the between-group variable of Experience for either change accuracy \[ F(1,10) = 1.91, p = .2, \eta^2_p = .16 \] or tracking accuracy \[ F(1,10) = 1.1, p = .32, \eta^2_p = .1 \]. There was, however, a main effect for the within-group variable Tracked Items in both change accuracy \[ F(1,10) = 32.71, p < .001, \eta^2_p = .39 \] (Figure 5) and tracking accuracy \[ F(1,10) = 80.98, p < .001, \eta^2_p = .22 \] (Figure 6), such that change detection throughout trials with 2-items tracked were identified at significantly higher rates, as well identifying the tracked items at the end of each trial. It important to mention that
the tracked 2-item trials failed to pass Levene’s Test for homogeneity of scores. This was most likely caused by the high rate of success by all participants in these trials. Additionally, there was a main effect found for within-group Prevalence on change accuracy \[F(1,10) = 6.32, p = .03, \eta^2_p = .39\] (Figure 5), which shows a higher rate of identifying changes throughout the high prevalence trials, compared to the low prevalence trials. However, the results for tracked accuracy \[F(1,10) = .14, p = .72, \eta^2_p = .01\] were not significant.

There were no significant two-way interactions for Experience and Prevalence \[F(1,10) = .16, p = .7, \eta^2_p = .02\], \[F(1,10) = 1.1, p = .32, \eta^2_p = .1\], Experience and Tracked Items \[F(1,10) = .21, p = .65, \eta^2_p = .02\], \[F(1,10) = .03, p = .86, \eta^2_p = 0\], or Prevalence and Tracked Items \[F(1,10) = 1.59, p = .24, \eta^2_p = .14\], \[F(1,10) = .13, p = .73, \eta^2_p = .01\], for change accuracy or tracking accuracy respectively. There was no significant three-way interaction between the independent variables for either the dependent variable of change accuracy \[F(1,10) = .75, p = .41, \eta^2_p = .07\] or tracking accuracy \[F(1,10) = 2.81, p = .13, \eta^2_p = .09\].

**Change Detection Task**

The most relevant dependent measure in the follow up change detection task is the accuracy in which participants recognized changes in trials that contained changes. A two-way between-subjects ANOVA was used to examine the effects of Experience (AVGP & NVGP) and most recent prevalence trial (Low & High) on change detection. A normality check was carried out on the residuals for change detection accuracy variable using Shapiro-Wilk tests, as well as Levene’s Test for homogeneity of variance, in which both criteria met assumptions.

A main effect was found for Experience \[F(1,8) = 5.48, p = .04, \eta^2_p = .41\], such that AVGP s performed significantly better in detecting changes than their NVGP counterparts in the change detection task (Figure 7). Most recent prevalence trial was not significant \[F(1,8) = 1.23,
\[ p = .3, \eta^2_p = .13 \]. There was also no significant interaction found between Experience and most recent prevalence trial \( F(1,8) = 1.837, p = .21, \eta^2_p = .19 \).

Although, the main effects for most recent prevalence were not significant AVGPs change detection was relatively the same across both most recent prevalence conditions with 41.67\% (SD = 14.43) and 43.75\% (SD = 12.5) for low and high respectively. An interesting trend, although not significant, was that NVGPs rate of detecting changes was slightly better when their most recent prevalence type was low 33.33\% (SD = 14.43), compared to when it was high 12.5\% (SD = 17.68).

**Discussion**

In this study, a dynamic MOT task was used with change detection to examine the effects of action video game experience and prevalence of changes on effective surveillance. A follow up change detection task was given to assess the effects of action video game experience and most recent prevalence trial on performance in a later task following the MOT trials. Although the small sample size reduced the power of the study and therefore could not test some of our predictions, there are still positive findings that can be taken away from the results of the study that can help guide future research on surveillance.

First, participants were expected to perform well on the 2-item tracking tasks for both change accuracy and tracking accuracy. Change and tracking accuracy was significantly better for 2-item trials compared to 4-items trials. These findings support previous research results that two items can be effectively tracked for changes (Wolfe, Place, & Horowitz, 2007; Wu & Wolfe, 2016). However, when the tracked number of items is raised to 4 performance in both change and tracking accuracy is reduced significantly. Tracking four items does not seem like an unreasonable number to keep track of during these types of tasks. Change accuracy for tracking 4
items was 53% \((SD = 20.68)\) and 32% \((SD = 18.01)\), for high and low prevalence respectively. These results do not support the claims of previous research that 4 items can be effectively tracked (Cowan, 2001). Discovering what contributes to one’s ability to effectively track items and recognize changes in them is an important endeavor to help better prepare individuals who use these skills in real world situations. More research manipulating the parameters of MOT tasks need to be conducted to assess the cognitive limits of individuals in these tasks to better understand the limits of these psychological processes.

Second, a novel contribution from this study was that there was a main effect of prevalence in the experimental paradigm of assessing change detection accuracy in a MOT task. These findings were predicted based on previous research assessing prevalence effects on identification of items of interest, but in static images (Wolfe, Horowitz, & Kenner, 2005; Wolfe et al., 2007; Wolfe et al., 2013; Evans, Birdwell, & Wolfe, 2013; Hout et al., 2015). Since low prevalence events happen infrequently, the prevalence effects found in the high prevalence MOT tasks trials will not aid in the recognition of real-world low prevalence events. The results, however, provide a basis for further studies to examine prevalence effects of change detection in MOT tasks. Further studies should be conducted to verify the prevalence effects found since some of the dependent variables were not normally distributed. However, finding significant effects in such a small sample is a positive sign for the experimental paradigm moving forward to assess prevalence of changes and their identification in MOT tasks.

The absence of a main effect of gaming experience could have been due to several factors. The non-significant effects of gaming experience could be caused due to the small sample of participants in this study. This would allow for follow up studies to be conducted to continue to assess the effects of action video game experience in sustained MOT tasks. Another
explanation could be that action video game experience may not help develop the sustained cognitive abilities associated with this task. Previous studies assessing the cognitive abilities of AVGPs have used shorter trials (Green & Bavelier, 2003; Green & Bavelier, 2006; Clark et al., 2011). Given the hours of action video games per week that participants needed to qualify as an AVGP, participants can play those games for extended periods of time. Further studies should therefore continue assessing the cognitive benefits, specifically benefits that endure over long periods of use, in future sustained research tasks.

It is important to acknowledge the limitations of this study due to the small sample size. Previous research including AVGPs has matched genders only when it was possible (Green & Bavelier, 2003) or excluded female gamers entirely (Clark et al., 2011), because it is very difficult to locate enough female gamers to have equal distribution of gamers by gender. Given the circumstances for this study, it was not possible to match genders in the two groups, which was originally planned for this research. It would have been difficult to parse out how exactly gender had influenced the data and caused these varying results from the initial analyses. Therefore, gender was not included in the analyses of the data for this study. Following studies should make special considerations to ensure that gender is addressed and controlled for in future research to limit these types of issues, especially if a small sample is being used, as the data sample was in the current study.

The final takeaway results from this research was from the follow up task, which replicated findings of previous research which indicated a relationship between action video game experience and enhanced performance in change detection tasks (Clark et al., 2011). Even though the follow up task was brief and only used 8 images with 4 change detection trials, there was still a significant difference in change detection based on AVGP experience. The non-
significant results for the effect of most recent prevalence type on the follow up change detection task suggest that short term prevalence experience may not help in identifying changes in later tasks. There were also no interactions found for experience with most recent prevalence type. A possible explanation for the lack of prevalence effects could be that the rates of prevalence used in this study were not at proper rates to observe any benefit. The inclusion of prevalence to change detections in MOT tasks in this study was a novel contribution to the field MOT research. For this reason, the rates in which high and low frequency events occurred may not have been at ideal levels in order to induce a prevalence effect for change detections in later tasks. One possibility is that this study may have generated an adverse effect of attentional overload, rather than the desired beneficial prevalence effect that was predicted. Tracking items in the high prevalence trials could have created some attentional overload where participants were not able to monitor items effectively due to the task being too taxing on cognitive resources (Coiera, Tombs, & Clutton-Brock, 1996). Further studies should be conducted to further investigate the possibility of an optimal rate of change in MOT tasks to create ‘prevalence training effects’ in later tasks. This is an important area to conduct further research in to further assess the possibility of prevalence training as a plausible avenue of training for surveillance tasks.

Taken together, despite a small pool of data, there were still some interesting implications that arose from the data analysis. The replication of main effects of item and prevalence in the MOT task demonstrate that the novel experimental paradigm of change detection in MOT tasks involving prevalence has merit and builds upon previous research and their findings. Video game experience may not have affected performance in the MOT task, but given the possibility of a potential design flaw creating attentional overload when tracking 4 items in high prevalence
trials, there is a possibility that follow up studies refining experimental parameters and design may find results the current study failed to generate. Future follow up studies should address some of the attentional overload issues, by assessing how well 3 items are tracked, as well as adjusting the rates of prevalence which were created without any prior literature to inform what proper rates should be for either condition. Overall, this study provides a new avenue in the field of MOT and change detection research by better replicating situations in which surveillance is conducted while assessing change detection and effective surveillance in a laboratory setting.
References


Figures

Figure 1

*Predicted MOT Tracking & Change Accuracy*

![Graph showing predicted MOT tracking and change accuracy. The blue solid lines represent AVGP (Action Video Game Player) scores, and the solid red lines represent NVGP (Non-Action Video Game Player) scores.](image)

*Note:* Predicted trends in data for rate of both tracking and change accuracy. The blue solid lines represent AVGP (Action Video Game Player) scores, and the solid red lines represent NVGP (Non-Action Video Game Player) scores.
Figure 2

*Predicted Rate of Change Trials Correctly Identified*

![Graph showing predicted rate of change trials correctly identified](image)

**Note:** Predicted result trends for rate of change detection accuracy in the follow up change detection task. The blue solid lines represent AVGP (Action Video Game Player) scores, and the solid red lines represent NVGP (Non-Action Video Game Player) scores.
Figure 3

*MOT Sample Stimuli*

Note: Samples of the stimuli used for the MOT task. Demonstrating how changes and masking of items during each trial occurred. (Wu & Wolfe, 2016)
Figure 4

*Change Detection Task Sample Stimuli*

*Note:* Sample cycle of a change trial. (Clark, Fleck, & Mitroff, 2011)
**Figure 5**

*MOT Change Accuracy Data*

*Note:* Change accuracy data for the MOT task. Change accuracy scores are the percent accuracy in identifying changes throughout each trial. The blue solid lines represent AVGP (Action Video Game Player) scores, and the solid red lines represent NVGP (Non-Action Video Game Player) scores. Error Bars: +/- 1 SE.
Figure 6

*MOT Tracking Accuracy Data*

Note: Tracking accuracy data for the MOT task. Tracking accuracy scores are the percent accuracy in identifying tracked items at the conclusion of each trial. The blue solid lines represent AVGP (Action Video Game Player) scores, and the solid red lines represent NVGP (Non-Action Video Game Player) scores. Error Bars: +/- 1 SE.
Figure 7

*Change Detection Task Accuracy Data*

Note: Accuracy in the change detection task. The scores are the accuracy rate of correctly identifying changes in the change trials. The blue solid lines represent AVGP (Action Video Game Player) scores, and the solid red lines represent NVGP (Non-Action Video Game Player) scores. Error Bars: +/- 1 SE.
## Appendix A – Video Game Questionnaire

### Video Game Playing Questionnaire – *DURING THE PAST YEAR*

<table>
<thead>
<tr>
<th>Ss ID: __________________________</th>
<th>Date: __________________________</th>
</tr>
</thead>
</table>

For each category of games, please rate:

1. Your estimated EXPERTISE in that category (1 = lowest, 7 = highest) – even if no experience, how do you think you would perform, compared to the general public?
2. Your average HOURS PER WEEK in that category for the past 12 months.
   - Ex: If you play 1.5 hrs/week, mark "1+ to 3"
3. The games you played and how old you were when you played them most

### ACTION_FIRST/THIRD PERSON SHOOTERS
(Call Of Duty, Halo, Battlefield, Half-Life, Overwatch, Counterstrike, ...

<table>
<thead>
<tr>
<th>Expertise: 1 2 3 4 5 6 7</th>
<th>Hours per week:</th>
<th>Games played most over the past year:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Never 0+ to 1</td>
<td>1+ to 3 3+ to 5 5+ to 10 10+</td>
</tr>
</tbody>
</table>

### ACTION_RPG/ADVENTURE
(The Witcher, Mass Effect, Fallout 4, Skyrim, GTA, Assassin’s Creed, Tomb Raider, The Last of Us, ...

<table>
<thead>
<tr>
<th>Expertise: 1 2 3 4 5 6 7</th>
<th>Hours per week:</th>
<th>Games played most over the past year:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Never 0+ to 1</td>
<td>1+ to 3 3+ to 5 5+ to 10 10+</td>
</tr>
</tbody>
</table>

### SPORTS/DRIVING
(Fifa, NHL, Mario Kart, Need for Speed, Forza, ...

<table>
<thead>
<tr>
<th>Expertise: 1 2 3 4 5 6 7</th>
<th>Hours per week:</th>
<th>Games played most over the past year:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Never 0+ to 1</td>
<td>1+ to 3 3+ to 5 5+ to 10 10+</td>
</tr>
</tbody>
</table>

### REAL-TIME STRATEGY/MOBA
(Starcraft, Warcraft (old ones: I, II & III), DotA, Command & Conquer, League of Legends, Age of Empires, ...

<table>
<thead>
<tr>
<th>Expertise: 1 2 3 4 5 6 7</th>
<th>Hours per week:</th>
<th>Games played most over the past year:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Never 0+ to 1</td>
<td>1+ to 3 3+ to 5 5+ to 10 10+</td>
</tr>
</tbody>
</table>

### NON-ACTION TURN-BASED ROLE-PLAYING/FANTASY
(Final Fantasy, Fable, Pokemon, Dragon Age, ...

<table>
<thead>
<tr>
<th>Expertise: 1 2 3 4 5 6 7</th>
<th>Hours per week:</th>
<th>Games played most over the past year:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Never 0+ to 1</td>
<td>1+ to 3 3+ to 5 5+ to 10 10+</td>
</tr>
</tbody>
</table>

### TURN-BASED STRATEGY/LIFE SIMULATION/PUZZLE
(Civilization, Hearthstone, The Sims, Restaurant Empire, Puzzle Quest, Bejeweled, Solitaire, Candy Crush, ...

<table>
<thead>
<tr>
<th>Expertise: 1 2 3 4 5 6 7</th>
<th>Hours per week:</th>
<th>Games played most over the past year:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Never 0+ to 1</td>
<td>1+ to 3 3+ to 5 5+ to 10 10+</td>
</tr>
</tbody>
</table>

### MUSIC GAMES
(Guitar Hero, Dance Dance Revolution, Rock Band, ...

<table>
<thead>
<tr>
<th>Expertise: 1 2 3 4 5 6 7</th>
<th>Hours per week:</th>
<th>Games played most over the past year:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Never 0+ to 1</td>
<td>1+ to 3 3+ to 5 5+ to 10 10+</td>
</tr>
</tbody>
</table>

### OTHER
(Games that don’t fit into any other category; Phone games, Browser games, Fighting games, etc.)

<table>
<thead>
<tr>
<th>Expertise: 1 2 3 4 5 6 7</th>
<th>Hours per week:</th>
<th>Games played most over the past year:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Never 0+ to 1</td>
<td>1+ to 3 3+ to 5 5+ to 10 10+</td>
</tr>
</tbody>
</table>
Video Game Playing Questionnaire – *BEFORE THE PAST YEAR*

Ss ID: ____________________________ Date: ____________________________

For each category of games, please rate:

1. Your estimated EXPERTISE in that category (1 = lowest, 7 = highest) – even if no experience, how do you think you would perform, compared to the general public?
2. Your average HOURS PER WEEK in that category for the past 12 months.
   Ex: If you play 1.5 hrs/week, mark “1+ to 3”
3. The games you played and how old you were when you played them most

**ACTION_FIRST/THIRD PERSON SHOOTERS** (Call Of Duty, Halo, Battlefield, Half-Life, Overwatch, Counterstrike ...)
Expertise: 1 2 3 4 5 6 7 Hours per week: *Never 0+ to 1 1+ to 3 3+ to 5 5+ to 10 10+*

Games played most over the past year:

**ACTION_RPG/ADVENTURE** (The Witcher, Mass Effect, Fallout 4, Skyrim, GTA, Assassin's Creed, Tomb Raider, The Last of Us, ...)
Expertise: 1 2 3 4 5 6 7 Hours per week: *Never 0+ to 1 1+ to 3 3+ to 5 5+ to 10 10+*

Games played most over the past year:

**SPORTS/DRIVING** (F1, NHL, Mario Kart, Need for Speed, Forza, ...)
Expertise: 1 2 3 4 5 6 7 Hours per week: *Never 0+ to 1 1+ to 3 3+ to 5 5+ to 10 10+*

Games played most over the past year:

**REAL-TIME STRATEGY/MOBA** (Starcraft, Warcraft (old ones: I, II & III), Dota, Command & Conquer, League of Legends, Age of Empires, ...)
Expertise: 1 2 3 4 5 6 7 Hours per week: *Never 0+ to 1 1+ to 3 3+ to 5 5+ to 10 10+*

Games played most over the past year:

**NON-ACTION TURN-BASED ROLE-PLAYING/FANTASY** (Final Fantasy, Fable, Pokemon, Dragon Age, ...)
Expertise: 1 2 3 4 5 6 7 Hours per week: *Never 0+ to 1 1+ to 3 3+ to 5 5+ to 10 10+*

Games played most over the past year:

**TURN-BASED STRATEGY/LIFE SIMULATION/PUZZLE** (Civilization, Hearthstone, The Sims, Restaurant Empire, Puzzle Quest, Bejeweled, Solitaire, Candy Crush, ...)
Expertise: 1 2 3 4 5 6 7 Hours per week: *Never 0+ to 1 1+ to 3 3+ to 5 5+ to 10 10+*

Games played most over the past year:

**MUSIC GAMES** (Guitar Hero, Dance Dance Revolution, Rock Band, ...)
Expertise: 1 2 3 4 5 6 7 Hours per week: *Never 0+ to 1 1+ to 3 3+ to 5 5+ to 10 10+*

Games played most over the past year:

**OTHER** (Games that don’t fit into any other category, Phone games, Browser games, Fighting games, etc.)
Expertise: 1 2 3 4 5 6 7 Hours per week: *Never 0+ to 1 1+ to 3 3+ to 5 5+ to 10 10+*

Games played most over the past year:
Appendix B – AVGP & NVGP Criteria

Video Game Questionnaire – Version March 2018

**NVGP criteria**

- **During** the Past Year
  - (1) At most 0-1 hour per week for Action_First/Third
  - (2) At most 0-1 hour per week for Action_RPG/Adv.
  - (3) At most 0-1 hour per week for Sport/Driving
  - (4) At most 0-1 hour per week for RTS/MOBA
  - (5) No more than 1-3 hours per week for each of the remaining game categories (Non-Action_RPG/FANTASY, Turn Based Strategy/LifeSim/Puzzle, Music games and Other)
  - (6) Using the maximum value in each hourly category (i.e. 0-1 = 1 hour; 1-3 = 3 hours), no more than 5 hours per week total across all game categories

- **Before** the Past Year
  - (1) At most 0-1 hour per week for Action_First/Third
  - (2) At most 0-1 hour per week for Action_RPG/Adv.
  - (3) At most 0-1 hour per week for Sport/Driving
  - (4) At most 0-1 hour per week for RTS/MOBA
  - (5) No more than 1-3 hours per week for any of the remaining game categories (Non-Action_RPG/FANTASY, Turn Based Strategy/LifeSim/Puzzle, Music games and Other)
  - (6) Using the maximum value in each hourly category (i.e. 0-1 = 1 hour; 1-3 = 3 hours), no more than 5 hours per week total across all game categories
AVGP criteria
Four options possible. All criteria MUST apply within each option

OPTION 1.
(1) At least 5+ hours for Action_First/Third during past year
OR
(2) At least 5+ hours for Action_RPG/Adv. during past year
AND
(3) At most 1-3 hours during past year on each game category listed: Turn Based Strategy/LifeSim/Puzzle, Music games and Other (that is all other game categories but RTS/MOBA and Sport/Driving)

OR

OPTION 2.
(1) 3-5 hours for Action_First/Third during past year
OR
(2) 3-5 hours for Action_RPG/Adv. during past year
AND
(3) At most 1-3 hours during past year on each game category listed: Turn Based Strategy/LifeSim/Puzzle, Music games and Other (that is all other game categories but RTS/MOBA and Sport/Driving)
AND
(4) At least 5+ hours on Action_First/Third before past year
OR
(5) At least 5+ hours for Action_RPG/Adv. before past year

OR

OPTION 3.
(1) 3-5 hours for Action_First/Third during past year
OR
(2) 3-5 hours for Action_RPG/Adv. during past year
AND
(3) At least 5+ hours in Sport/Driving during past year
AND
(4) At most 1-3 hours on each remaining game category during past year: Turn Based Strategy/LifeSim/Puzzle, Music games and Other (that is all other game categories but RTS/MOBA)

OR

OPTION 4.
(1) 3-5 hours for Action_First/Third during past year
OR
(2) 3-5 hours for Action_RPG/Adv. during past year
AND
(3) At least 5+ hours in RTS/MOBA during past year
AND
(4) At most 1-3 hours on each remaining game category during past year: Turn Based Strategy/LifeSim/Puzzle, Music games and Other (that is all other game categories but Sport/Driving)
**Tweener criteria**

Participants do not fit into any categories (they are not pure NVGP nor are they experts in any game genres that we sample including others)

1. do not qualify as a NVGP
   AND
2. Play strictly less than 5 hours in all of the game genres listed in VG questionnaire, for both During the past year AND before the past year: