Memory abilities in action video game players

Ashley F. McDermott a, Daphne Bavelier a,b, C. Shawn Green a,c,*

a Brain & Cognitive Sciences, University of Rochester, Meliora Hall, Box 270268, Rochester, NY 14627-0268, United States
b Psychology Section, FSEE, University of Geneva, Geneva, Switzerland
c Psychology Department, University of Wisconsin-Madison, Brogden Psychology Building, 1202 W. Johnson St., Madison, WI 53706, United States

A R T I C L E   I N F O
Article info:
Available online 21 February 2014
Keywords:
Action video games
Memory
Speed-accuracy trade-off

A B S T R A C T
Action video game players (AVGPs) have been shown to outperform non-action video game players (NVGPs) on tasks of perception and attention. Here we set out to investigate if these benefits also extend to another cognitive domain, memory. In particular, while there is previous evidence suggesting AVGPs demonstrate better visual short-term memory, it is unclear whether this extends to long-term memory processes or indeed, whether these enhancements are due to memory per se or are instead reflective of changes in speed of processing or strategy. Using four tasks that tap distinct areas of memory processing we found evidence for greater speed of processing and enhanced visual short-term memory in AVGPs and compared to NVGPs. However, we found either equivalent or possibly decreased performance in AVGPs in tasks related to long-term memory access. Furthermore, differences in strategy were noted across tasks, in particular differences in the tradeoff between speed and accuracy, which calls for a closer investigation of how task instructions bias performance in future studies.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

From the time that video games entered mass culture in the mid-to-late 1980s and early 1990s there has been significant scientific interest in the possible perceptual and cognitive effects of video game play (De Lisi & Cammarano, 1996; Dorval & Pepin, 1986; Drew & Waters, 1986; Dustman, Emmerson, Steinhaus, Shearer, & Dustman, 1992; Gagnon, 1985; Goldstein et al., 1997; Greenfield, 1984; Greenfield, DeWinstanley, Kilpatrick, & Kaye, 1996; Lowery & Knirk, 1982; McClurg & Chaille, 1987; Okagaki & French, 1994; Subrahmanyan & Greenfield, 1994; Yuij, 1996). In particular, over the past decade a significant body of research has documented the beneficial effects of playing one particular sub-genre of games known as “action video games” (see Green & Bavelier, 2008, 2012; Spence & Feng, 2010 for reviews; or Powers, Brooks, Aldrich, Palladino, & Alfieri, 2013 for a recent meta-analysis). Playing these fast-paced, perceptually, attentionally, cognitively, and motorically demanding games has been shown to enhance a variety of abilities from sensory and perceptual skills (Donohue, Woldorf, & Mittroff, 2010; Li, Polat, Makous, & Bavelier, 2009) to attentional skills (Chisholm, Hickey, Theeuwes, & Kingstone, 2010; Chisholm & Kingstone, 2012; Green & Bavelier, 2003, 2006a, 2007; West, Stevens, Pun, & Pratt, 2008) to more cognitive skills (Barlett, Vowels, Shanteau, Crow, & Miller, 2009; Boat, Kramer, Simons, Fabiani, & Gratton, 2008; Cain, Landau, & Shimamura, 2012; Clark, Fleck, & Mitroff, 2011; Colzato, van den Wildenberg, Zmigrod, & Hommel, 2012; Colzato, van Leeuwen, van den Wildenberg, & Hommel, 2010; Feng, Spence, & Pratt, 2007; Ferguson, Cruz, & Rueda, 2007; Green, Sugarman, Medford, Klobusicky, & Bavelier, 2012; Karle, Watter, & Shedden, 2010; Strobach, French, & Schubert, 2012) and even to decision making (Green, Pouget, & Bavelier, 2010). Furthermore, the neural underpinnings of many of these effects are beginning to be uncovered (Bavelier, Achtman, Mani, & Foecker, 2012; Koepp et al., 1998; Krishnan, Kang, Sperling, & Srinivasan, 2012; Mishra, Zinni, Bavelier, & Hillyard, 2011). Action video game play has also been repeatedly shown to result in faster reaction times (Bialystok, 2006; Castel, Pratt, & Drummond, 2005; Dye, Green, & Bavelier, 2009b; Nelson & Strachan, 2009). And although the majority of the work has been conducted using college-aged individuals, there is evidence that similar effects can be induced in both children and the elderly (Belchior, 2007; Dye, Green, & Bavelier, 2009a; Olson, 2010; Trick, Jaspers-Fayer, & Sethi, 2005). Critically, a causal relationship between the observed effects and action game training has been repeatedly demonstrated (Cohen, Green, & Bavelier, 2007; Feng et al., 2007; Green & Bavelier, 2003, 2006a, 2007; Green et al., 2010; Li et al., 2009; Spence, Yu, Feng, & Marshman, 2009), thus leading to significant interest in the potential of using action games for rehabilitative or training...
purposes (Li, Ngo, Nguyen, & Levi, 2011; Rosser et al., 2007; Schlic-kum, Hedman, Enochsson, Kjellin, & Fellander-Tsai, 2008; Schlic-kum, Hedman, Enochsson, Kjellin, & Fellander-Tsai, 2009).

However, an area that remains to be thoroughly explored con-cerns the impact of action video game play on aspects of memory. Here the available research is much sparser. Sungur and Boduroglu (2012) recently reported that avid action video game players (AVGPs) form more accurate memory representations of objects than NVGPs. In a short-term memory task, AVGPs were able to identify the color of an object on a color wheel more precisely than NVGPs. It was also found that even when AVGPs made errors in reporting the location of the peripheral target in a visual search task, their errors were closer to the correct location, indicating that they had a better representation of the target location than those who do not play action games (non-action video game players – NVGPs). Using a standard N-back task, Boot et al. (2008), saw faster RTs, but equivalent accuracy in AVGPs, while Colzato et al. (2012) reported both higher accuracy and faster RTs in AVGPs in 1-back and 2-back versions of the N-back task. And finally, both Green and Bavelier (2006b), using an object enumeration paradigm, and Boot et al. (2008), using the Luck and Vogel (1997) change detec-tion task, reported a benefit in visual short-term memory in AVGPs.

The goal of this paper is to add to our understanding of the ef-fects of action game play on memory by examining AVGP perform-ance in four tasks thought to tap fundamentally different aspects of memory (noting here and throughout the paper that although previous studies have specifically demonstrated a causal link between differences in cognitive function and action game experience via training studies, the present study utilizing AVGPs and NVGPs can only demonstrate a correlation that would be con-sistent with such a causal relationship and not a causal link in and of itself). In particular, it is unclear a priori whether action games truly place demands on the memory system (indeed, it is more common in this genre for information to be readily available either on the screen, in menus, or via various types of epistemic actions). Thus, the tasks that were selected allow the assessment of not just memory capacity, but also speed of access to memory traces, as well as the ability to flexibly manipulate or inhibit memory traces.

Previous evidence suggests that action gaming may alter these pro-cesses during perceptual and attentional tasks, and thus an out-standing question is whether faster access, greater flexibility and more efficient inhibition of distractors in action gamers also ap-ply during memory processes.

The first of the four tasks utilized was the Posner letter identity task, which measures long-term memory access speed (Posner & Mitchell, 1967). Here, subjects were asked to make same/different judgments on two letters either on the basis of whether the two letters were physically identical (i.e. the same letter name and the same case) or on the basis of whether the two letters had the same letter name (but may differ in case). The Physical Identity task can be performed solely on the basis of the sensory input without the need to access long-term memory. The Name Identity task, on the other hand, requires retrieving the letter name from long-term memory. Long-term memory access time is thus typi-cally measured as the difference in reaction times between the two tasks (Posner & Mitchell, 1967).

The second task we utilized measured proactive interference (PI), or the ability to control the saliency of internal representations on the fly as task demands change (e.g., Kane & Engle, 2000; Rang-anath, Johnson, & D’Esposito, 2003). In day-to-day life, it is often necessary to distinguish the relevant memory for a given task from other related, but currently irrelevant memories. For example, when your phone number changes upon moving, you need to pro-vide the new number to friends and associates. Doing so requires giving them your current number rather than your old number or other phone numbers you have learned. In order to test PI, subjects saw a set of 4 letters and after a short delay were given a probe letter. The probe letter may have been in the set (a positive trial) or not (a negative trial). If the probe was not in the set of 4 letters, then it may have appeared in other sets seen recently, which would create interference. A subject’s ability to correctly re-pond on such trials indicates how well they can resolve the interference.

The third task employed a version of an N-back task that makes use of space to allow the probing of a range of N-back values. N-back tasks are a strong measure of working memory, or in other words, of the ability to both store and manipulate internal rep-re-sentations. As such, N-back tasks have been a paradigm of choice in studies that focus on improving working memory, as well as in establishing links between working memory and fluid intelli-gence (Jaeggi, Buschkuehl, Jonides, & Perrig, 2008; Unsworth, Brewer, & Spillers, 2009). Unlike traditional span tasks, where items simply need to be retained in memory, N-back tasks require updating of the information being held in memory with each added item. Whenever a new item is presented, the participant no longer needs to remember the item that was n + 1 back in the list, but needs to shift the position of each memory item, much like a sliding window that holds only n items.

Our fourth and final task was a short-term memory task similar to ones employed by Luck and colleagues (e.g. Vogel, Woodman, & Luck, 2001; Zhang & Luck, 2008) and used to estimate memory capacity. This type of task requires maintenance of information in memory, but does not require the additional manipulations needed in working memory tasks like the N-back. This allows for measuring capacity of memory in situations that only require re-call. This task also differed from the previous three in that accuracy rather than speed was the clear emphasis of the subjects (in the previous three tasks subjects needed to respond both accurately and as quickly as possible). In this task, subjects viewed a display of colored lines in different orientations, then after a delay were shown a probe from the display and asked to report whether or not the probe matched the item presented at the same location in the original display (change detection task).

2. Experiment 1: Posner letter identity task

The Posner letter identity task is a basic cognitive task that has traditionally been used to examine the relationship between speed of information processing and psychometric intelligence (Goldberg, Schwartz, & Stewart, 1977; Hunt, 1980; Keating & Bobbitt, 1978). In this task, subjects are presented with two letters that can differ in two ways – the letters can be either capital or lower-case and can be either the same letter or not. In one block of trials, the Physical Identity block, subjects have to respond as to whether the two letters are physically identical (i.e. have both the same let-ter name and the same capitalization) or not. In the second block, the Name Identity block, they are asked to report whether the let-ters have the same name or not. Thus, in the Physical Identity con-dition, subjects can make their decisions based entirely on the physical appearance of the stimuli. In the Name Identity condition however, they must retrieve the name of both of the letters from long-term memory. By subtracting the average RT in the Physical Identity block from the Name Identity block, this paradigm pro-vides an estimate of the time required to access long-term mem-ory, and more specifically, the speed of long-term memory access for highly over-learned information (e.g. letters – Posner & Mitch-ell, 1967). AVGPs have been found to have faster speed of process-ing in perceptual and attentional tasks (Dye et al., 2009b; Green et al., 2010), so it is a natural extension to investigate whether this population will also show faster information processing when asked to retrieve internal memory representations.
2.1. Methods

2.1.1. Participants
All participants were undergraduates at the University of Rochester and were recruited through electronic or physical advertisements and were paid for their participation in the experiment. Separate advertisements were distributed to recruit participants who had a lot of experience playing action video games and to recruit participants who did not have experience with action games. Once they arrived in the lab and before participating in the study, all participants filled out a survey about their video game playing habits, which allowed us to divide them into AVGPs and NVGPs at the time of subject recruitment. As in our past studies, AVGPs were defined as subjects who played 5 h or more per week of action video games for the past year (mean of 5–10 h per week). On average NVGPs played little to no video games and, crucially, did not play any action video games (played less than 1 h per week). They were also tested to verify that they had normal or corrected-to-normal vision. Some of the participants performed multiple tasks and overlap between subjects is reported in each experiment discussed in this paper. Three participants (1 NVGP, 2 AVGPs) consistently performed poorly on all tasks they completed (including experiments 1 and 3 presented here) and were excluded from all analyses. The exclusion criteria are reported for each experiment separately. Twenty-eight male AVGPs (mean age 19.8, played 5–10 h per week on average) and 25 male NVGPs (mean age 20.2, averaged less than 1 h of action video game play per week) participated in the Posner letter identity task. The sample here, and in subsequent experiments, is overwhelmingly male due to the relative rarity of female action game players. This necessity limits our ability to determine whether the effects are the same or different across the sexes. However, previous work suggests that females are, at a minimum, equally susceptible to video game training, and perhaps even show greater benefits of training than do males (Feng et al., 2007).

2.1.2. Design and procedure
The design of this task was based on the version of the Posner Name Identity task from Neubauer, Riemann, Mayer, and Angleitner (1997). The experiment consisted of 2 blocks comprised of 64 trials each with 10 practice trials preceding each block. In both blocks there was an inter-trial interval of 1 s followed by a fixation cross that appeared for 1 s, warning the subject that the next trial was about to begin. On each trial 2 letters were presented. Each letter was 1 deg of visual angle wide and 2 deg tall. The letters were located on either side of center and remained on the screen until the subject responded. The upper and lowercases of letters A, B, C, and E were used and the pair of letters could differ in Name Identity and in Physical Identity. More specifically, the letter pairs fit into one of three categories: “name same, physical same” (NSPS, i.e. ‘Aa’ or ‘b b’), “name same, physical different” (NSPD, i.e. ‘a A’ or ‘B b’), or “name different” (ND, i.e. ‘A B’ or ‘A B’). Trials were counterbalanced so that each of the 8 possible letters (4 identities, 2 cases) appeared in each side of a pair equally often and that the same letter was not in the same position for consecutive trials. Subjects responded by pressing the ‘s’ key for same and the ‘k’ key for different on a standard keyboard.

As in the Neubauer et al. (1997) study, the number of same and different trials was matched within each block. In the first block, subjects reported on the Physical Identity of the letters, so the block contained 32 NSPS trials, 16 NSPD trials, and 16 ND trials with all trial types intermixed. For the second block, subjects responded based on the Name Identity of the letters, resulting in 16 NSPS trials, 16 NSPD trials, and 32 ND trials. Subjects were instructed to be as fast and as accurate as possible.

2.2. Results

In addition to the 3 subjects excluded from all tasks, 1 NVGP and 1 AVGP were excluded from all analyses because they each had mean RTs that were slower than 2 standard deviations from the mean of the group. Accuracy analyses included all trials. Reaction time analysis focused on correct trials only, with anticipation (less than 100 ms) or clear lapses (greater than 3 s) removed (<.001% of the trials). We ran 3 (trial type: NSPS, NSPD, ND) × 2 (gamer status: AVGP, NVGP) × 2 (task: Physical Identity, Name Identity) mixed model ANOVAs on accuracy and on mean RTs respectively.

Accuracy data are reported in Table 1. There were no significant main effects of task, however a significant effect of trial type was found (F(2,98) = 12.258, p < .0005, η² = .200), with accuracy being highest in the NSPS condition. There was also a significant interaction between trial type and task (F(2,98) = 9.150, p < .0005, η² = .157). This interaction reflects lower accuracy when letters had the same name but different physical identities (NSPD condition) in the Name Identity task than in the Physical Identity task, while accuracy in the other trial types was comparable across the two tasks. There was no effect of gamer status on accuracy (p > .1), although there was a significant interaction between gamer status and task (F(1,49) = 4.477, p = .039, η² = .084) as AVGPs tended to be less accurate on the Physical Identity task than NVGPs, but more accurate than NVGPs on the Name Identity Task.

Mean RT data are reported in Table 1. The expected main effect of task was observed with the Name Identity task being slower than the Physical Identity task (F(1,49) = 8.441, p = .005, η² = .147). There was a main effect of trial type (F(2,98) = 34.027, p < .0005, η² = .410), as well as a significant interaction of trial type by task (F(2,96) = 14.786, p < .0005, η² = .232). RTs were shorter when physically identical letters were presented (NSPS), as expected. The difference between NSPS and the other two trial types was larger for the Name Identity task than the Physical Identity task, similar to what was found in the accuracy analysis. There were significant interactions between trial type and gamer status (F(2,98) = 3.225, p = .044, η² = .062) and trial type, gamer status, and task (F(2,98) = 3.105, p = .049, η² = .060) due to a larger difference between the groups in the NSPD condition for the Name Identity task. AVGPs also showed a trend to be faster in all conditions (overall means ± standard error: AVGPs = 542 ± 17 ms; NVGPs = 580 ± 21 ms; F(1,49) = 2.924, p = .094, η² = .056 – although, given the strong a priori hypothesis that AVGPs would be faster than NVGPs a one-tailed test would be appropriate here, we prefer to utilize two-tailed statistics throughout in order to be consistent with the previous literature).

As discussed earlier, one of the features of this task is that it provides a measure of memory access speed (Neubauer et al., 1997), which is estimated by computing the difference between overall reaction time on the Name Identity task and the Physical Identity task. We ran an independent-samples t-test on the memory access time measure (NI–PI) and found no difference between the two groups (p > .4). Because both accuracy and RT showed interactions based on trial type, we also calculated the difference between tasks for each condition, to test if the global measure might be masking some effects of access speed. We ran a 3 (trial type) × 2 (gamer status) ANOVA on the RT difference between NI and PI. We found an effect of trial type (F(2,98) = 14.786, p < .0005, η² = .232) and a trial type by gamer status interaction (F(2,98) = 3.105, p = .049, η² = .060), which was driven by the NSPD condition. In that condition, NVGPs showed the expected cost (~50 ms) associated with performing the Name Identity task versus the Physical Identity task, while AVGPs showed essentially no cost (~4 ms).
2.3. Discussion

The Posner letter identity task provides a measure of how efficiently participants can compare items based on their appearance versus using over-learned knowledge stored in memory, such as letter names. Overall we found the differences between conditions usually reported. Participants were faster on the Physical Identity than the Name Identity in line with the proposal that the latter requires additional time to access well-learned long-term memories for letter names. Trials in which the letters were physically identical (NSPS condition) were the easiest in both tasks, with the highest accuracy and the shortest RTs, which makes sense given that the answer would be the same whether participants used physical or name information. Participants were less accurate when the same letter displayed in different case was used (NSPD condition). For these stimuli, the answer differed depending on task – “same” for the Name Identity task and “different” for the Physical Identity task, which creates a conflict and interference from the other task instructions. This pattern of results has also been reported before. More relevant for our aim, overall faster reaction times in AVGPs were observed with accuracy remaining comparable across groups. The faster RTs for AVGPs suggest that they are faster to retrieve over-learned knowledge. Given that AVGPs retrieve letter names quickly, it may be that the lower accuracy in the NSPD condition for the Physical Identity task and no cost in RT for Name task on the NSPD condition are indications of AVGPs experiencing a Stroop effect, where the Name Identity is overriding the Physical Identity, which we do not see for NVGPs. This would be consistent with AVGPs having fast memory access speed, but it also suggests that AVGPs may be more susceptible to conflict, as conflicting memory traces are more readily accessed. In order to test if AVGPs have more difficulty resolving interference in memory, both populations were compared on a standard proactive interference task.

3. Experiment 2: Proactive Interference task

Proactive interference resolution is a key aspect of response inhibition and it is commonly tested using the proactive interference recent probes task (Jonides & Nee, 2006). Proactive interference arises when the memory for recent items or recent memories interfere with the currently relevant memory set. In Experiment 1 there was some evidence that AVGPs may be more susceptible to interfering information. There is also some evidence in the literature that AVGPs may differ from NVGPs when it comes to proactive control (Bailey, West, & Anderson, 2010). We further investigate this issue here with the recent probes task, which more directly tests for differences in proactive interference resolution.

3.1. Methods

3.1.1. Participants

Twenty-five AVGPs (1 female; mean age 19.8) and 23 NVGPs (1 female; mean age 20.0) were paid for their participation in this task. All were undergraduates at the University of Rochester and they were classified as AVGPs or NVGPs using the same criteria discussed above with AVGPs averaging 5–10 h per week of playing action video games and NVGPs playing 0 h per week. Eleven AVGPs and 17 NVGPs also participated in the Posner task and 1 AVGP and 1 NVGP were also included in the N-back task.

3.1.2. Design and procedure

This experiment was modeled after that used by Nelson, Reuter-Lorenz, Sylvester, Jonides, and Smith (2003), but the timing was adjusted to suit a behavioral study rather than a brain imaging study. Each trial was preceded by a fixation cross, presented for 500 ms, and then a 1500 ms blank followed by the 500 ms presentation of four lowercase consonants in a 2 × 2 grid. After the grid there was a 3 s blank before an uppercase consonant probe was presented in the center of the screen. Subjects made a keyboard press to indicate whether the letter had been in the previous grid or not. There were 96 positive (P) match trials where the probe had been in the previous grid. For the other 96 trials, the probe letter had not been presented in the previous grid; however, it could have been presented earlier. These negative match trials were divided into 4 categories:

- Unfamiliar (UF), in which the probe had been neither a stimulus nor probe in the previous two trials.
- Familiar (F), in which the probe had been a stimulus, though not a probe, in the previous trial, but (crucially) had been neither stimulus nor probe in the trial before the previous one.
- Highly familiar (HF), in which the probe had been a stimulus, though not a probe, in the previous trial and the trial before.
- Response conflict (RC), in which the probe had been a positive probe in the previous trial.

Trial type was balanced over the course of the experiment. Unfamiliar negative trials are expected to be a baseline and should not show any reductions in accuracy. Accuracy would be expected to decrease in the familiar negative trials and show a more pronounced difference in the highly familiar negative trials. The amount of interference in the response conflict trials is typically equal to that in the highly familiar trials (Nelson et al., 2003).

3.2. Results

One AVGP was removed from these analyses because of RTs on all conditions that were slower than 2 standard deviations from the mean of the group. RTs analyses were carried out on correct trials. There were no anticipations (RTs less than 100 ms) or lapses (RTs greater than 3000 ms). We ran a mixed model 2 (gamer status: AVGP, NVGP) × 5 (trial type: RC, UF, F, HF, P) ANOVA for accuracy and for mean RT. The results are summarized in Fig. 1.
There was a significant effect of trial type (accuracy: $F(4,184) = 20.409$, $p < .0005$, $\eta^2_g = .307$; RT: $F(4,184) = 14.324$, $p < .0005$, $\eta^2_g = .237$) for both accuracy and RT. There was an effect of gamer status for RT ($F(1,46) = 4.778$, $p = .034$, $\eta^2_g = .094$) and a marginal effect of gamer status for accuracy ($F(1,46) = 2.858$, $p = .098$, $\eta^2_g = .058$) with AVGPs performing the task more quickly with a trend for being less accurate than NVGPs (see Fig. 1). The interaction between trial type and gamer status was also marginal for accuracy ($F(4,184) = 2.388$, $p = .053$, $\eta^2_g = .049$) and significant for RT ($F(4,184) = 2.519$, $p = .043$, $\eta^2_g = .052$), reflecting again slightly faster RTs in AVGPs, especially when conflict needs to be resolved, with a trend for a greater error rate in AVGPs than NVGPs for response conflict trials.

Speed-accuracy tradeoffs are notoriously difficult to interpret, especially in tasks where performance is clustered near ceiling. To shed more light on the present speed-accuracy tradeoff, we normalized RTs by accuracy, a method used in the developmental literature (Ahkbar & Enns, 1989). This normalization is used to correct for baseline differences in RTs, for instance when comparing children, who are typically slower to respond, to adults. For each subject and condition, the RTs were transformed into proportions based on accuracy, such that normRT is equal to RT divided by proportion correct (Table 2). We ran a mixed model 2 (gamer status: AVGP, NVGP) x 5 (trial type: RC, UF, F, HF, P) ANOVA on these normalized RTs. The effect of trial type remained ($F(4,184) = 25.140$, $p < .0005$, $\eta^2_g = .353$), but there was no longer an effect of gamer status ($F(1,46) = 2.299$, $p = .136$, $\eta^2_g = .048$) and no interaction ($F(4,184) = .646$, $p = .630$, $\eta^2_g = .014$).

3.3. Discussion

When looking at proactive interference resolution, AVGPs were faster than NVGPs. AVGPs also appeared to answer at the same speed, regardless of the amount of interference, while NVGPs appeared to slow down as interference increased and could only approximate the speed of AVGPs when there was no interference (in the unfamiliar condition). The accuracy of AVGPs was affected by this speed, as they were less accurate on familiar and response conflict trials. When normalizing differences in speed by accuracy, the two populations did not differ significantly, suggesting a trade-off between speed and accuracy across populations in this experiment. Whether a speed-accuracy tradeoff should be considered as detrimental or beneficial to behavior is a challenging issue. In this present experiment, AVGPs show significantly greater speed but only a marginal loss in accuracy, making it difficult to conclude that there is actual poorer interference resolution in AVGPs. To address this issue further, future studies are needed that are designed in such a way that speed and accuracy can be better disentangled.

Proactive interference resolution is closely related to working memory abilities (Jonides & Nee, 2006) since using working memory often requires online interference resolution as well as short-term memory storage. By testing these two populations on an N-back task, which has both storage and interference resolution requirements, it will help to clarify whether there is a difference in susceptibility to interference between the two groups.

4. Experiment 3: N-back task

A task commonly used to study working memory is the N-back task, which requires subjects to report whether or not a current probe matches a probe presented n trials previously (Dahlin, Neely, Larsson, Backman, & Nyberg, 2008; Jaeggi et al., 2008; Li et al., 2008). This task therefore requires the participant to both remember a certain number of items, and also update which items should be remembered. Thus, effective performance requires the subject to inhibit the memory of recent trials that were more than n trials before the current trial, while enhancing the memory of the most recent probes. The version of the N-back we used was adapted from Verhaeghen, Cerella, and Basak (2004). The letter stimuli were presented in multiple columns, with the number of columns corresponding to the level of N-back studied. Typically N-back tasks take the form of 2- or 3-back; however, using the Verhaeghen et al. (2004) design allowed us to test a greater range of n’s, from the usual 1-back up to 7-back, while minimizing ceiling or floor effects.

4.1. Methods

4.1.1. Participants

We used the same screening procedures as discussed for Experiment 1. Thirteen AVGPs (1 female; mean age 19.9) and 11 NVGPs (1 female; mean age 20.7) were paid for their participation in our spatial N-back task. The AVGPs averaged playing 5–10 h per week of action video games and the NVGPs played 0 h of action games per week. One of the AVGPs participated in all tasks except the visual short-term memory task and 1 NVGP also participated in the visuospatial short-term memory task and the proactive interference task.

Table 2

<table>
<thead>
<tr>
<th>Trial type</th>
<th>normRTs in ms mean (std. error)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AVGPs</td>
</tr>
<tr>
<td>Positive</td>
<td>885 (48)</td>
</tr>
<tr>
<td>Unfamiliar</td>
<td>797 (43)</td>
</tr>
<tr>
<td>Familiar</td>
<td>849 (46)</td>
</tr>
<tr>
<td>Highly familiar</td>
<td>882 (45)</td>
</tr>
<tr>
<td>Response conflict</td>
<td>1002 (62)</td>
</tr>
</tbody>
</table>

Fig. 1. The accuracy and the reaction times for the proactive interference recent probes task. AVGPs have higher error rates than NVGPs, particularly in the response conflict condition (a), as well as faster RTs in all conditions (b). The conditions are labeled as P = positive, UF = unfamiliar, F = familiar, HF = highly familiar, RC = response conflict.
4.1.2. Design and procedure

An N-back task was adapted from the work of Verhaeghen et al. (2004). The set of possible stimuli were all 20 consonants in the English alphabet. Stimuli were arranged in a two-dimensional grid with the number of columns varying (from 1 to 7) as a function of the value of $n$ in the N-back. The number of columns ($n$) corresponded to working memory load, so the subjects had to respond whether the letter was the same as the previous letter that appeared within that column, which would have appeared $n$ trials earlier (Fig. 2). In each grid, there were $20 + n$ stimuli. In other words, the first letter of each column appeared, which did not require a response, then $20$ other letters appeared across the columns, requiring subjects to make a total of 20 responses, regardless of the number of columns.

Subjects were instructed to be as fast and as accurate as possible.

First, participants practiced one block with only one column and one block with three columns. If a participant required more practice (2 subjects did not initially understand the task), the practice session of 2 blocks was repeated. Participants then completed 2 experimental sessions. Each session consisted of 35 blocks—5 at each N-back level. The N-back levels increased throughout the session, starting with 5 blocks of one column, then 5 of two columns, and so forth. Between sessions participants completed other, unrelated tasks that totaled 12 min in length.

4.2. Results

Two subjects (1 AVGP, 1 NVGP) were removed because their reaction times were more than 2 standard deviations slower than the mean of their respective groups. RTs analyses were carried out on correct trials only after removal of anticipation (RTs less than 100 ms) and lapses (RTs greater than 2700 ms). Analyses proceeded by running separate 2 (gamer status: NVGP, AVGP) $\times 7$ (number of columns: 1–7) $\times$ 2 (session) ANOVAs for accuracy and for mean RTs.

As expected, for accuracy there was a main effect of number of columns ($F(6,120) = 51.756, p < .0005, \eta^2_p = .721$) with accuracy declining as the number of columns, and thus the length of the N-back, increased (Fig. 3). No other effect was significant ($ps > .1$).

For RTs, the expected main effect of number of columns was present ($F(6,120) = 21.677, p < .0005, \eta^2_p = .520$). There was also a significant interaction between session and columns ($F(5,100) = 2.346, p = .035, \eta^2_p = .105$) whereby subjects increased their response speed more for higher numbers of columns in the second session. No other effects were significant ($p’s > .1$).

In tasks where RT is measured, AVGPs are typically consistently faster than NVGPs. Assuming this a priori hypothesis led us to inspect the RT data more closely. It appeared that there was a difference between the two groups for all but the easiest column. To investigate this, we ran additional ANOVAs for accuracy and RT, excluding the 1 column (1-back) condition. The findings for accuracy did not change, with no effect of session or gamer status and the expected effect of number of columns remaining ($F(5,100) = 56.445, p < .0005, \eta^2_p = .738$). In the RT analysis, the effects of session ($F(1,20) = 46.158, p < .0005, \eta^2_p = .698$) and number of columns ($F(5,100) = 3.854, p = .035, \eta^2_p = .162$) both also remained. In addition, there was a trend toward the expected significant effect of gamer status (overall mean $\pm$ standard error RTs: AVGPs = 930 $\pm$ 36 ms; NVGPs = 1062 $\pm$ 72 ms; $F(1,20) = 3.027, p = .096, \eta^2_p = .131$) with AVGPs being faster than NVGPs.

4.3. Discussion

This variant of the N-back task allowed us to compare how AVGPs and NVGPs manipulate and update working memory, starting with an easy task and increasing difficulty through loads that are higher than what humans can usually manage. Accuracy analyses revealed equal accuracy between groups, while RT analyses indicated that, as expected, AVGPs responded faster than NVGPs...
on this task, with this difference being masked when performance is near ceiling in the 1 column (1-back) condition (unlike the task utilized by Colzato et al. (2012), where subjects were not at ceiling for either 1-back or 2-back). This result mirrors the trend found by Boot et al. (2008) on their N-back task, where they found that AVGPs and NVGPs were equally accurate and the AVGPs showed a trend to be faster.

The N-back task requires juggling multiple memory abilities, so it is difficult to parse which of these, if any, may be responsible for the slightly faster reaction times in AVGPs. It is possible that AVGPs greater speed may be due to an initial faster perceptual decoding of each stimulus presented, calling for studies that better partition the different components of memory. With this caveat in mind, we turned to a test of memory capacity where speed is not of concern.

5. Experiment 4: Visual short-term memory task

The previous three tasks required subjects to be as fast and as accurate as possible. Under such conditions, AVGPs are faster than NVGPs, but also display marginal loss in accuracy at times. Such speed-accuracy tradeoffs are difficult to interpret, as it remains unclear whether, when freed of timing constraints, AVGPs would indeed show greater memory capacity, or rather would keep showing trends for lesser accuracy. To address this issue, we used a visual short-term memory task commonly used to estimate short-term memory capacity (Awh, Barton, & Vogel, 2007; Vogel et al., 2001; Zhang & Luck, 2008). Using a variant of this task, Boot et al. (2008) found differences between AVGPs and NVGPs. In addition several other results point to an advantage in AVGPs in visual short term memory tasks as witnessed by (i) an AVGP advantage in the multiple object tracking task (Green & Bavelier, 2006b; Sun, Ma, Bao, Chen, & Zhang, 2008), (ii) more accurate enumeration for high numerosity displays (Green & Bavelier, 2003, 2006b), (iii) or better memory for the color of objects (Sun et al., 2008).

5.1. Methods

5.1.1. Participants

Twelve AVGPs and 13 NVGPs (all male) were paid for their participation. Five of the NVGPs also participated in the Posner letter identity task and 3 of those participated in the proactive interference.

5.1.2. Design and procedure

The design and procedure was based on Vogel et al. (2001). The participant initiated each trial by pressing the spacebar. Then they were presented with a display containing colored, oriented lines for 1000 ms. There were 8 possible set sizes (1, 2, 3, 4, 5, 6, 7, 8, or 10) and the lines appeared at random locations within a 5 x 5 position grid. Each line was drawn in 1 of 4 colors (red, green, blue, or white) with 1 of 4 possible orientations (0°, 45°, 90°, or 135°). This was followed by a blank screen for 900 ms, then a single probe line was displayed. The probe was in the same position as one of the lines in the first display. On 50% of the trials, one feature (color or orientation) of the probe was altered. Participants indicated whether they thought the probe had changed or not. The response display screen remained on until the participant answered or for 2000 ms. Subjects were instructed to be as accurate as possible; they were explicitly told not to worry about speed, but rather focus on accuracy.

Participants were asked to report changes in color during one block and changes in orientation during the other, with the order of the blocks counterbalanced between participants. Set sizes were balanced across each of two blocks, with 16 trials at each of the 9 set sizes. The task irrelevant feature of the probe never changed. The participant pressed the ‘n’ key (labeled ‘Y’ for yes) to indicate a change, and the ‘m’ key (labeled ‘N’ for no) to indicate no change.

5.2. Results

We removed 1 NVGP due to consistently poor performance — a sensitivity index (d’) of zero for all set sizes higher than 3, and comparatively low d’ even at very small set sizes. Performance between groups were compared by first analyzing d’ in a 2 (condition: color or orientation) x 8 (set size) x 2 (group: NVGP, AVGP) ANOVA. An expected main effect of set size was found (F(7,154) = 39.214, p < .0005, ηp2 = .641) with d’ decreasing as set size increased. The main effect of group was not significant, but there was a significant set size x group interaction (F(7,154) = 2.547, p = .017, ηp2 = .104) indicating an advantage in AVGPs as set size increases. A main effect of condition was also present with subjects being more accurate in the color condition than the orientation condition (F(1,22) = 70.960, p < .0005, ηp2 = .763) as well as an interaction between condition and set size (F(7,154) = 4.345, p < .0005, ηp2 = .165) but no interaction with group (p > .1). All other effects were nonsignificant (ps > .1) (see Fig. 4).

Between groups performance was also compared using the memory capacity index K (k = set size x (Hit Rate–False Alarm Rate)) as in Vogel et al. (2001). A 2 (condition) x 8 (set size) x 2 (group) ANOVA indicated main effects of condition (F(1,22) = 48.162, p < .0005, ηp2 = .686) and set size (F(7,154) = 30.318, p < .0005, ηp2 = .579), as well as interactions of condition x set size.
was entirely on accuracy, AVGPs were more accurate than NVGPs. In the visual short-term memory task, where the subjects’ focus was unimportant, is consistent with the broader literature on AVGPs (i.e. faster RT and equivalent accuracy shown in tasks with a strong RT component, higher accuracy in tasks with no RT component).

5.3. Discussion

Using a visuospatial short-term memory task designed to assess span, AVGPs were found to outperform NVGPs. AVGPs were more accurate than NVGPs for set sizes that are higher than 4, the typical memory capacity found in these tasks (Cowan, 2001). This suggests that AVGPs have better visuospatial memory performance than NVGPs. The fact that AVGPs showed higher accuracy in this task, where RT was unimportant, is consistent with the broader literature on AVGPs (i.e. faster RT and equivalent accuracy shown in tasks with a strong RT component, higher accuracy in tasks with no RT component).

6. Conclusions and general discussion

The goal of this paper was to add to our understanding of the effects of action video game experience on memory. The Posner letter identity task revealed that AVGPs access well-learned memories faster than NVGPs, but this faster access appears to come with an increase in some interference effects. However, in the proactive interference task, the majority of the differences between AVGPs and NVGPs appear to be due to differences in the relative weighting of speed versus accuracy, rather than true differences in the resolution of conflict per se. In the N-back task AVGPs were again faster, but there were no group differences in accuracy, while in the visual short-term memory task, where the subjects’ focus was entirely on accuracy, AVGPs were more accurate than NVGPs.

Overall, in the tasks where subjects were asked to perform the task quickly, AVGPs were consistently faster than NVGPs, a finding in line with the rest of the literature (Bialystok, 2006; Castel et al., 2005; Chisholm et al., 2010; Green & Flowers, 2003; Greenfield et al., 1996; Karle et al., 2010). In a previous meta-analysis Dye et al. (2009b) demonstrated a consistent degree of speedup across the many attentional and perceptual tasks where such advantages have been noted. When our tasks are outlined on this plot, we see that they line up well with the attentional and perceptual tasks, indicating that AVGPs show the same degree of speeding up in these memory tasks (Fig. 6). Interestingly, in the present study, this greater speed of response was accompanied by some loss in accuracy, whereas the attentional and perceptual tasks analyzed by Dye et al. (2009a) revealed faster reaction times in the face of comparable accuracy. Speed without accuracy loss suggests changes in the sensitivity of processing. In contract, completing tasks quickly but with more errors suggests strategy differences (either implicit or explicit) that could be based in playing fast-paced action games. As Nelson and Strachan (2009) found, people perform tasks more quickly after playing an action video game and for the novice players in that study, this strategy also led to more errors. It could be the case that AVGPs typically use speed as a built-in component of their strategy. Having played enough fast paced games, they may also have developed some ability to minimize the detriment in accuracy from their increased speed. When AVGPs did have lower accuracy than NVGPs, it was in the face of some manner of interference, both in Experiments 1 and 2. In Experiment 1 the interference had a Stroop-like nature, where AVGPs appeared to experience interference from over-learned knowledge (i.e. letter names) when responding about physical attributes, which may be tied to the faster memory access speed that AVGPs displayed. In Experiment 2, the only trials where AVGPs accuracy was substantially lower were when the same probe they just responded to required a different response. Changing response mapping to the same item is not common in action games, and just as they may be learning their speed from playing these games, they may discount this small number of trials in their strategy to complete the task. It is also possible that there is something about interference in response mapping that is more difficult for AVGPs to resolve even though they do not show a detriment for other forms of interference resolution. Clark et al. (2011) also report strategy differences between AVGPs and NVGPs in a change detection task. Specifically, AVGPs used a broader search strategy, which led to increased detection performance in the AVGP group. Gaining a better understanding of the extent to which AVGPs are capable of adaptively modifying selection strategies given current task demands should be a fruitful avenue of research.

We also chose to look at the N-back because changes in working memory performance have been reported for other populations with this type of task. Watson and Strayer (2010), for instance, found a population they refer to as “supertaskers” who do not show a detriment in performance when performing a dual task compared to doing each task separately. These supertaskers also performed better on a dual N-back task than controls, indicating increased capacity and flexibility in working memory. Ophir, Nass, and Wagner (2009), in contrast, found that high media multi-taskers, or individuals who report using multiple media simultaneously, were less accurate on a 3-back task than low media multi-taskers. Specifically, high media multi-taskers had more false alarms to previously seen letters that were not the target than the low media multi-taskers. In both of these cases, the population of interest showed different levels of accuracy establishing that such changes in capacity are certainly possible. However, the present study shows no evidence that AVGPs have a greater working memory capacity as measured by the N-back task.

In the visual short-term memory task AVGPs outperformed NVGPs and their performance was suggestive of better memory performance. However, given the interference trends found in the other experiments, it is also possible that AVGPs have greater short-term memory abilities, but when a task requires response conflict resolution, as is the case with working memory tasks like the N-back task or proactive interference, the group differences are eliminated due to AVGPs poorer interference resolution. While a trend for poorer interference resolution in AVGPs may explain some of our results, an alternate view is that the higher availability of memory traces in AVGPs triggers higher levels of interference, and thus creates more demands on interference resolution for this
population (see Green & Bavelier, 2003 for a similar argument when comparing Flanker Compatibility Effect between AVGPs and NVGPs). The present set of experiments provide directions for future research but remain tentative as the size of the group differences reported when it comes to differences in conflict resolution are relatively small. Finally, while Experiment 4 establishes greater visuospatial short-term memory skills in AVGPs, the contrast between this experiment and the first three may lay with the use of verbal versus visuospatial stimuli. In the first three experiments, including the N-back task, the stimuli were letters, while in Experiment 4 the stimuli were purely visuospatial in nature. AVGPs do show superior object representation (Sungur & Boduroglu, 2012), but this ability would not necessarily hold with letters, which are well learned. In this view, part of the possible AVGP advantage would be removed when using verbal stimuli.

To distinguish between these possibilities, AVGPs and NVGPs need to be more systematically compared on visuospatial versus verbal memory tasks. Finally, and as noted in the introduction, the results above are purely cross-sectional in nature and thus cannot be used to infer a causal relationship between action gaming and those few differences that were observed. Furthermore, as our AVGPs population played essentially no games, we cannot positively identify, assuming a causal relationship is present, whether the critical experiences are unique to the action genre or exist in multiple genres.

Overall, these results indicate comparatively little change in memory functions in AVGP, except for greater visuospatial memory. The weak or null group differences in the other memory tasks presented here contrasts with previously described effects on perception, attention or cognition described in the introduction. It is worth noting that this finding of little to no difference in AVGPs as compared to NVGPs provides evidence against a recent hypothesis that the effects previously seen in AVGPs are due solely to motivational effects that arise due to non-covert recruitment (see: Schubert & Stroebach, 2012 for a rebuttal as well). Indeed, the current results would be difficult to reconcile with such a hypothesis, as the AVGP and NVGP subjects were recruited in the same manner for all four experiments, matching that used in most previous studies and thus should have had the same motivational state as in those previous studies. Instead, the findings are more consistent with the view that action video game experience primarily effects those cognitive processes placed under the most extreme loads during play (i.e. perception, fast-decision making, etc.). However, it may still be the case that playing games from genres that place more significant demands on the memory system (e.g. real-time strategy) would result in clear improvements in the memory system.

Acknowledgments

ONR N 00014-07-1-0937, NEI EY016880, The McDonnell Foundation.

References


