Master thesis on Brain and Cognition

Universitat Pompeu Fabra

Association Between Action Video Game Playing Experience and Visual Search in Real-life Multisensory scenes

Mohammad Hamzeloo

Supervisor: Salvador Soto-Faraco

Co-supervisor: Daria Kvasova

September 2020

Universitat Pompeu Fabra Barcelona

Table of Contents

Dedication

This thesis is dedicated to my wonderful wife Yasamin and our two children, Nazanin and Rayan, whose unyielding love, support and encouragement have enriched my soul and inspired me to pursue and complete this research.

Acknowledgments

First and foremost, I would like to express my deep and sincere gratitude to my thesis supervisor, Dr. Salvador Soto-Faraco, Professor and head of Multisensory Research group at Universität Pompeu Fabra, for giving me the opportunity to do research and providing invaluable guidance throughout this research. His dynamism, vision, sincerity and motivation have deeply inspired me. He has taught me the methodology to carry out the research and to present the research works as clearly as possible. It was a great privilege and honor to work and study under his guidance. I am extremely grateful for what he has offered me.

I would also like to thanks my co-supervisor, Dr. Daria Kvasova, for her valuable guidance in designing the experiment. Without access to her materials for the experiment, this research would not have been possible.

Secondly, I would like to acknowledge and thank all members of MRG group, each of whom has provided patient advice and guidance throughout the research process.

Besides, I would like to thank my committee members: Prof. Núria Sebastián Gallés, Dr. Adrian Fernando Ponce Alvarez, and Dr. Pau Alexander Packard for their brilliant comments and suggestions that would help me to improve.

Finally, and most importantly, huge thank you to my wife Yasamin; without her supports none of this would indeed be possible.

Abstract

Prior studies investigating the effects of playing action video games on attentional control have demonstrated improved performance on a variety of basic tasks. Given this evidence, there has been growing interest in using this kind of entertainment as a tool to improve everyday attentional control. However, as of yet, there is little evidence indicating that the cognitive benefits of playing action video games generalize to complex settings with multisensory integration – a fundamental characteristic of our natural, everyday life environments. The present study addresses the generalization of attentional control enhancement due to AVGP experience to real-life scenarios by comparing the performance of action video-game players (AVGPs) with non-players (NVGPs) on a visual search task using real-life, dynamic audio-visual scenes. To this end, a questionnaire collecting data on gaming habits and sociodemographic data as well as a visual search task was administered online to a gender-balanced international sample of 60 participants of age 18 to 30 years. According to the standard hypothesis, AVGPs outperformed NVGPs in the search task overall by showing faster reaction times without sacrificing accuracy. In addition, in replication of previous findings, semantically congruent cross-modal cues benefit performance, but incongruent cross-modal cues cannot hinder performance. However, according to our results, despite the overall advantage in search, AVGPs cannot exploit multisensory cues more efficiently than NVGPs. Exploratory analyses with gender as a factor indicated that the advantage of AVG experience is constrained to males. Furthermore, the finding from the general analysis of data demonstrated that females provide more impulsive responses in distractor-consistent and neutral conditions than males. These findings provide some evidence that benefits associated with AVG experience extend to more complex naturalistic context, but the use of AVG for attentional control training should be investigated more systematically in terms of potential gender differences.

Introduction

Over the last decade, the rapid development in smart technology has produced an explosion of mobile gaming, which pushed video-game playing to the forefront of scientific attention. Although video game playing is associated with improvements in a variety of cognitive and perceptual skills (Hisam et al., 2018; A. C. Oei & Patterson, 2015; Powers, Brooks, Aldrich, Palladino, & Alfieri, 2013; Strobach, Frensch, & Schubert, 2012), a significant body of work in this domain has established that a causal relation exists between the playing action video games and attentional control (D. Bavelier & Green, 2019). It is proposed that attentional control improvement due to playing action video games can be considered as training-related transfer between different perceptual and cognitive abilities (Ducrocq, Wilson, Vine, & Derakshan, 2016; C. Green, Gorman, & Bavelier, 2016). However, it remains unknown whether the attentional control improvement in AVGPs generalizes beyond simplified laboratory conditions to tasks with the complexity of realistic, multisensory conditions. We set out to address this question.

The literature defines action video games as a type of game that requires the processing of a large amount of visual information, presented rapidly over a wide field of view, and often requires the simultaneous tracking of multiple targets under high attention demands (C. S. Green & Bavelier, 2006). These games include the so-called first- or third-person shooter games or adventure games such as *Medal of Honor*, *Call of Duty* or *Grand Theft Auto (GTA).*

Attentional control is a key mechanism of adaptive behavior and can be defined as our ability to stay focused on task-relevant information while resisting distraction, and being responsive to changes in the environment that require re-orienting the focus to new sources of relevant information (Engle, 2002). Thus, attentional control not only enables selective attention (focusing on spatial locations or objects that are goal-relevant while minimizing task-irrelevant information) but also the capacity to shift between selective and divided attention to allow consistent monitoring of changes in the environment (Corbetta & Shulman, 2002). One of the first scientific studies related to visual attention in Action Video Game Players (AVGPs), compared to Non-Video Game Players (NVGPs), demonstrated lower attentional cost for AVGPs when targets appeared at low probability positions in a stimulus detection paradigm (Greenfield, DeWinstanley, Kilpatrick, & Kaye, 1994). Based on this finding, Greenfield et al. suggested that playing action video games can foster the skills of allocating and dividing attention, thereby improving visual search performance.

Over the past decade, researchers have taken an interest in visual search advantages in AVGPs. For example, AVGPs outperform NVGPs on traditional visual search tasks (Castel, Pratt, & Drummond, 2005; Hubert-Wallander, Green, Sugarman, & Bavelier, 2011), flanker/load tasks (M. G. Dye, C. S. Green, & D. Bavelier, 2009; C. S. Green & Bavelier, 2006; S. Green & D. Bavelier, 2003; Irons, Remington, & McLean, 2011; Xuemin & Bin, 2010), distraction-based tasks (Joseph D Chisholm, Hickey, Theeuwes, & Kingstone, 2010; Rupp, McConnell, & Smither, 2016), and a change detection task (Clark, Fleck, & Mitroff, 2011; Durlach, Kring, & Bowens, 2009). These findings have suggested that action video game experience enhances various aspects of top-down attention, and the effect can be seen both in cross-sectional and in intervention studies (Bediou et al., 2018; Joseph D. Chisholm & Kingstone, 2015a; Schubert et al., 2015). Although it has been thought that these changes are a result of changes in “selective attention”, more recently, enhancement in various additional aspects of top-down attention is considered. For example, changes in attentional control or in the capacity to swiftly switch between attentional modes based on task demands (D. Bavelier & Green, 2019).

The vast majority of studies in this area have compared AVGP and NVGP performance on traditional paradigms with basic visual stimuli, and only a handful of studies have considered other sensory modalities in their experimental designs. In a cross-modal study, Donohue, Woldorff, and Mitroff (2010) compared AVGPs and NVGPs in an audio-visual simultaneity judgment task. The results showed that AVGPs were more accurate to distinguish whether simple visual and auditory pairs occurred in synchrony or slightly offset in time. They also revealed an enhanced ability to determine the temporal order of the different modalities in cross-modal stimuli. In a study measuring auditory decision making, AVGPs were found to be faster compared to NVGPs at indicating the ear to which the sound was presented, especially at low signal to noise ratios (C. S. Green, Pouget, & Bavelier, 2010). In a recent study using a highly demanding auditory discrimination task, AVGPs managed to detect auditory targets and to distinguish them from auditory non-target standards more accurately than NVGPs (Föcker, Mortazavi, Khoe, Hillyard, & Bavelier, 2019).

Despite the growing body of research on the benefits of action video game experience, there have been little studies investigating whether the cognitive benefits observed in the laboratory conditions can also be seen in more complex, naturalistic contexts. This step would be important to understand how the putative superiority of AVGPs in certain tasks transfers to real-life environments that are complex, multisensory, and often semantically meaningful (Soto-Faraco et al., 2019). Recent studies on cross-modal interactions in attention orienting have highlighted that in real-life scenarios, not only temporal and spatial congruence between stimuli across modalities plays a functional role in the control of attention, but also semantic correspondence can facilitate detection and recognition performance (Kvasova & Soto-Faraco, 2019; Roberts & Hall, 2008; Spagna, Mackie, & Fan, 2015). Cross-modal semantic facilitation has been shown in a variety of tasks, including audio-visual matching task (Chen & Spence, 2010; Hein et al., 2007; Iordanescu, Guzman-Martinez, Grabowecky, & Suzuki, 2008; Laurienti et al., 2003), visual awareness (Chen, Yeh, & Spence, 2011; Hsiao, Chen, Spence, & Yeh, 2012), spatial attention (List, Iordanescu, Grabowecky, & Suzuki, 2014; Mastroberardino, Santangelo, & Macaluso, 2015; Pesquita, Brennan, Enns, & Soto-Faraco, 2013), and object search in real-life scenes (Kvasova, Garcia-Vernet, & Soto-Faraco, 2019).

Although evidence to support the idea that benefits of action video game experience can be found in real-life performances is very limited, C. S. Green and Bavelier (2015) suggest that action video game experience induces a form of 'learning to learn', whereby it enables players to suppress sources of noise or distraction efficiently while they extract task-relevant information very fast and more precisely. Based on this hypothesis, one outstanding question is whether the 'learning to learn' ability that putatively emerges as a result of action video game experience can be observed in real-life search. That is, for example when looking for an object in complex, naturalistic scenes. Because naturalistic environments are multisensory, a second interrelated question is whether AVGPs can benefit more efficiently than NVGPs from cross-modal semantic congruence between visual events and sounds when they search for an object in real-life scenes. To answer the two questions posed above, we conducted a study using a visual search task on realistic multisensory scenes adapted from those used in the study by Kvasova et al. (2019).

In the visual search task, targets consist of usual everyday life objects embedded in video clips of natural scenes. For example, participants were asked to look for a dog in a city street scene. A characteristic object sound (consistent with the search target, with a distractor object present in the video clip, or unrelated to the objects in the scene) was presented by stimulus onset, mixed with ambient noise. Reaction times as well as visual search accuracy were measured.

We hypothesized that, if AVGPs benefit from more efficient attentional control to direct their attention to target relevant information while minimizing target irrelevant information in multisensory environments, they would outperform NVGPs in the task of searching objects in real-life scenarios. We advance that this advantage would be observed in both faster reaction times and/or more accurate responses overall in the task. In addition, if AVPGs have learned to use environmental multisensory cues more efficiently, then we expect a cross-modal advantage for AVGPs. That is, in AVGPs the improvement in reaction times in the target-consistent condition with respect to the distractor-consistent condition will be proportionally larger, compared to the NVGPs. We expect this because previous studies revealed that AVGPs can resist distraction more efficiently in high-load perceptual scenarios (M. W. G. Dye, C. S. Green, & D. Bavelier, 2009; S. Green & D. Bavelier, 2003), whilst NVGPs show a reduction in the magnitude of the flanker effect quickly (D. Bavelier & Green, 2019).

Materials and Methods

Experimental Design. The design includes two independent variables: Video-game experience (across subjects) and sound-target relation (within-subjects). Action video game experience was measured using the Bavelier Lab Gaming Questionnaire, version November 2019. Based on the questionnaire adapted from C. S. Green et al. (2017), we had two groups of participants according to their experience: AVGPs and NVGPs. An AVGP is a person who plays 5+ hours per week of First/Third-person shooter and/or action-RPG/adventure genre of games. An NVGP is a person who plays at most 1 hour per week of any genre of video games. This categorization of video game playing habits is based on Li, Polat, Makous, and Bavelier (2009).

The sound-target relation, manipulated as an independent factor within-subjects, relates to the semantic relationship between the visual search target in the task and the object sounds on each trial. This variable had 3 levels: target-consistent sound, distractor-consistent sound, and neutral sounds. In the target-consistent condition, the sound was matched with the target object. In the distractor-consistent condition, the sound was matched with the distractor (a non-target object present in the scene), and in the neutral condition, the object sound was not semantically congruent with any object in the scene. In order to be able to measure false-alarm rates and to balance the response types (50% yes, 50% no) in the task, the task included additional catch trials in which targets are not present (hence, the correct response was 'NO'). Within these catch trials, the object sound had 3 levels, like in experimental trials: target-consistent sound (here, the sound is consistent with the designated search object, which is not present in the video clips), distractor-consistent sounds, and neutral sounds. The dependent variable of our design was visual search object performance, that was measured with reaction times of correct responses in target-present trials, and with search accuracy in all experimental (target-present) conditions as hit rates. In addition, we measured the false-alarm rate in catch trials as a reality check for accuracy.

Participants. We used G Power to estimate the sample size for the study, taking a repeated-measures ANOVA within-between interactions with a medium effect size f=0.25, and α level= 0.05. The total sample size for a statistical power of 0.95 was estimated at N=54 (Tomczak, Tomczak, Kleka, & Lew, 2014). By considering drop-out and inclusion criteria, we enrolled 60 individuals around the world. Most of them were from European countries and some other participants from South American countries, the US, Canada, and one from South Africa. According to their responses to a questionnaire asking about video game play habits of different genres during the previous 12 months (Bavelier Lab Gaming Questionnaire), we had 30 AVGPs and 30 NVGPs. The mean age was 23.78 and SD = 3.12. There was not a significant age difference between the two groups ($t = 0.115$, $df = 28$, $p > 0.05$), and the two groups were gender-balanced (15 males and 15 females in each group).

Otherwise, general inclusion criteria were: (1) having a normal or corrected-to-normal vision and hearing, (2) having a good quality of internet connection and access to the experiment from a laptop or personal computer, (3) the false alarm rate in catch trials (trials in which the search target was not present) was less than 15%, and (4) the average accuracy in the three experimental conditions was more than 70%.

Stimuli. The materials for the visual search task (target objects, sounds, and video clips) in naturalistic environments were selected from those used in Kvasova et al. (2019) (see Fig. 1, for an example). There was a set of 156 different video-clips extracted from movies, TV shows, advertisements, and others recorded by Kvasova et al. from everyday life scenes. Video clips were recorded/played in color, with 30 fps and 1024 * 768 pixels resolution. All video-clips were edited to be 2s duration fragments, without fade-in or -out. Twelve videos were used for training trials, 72 videos were used for the three experimental conditions (see Table 1), and 72 further videos were used for catch trials in which the search target object is not present. The original sounds of the videos had been replaced with background noise created by the superposition of various everyday life sounds. In each video clip in experimental trials, there are at least two visual objects that have a familiar characteristic sound (such as musical instruments, animals, tools, ...). One of these objects was designated the search target and the other the distractor. In both cases, the target/distractor objects were visible but not a part of the main action in the scene. For instance, if a person is talking on the cellphone as a main action in the scene, the cellphone cannot be a target object. Both target and distractor objects were present throughout the video clip. The contribution of the two designated objects of each video as a target or distractor was counterbalanced across the experimental design to compensate for potential biases related to specific objects. To reach this goal, we used each video of experimental trials to be in the three different experimental conditions by combining the video with target-consistent sound, distractor-consistent sound, and neutral sound while one of the two target objects present in the screen before the video to create six equivalent versions of an experimental trial (2 target objects * 3 conditions). In each video clip in catch trials, there was at least one visual object that has a characteristic sound, that was used as a distractor object. We created 3 equivalent versions of catch trials from each video, in the same way, so that each catch video appeared in the three conditions across different versions of the experiment. By 6 versions of experimental trials and 3 versions of catch trials, we had six versions of the task with the same length. Each version was presented to 10 participants with different random order of videos. Characteristic sounds were semantically compatible with the target/distractor object, depending on the particular condition, and were presented centrally, providing no cue for object location. All the sounds were normalized at origin to have equivalent SPL and were presented for 600ms.

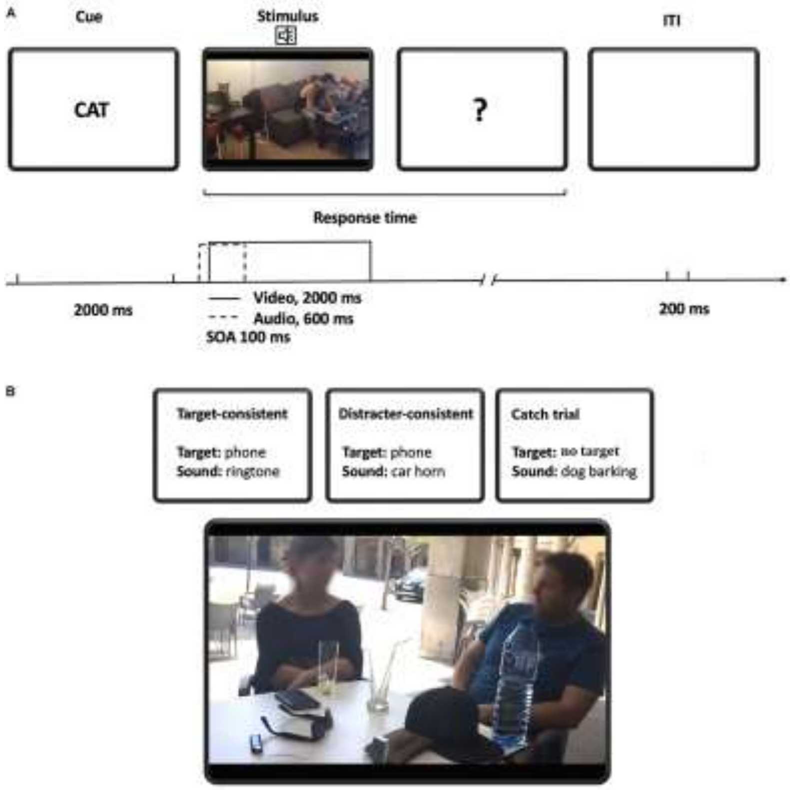

Fig 1. (A) Sequence of events in the experiment. The trial started with the presentation of target word for 2000 ms. The target word was followed by the auditory cue and video. Auditory cue was presented 100 ms before the video was shown (SOA = 100 ms) and lasted for 600 ms while the video lasted for 2000 ms. There was no time limit for the participant response. 200 ms after the participant had responded a new target word was presented. (B) Example of conditions. In this example of stimulus, the possible targets are a mobile phone and a car. If the target is a mobile phone, in the target-consistent condition the sound will match the target, in the distractor-consistent condition the sound will match the distractor (a car), and in the catch trial the sound will not match any object of the scene (e.g., dog barking). The image is a frame of the video clip filmed by the research group.

Procedure. The coronavirus outbreak has triggered us to run the experiment online. We built the experiment using Builder under Psychopy package (v. 2020.1.3) and ran it under the Pavlovia.org platform. Psychopy also is the only package with reaction time precision under 4 ms online (Sauter, Draschkow, & Mack, 2020). We will recruit participants from the Prolific.co. platform. Each participant was asked to read the informed consent and confirm their agreement to participate in the experiment voluntarily. They were able to exit the experiment at any moment by pressing the 'Escape' key in their keyboard. After consent approval, they filled a form of demographic data (see Appendix A) and the Video Game Questionnaire in the first part of the study. If they felt in each group category (AVGP or NVGP), they have received an invitation for the second part of the study. By clicking on the invitation link, the instruction of the main task appeared on the screen and they were asked to do the task in indoor conditions with dim illumination, turn their phone off in order to avoid distractions during the task. They were also asked to increase the volume of their device up to 80% or to some extent in which they can easily hear the sounds. By pressing the space bar, they entered a 12-trial training block before the beginning of the experiment. Feedback related to the participant's response was provided in the training trials to make sure that participants understand the task. No feedback was provided during the experimental blocks.

Table 1: Distribution of trials and conditions

	Condition (type of trial)	**Description**	**Number of trials**
Experimental trials	Target-consistent	The target is present and the sound is congruent with the target object	24
	Distractor-consistent	The target is present and the sound is congruent with a distractor object	24
	Neutral sound	A target is present and the sound is not related to any object in the video	24
Catch trials	Target-consistent	The target is not present and the sound is congruent with the target word	24
	Distractor-consistent	The target is not present and the sound is congruent with a distractor object	24
	Neutral sound	The target is not present and the sound is not related to any object	24
Total			144

Each trial started with a cue word designating the search target object, printed in the middle of the screen for 2000ms. This cue was followed by the video clip, with the corresponding sound. Auditory cues (target-consistent, distractor-consistent, or neutral sound) begin slightly ahead of the video, by 100ms and last for 600ms. The video was shown for 2000ms. The participant will judge whether or not the target object is present in the video clip. They will respond by pressing the Y key as Yes with their left index finger in the case of finding the target object in the video or pressing the N key as No with their right index finger in the case of the target object will not be presented. Participants were instructed to respond as fast as possible, but there was no time limitation for the trial response. A question mark was presented after the video offset, and until a response was made. There will be a 200ms blank screen between trials.

When they finished the task, they were returned to the Prolific website where a successful submission appeared in their account that waiting for the researcher’s approval. If the accuracy of the data they provided was more than 85% of all catch trials, and 70% of all experimental trials, their submission was approved and they were paid 2.5 £ as a compensation for their participation. We collected 85 data but 25 (17 NVGPs and 8 AVGPs) of them were excluded from analyses because of the accuracy criteria. Because of the high rejection rate, we continued to run the experiment until we had 60 valid data to enter in the analyses.

Data analyses

Data pre-processing. To eliminate other processes outside the interest of the study, including fast lucky guesses, delayed responses due to subject’s inattention, or guesses based on the subject’s failure to reach a decision, we considered an outlier filter for RTs +/-2SD around the mean of each condition for each subject: neither RT nor accuracy data were be analyzed for these outlier trials. In some of the analyses, where within-subject factors are tested, data from neutral trials were used to normalize the data from the conditions of interest across participants, in order to reduce inter-individual differences and concentrate on the effects of interest. The normalization was done according to the formula (1) for each subject (i) and condition of interest (j):

$$(1)\ NormRT\_j\_i = RawRT\_j\_i / NeutralRT\_i$$

Where *NormRT_j_i* is the normalized RT for subject *i* in condition *j*, *RawRT_j_i* is the mean of raw RT for subject *i* in condition *j*, and *NeutralRT_i* is the mean RT in the neutral sound condition of subject *i*.

Accuracy data. Signal Detection was used to measure precision of responses from the conditions of interest across participants by calculating *d’* according to formula (2) for each subject (i) and condition of interest (j):

$$(2)\ d'\_j\_i = z(hit\_rate\_j\_i) - z(fa\_rate\_j\_i)$$

where d'_j_i is the d prime for subject i in condition j, hit_rate_j_i is yes responses / total responses for subject i in condition j of experimental trials, fa_rate_j_i is yes responses / total responses for subject i in condition j of catch trials, z(hit_rate_j_i) and z(fa_rate_j_i) is standardized z score for hit_rate_j_i and fa_rate_j_i. If hit_rate or fa_rate is equal to 0, it was replaced with 1/total_responses; and if hit_rate or fa_rate is equal to 1 it was replaced with (total_responses-1)/total_responses.

Additionally, data from neutral trials were used to be able to explore whether the congruence effects were seen in the conditions of interest, if any, are due to cross-modal benefit in target detection, cross-modal interference from distractors, or both.

Results

Results of pre-registered analyses. To evaluate our first hypothesis, whether the benefit of action video game experience can transfer to real-life scenarios, we expect a superiority of AVPGs in RTs and/or precision (*d*'). To test this, we used the inter-subject averages per group (AVPG, NVGP) across all other conditions pooled and performed a one-sided t-test on RTs of correct responses to experimental trials (filtered), and on d' scores. The result shows significant differences between groups on raw RTs ($t_{(3383)}$ = 7.10, $p < 0.001$, Cohen's d=0.39) (Figure 2A), whereas no such difference occurred in the d' total ($t_{(58)}$ = 0.68, $p = 0.245$) (Figure 2B). These results confirmed the first part of our first hypothesis and indicate that AVGPs responded faster, albeit equally accurate, than NVGPs.

Table 2: Descriptive statistics (means and standard deviations)

Variable	AVGP (n=30)			NVGP (n=30)			Total		
	Male	Female	Total	Male	Female	Total	Male	Female	Total
Age	25(3)	24(3)	24(3)	25(4)	24(3)	24(4)	25(3)	24(3)	24(3)
RTs for TC	1.40(.12)	1.42(.28)	1.41(.21)	1.73(.40)	1.58(.24)	1.65(.33)	1.56(.34)	1.50(.27)	1.53(.30)
RTs for DC	1.46(.13)	1.45(.46)	1.45(.33)	1.70(.40)	1.65(.32)	1.68(.36)	1.58(.32)	1.55(.40)	1.56(.34)
RTs for NC	1.40(.13)	1.53(.44)	1.47(.33)	1.74(.39)	1.63(.22)	1.69(.32)	1.57(.34)	1.58(.34)	1.58(.34)
RTs for ET	1.42(.44)	1.51(.61	1.46(.53)	1.73(.59)	1.62(.53)	1.68(.56)	1.58(.54)	1.56(.58)	1.57(.56)
Norm RTs for TC	1.00(.10)	0.95(/11)	0.97(.11)	1.00(.13)	0.97(.08)	0.99(.11)	1.00(.11)	0.96(.09)	0.98(.11)
Norm RTs for DC	1.05(.10)	0.95(.10)	1.00(.11)	0.98(.09)	1.01(.10)	0.99(.09)	1.01(.10)	0.98(.10)	0.99(.10)
D prime for TC	2.52(.52)	2.68(.33)	2.60(.43)	2.72(.51)	2.51(.47)	2.61(.49)	2.62(.52)	2.60(.41)	2.61(.46)
D prime for DC	2.79(.40)	2.36(.40)	2.57(.45)	2.59(.56)	2.25(.60)	2.42(.60)	2.69(.49)	2.31(.51)	2.50(.53)
D prime for NC	2.71(.53)	2.33(.59)	2.52(.58)	2.72(.48)	2.16(.40)	2.44(.52)	2.72(.49)	2.24(.50)	2.48(.55)
D prime Total	2.8(.55)	2.52(.44)	2.66(.51)	2.77(.59)	2.38(.39)	2.58(.53)	2.78(.56)	2.45(.42)	2.62(.52)

TC= Target_consistent condition; DC= Distractor_consistent condition; NC= Neutral condition; ET= Experimental Trials;

To evaluate our second hypothesis, whether AVGPs can benefit more from cross-modal cues and/or are more resistant to cross-modal distractors, we expect an interaction between Group and Sound Condition. We entered normalized RTs into a repeated-measures analysis of variance (ANOVA) with the between-participants factor of Group (AVGPs vs. NVGPs) and the within-participants factor Sound Condition (target-congruent and distractor congruent, normalized). The interaction of group*condition was not significant ($F_{(1, 58)} = 0.218$, $p = 0.64$) (Figure 2C). We ran another repeated measures ANOVA on *d'* scores with the same factors and the result of the interaction between groups and sound conditions was not significant ($F_{(2, 57)} = 0.526$, $p = 0.59$) (Figure 2D). The analyses did not confirm our second hypothesis and suggest that AVGPs do not seem to benefit more or be more interfered from cross-modal cues than NVGPs in speed or accuracy.

Exploratory analyses. In addition to the hypothesis-driven analyses described above, we performed two further exploratory analyses which are not addressed directly at testing our initial hypothesis:

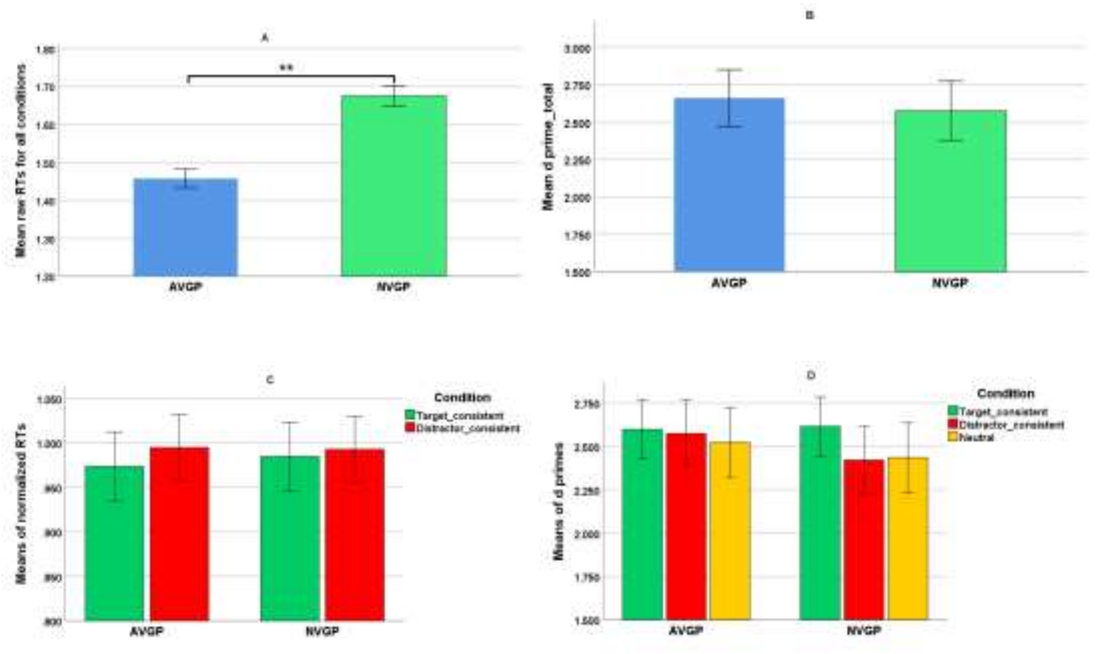

Figure 2: (A) Visual search reaction times toward a target and error rates were plotted for Action Video Game Players (AVGP) and Non-Video Game Players (NVGP). Error bars indicate the standard error of means. Asterisks indicate significant difference in the conditions between two groups (***p*-value<0.01). (B) Visual search accuracy (d prime) toward a target and error rates were plotted for the two groups. Error bars indicate the standard error of means. The differences between the two groups in the conditions were not significant. (C) Normalized reaction times in the target-consistent condition and distractor-consistent condition were plotted. The differences between the two groups in the conditions were not significant. (D) D primes were plotted in the three conditions and the differences between the two groups in the conditions were not significant.

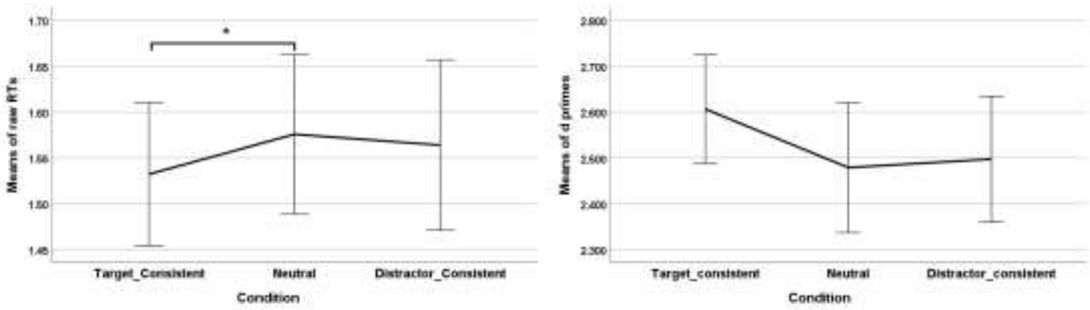

Figure 3: (A) Raw reaction times in the three experimental conditions were plotted for all data. Error bars indicate the standard error of means. Significant differences are indicated by Asterisks (*p-value<0.05). (B) Visual search accuracy (d prime) for the three experimental conditions and error rates were plotted for all data.

(1) Comparison between neutral and target-consistent condition, and between neutral and distractor consistent condition. This analysis shall be informative as to whether congruent cross-modal cues benefit performance, incongruent cross-modal cues hinder performance, or both. We ran two one-sided paired samples t-tests to compare raw RTs of target-consistent condition and distractor consistent condition to raw RTs of neutral condition. The results show a significant difference on raw RTs of target-consistent condition and neutral condition ($t_{(59)} = 1.849$, $p = 0.033$, Cohen's d=0.16), but the difference between raw RTs of distractor-consistent condition and neutral condition was not significant ($t_{(59)} = 0.602$, $p = 0.275$) (Figure 3A). We ran another two one-sided paired samples t-tests to compare d primes of target-consistent condition and distractor consistent condition to d prime of neutral condition. The results show a marginal tendency for higher accuracy in target-consistent condition than neutral condition but not significant ($t_{(59)} = 1.405$, $p = 0.083$, Cohen's d=0.26) (Figure 3B). As for RTs, there was no significant difference between d prime of distractor-consistent condition and neutral condition ($t_{(59)} = 0.266$, $p = 0.395$). The results indicate that congruent cross-modal cues benefit performance, but incongruent cross-modal cues do not hinder performance, with respect to neutral conditions. These results replicate previous findings from Kvasova & Soto-Faraco (2019).

(2) Gender-dependent effects. Given that we have collected data from a gender-balanced sample for both AVGP and NVGP groups, we entered gender as a factor in analyses parallel to the ones described in the pre-registered analyses above. There was a nearly significant interaction for group*condition*gender ($F_{(1, 56)} = 3.07$, $p = 0.085$) when we ran a repeated measures ANOVA on normalized RTs. Owing to this marginal effect, we broke down the analysis in the two ANOVAS, one for each gender on the

group*condition interaction. Neither the analysis for male ($F_{(1, 28)} = 2.58$, $p = 0.12$) (Figure 4A) nor for female ($F_{(1, 28)} = 0.79$, $p = 0.38$) (Figure 4B) were significant. When we took a carefully look at the graph A and B in the figure 4, it seems that there are some differences between female and male AVGPs. We did some follow up analyses by running a repeated measures ANOVA on normalized RTs just for AVGP and gender as a factor. We found a significant interaction for gender in the tests of between-subjects effects on normalized RTs ($F_{(1, 57)} = 5.84$, $p = 0.022$) (Figure 4C). Post hoc t-tests show a

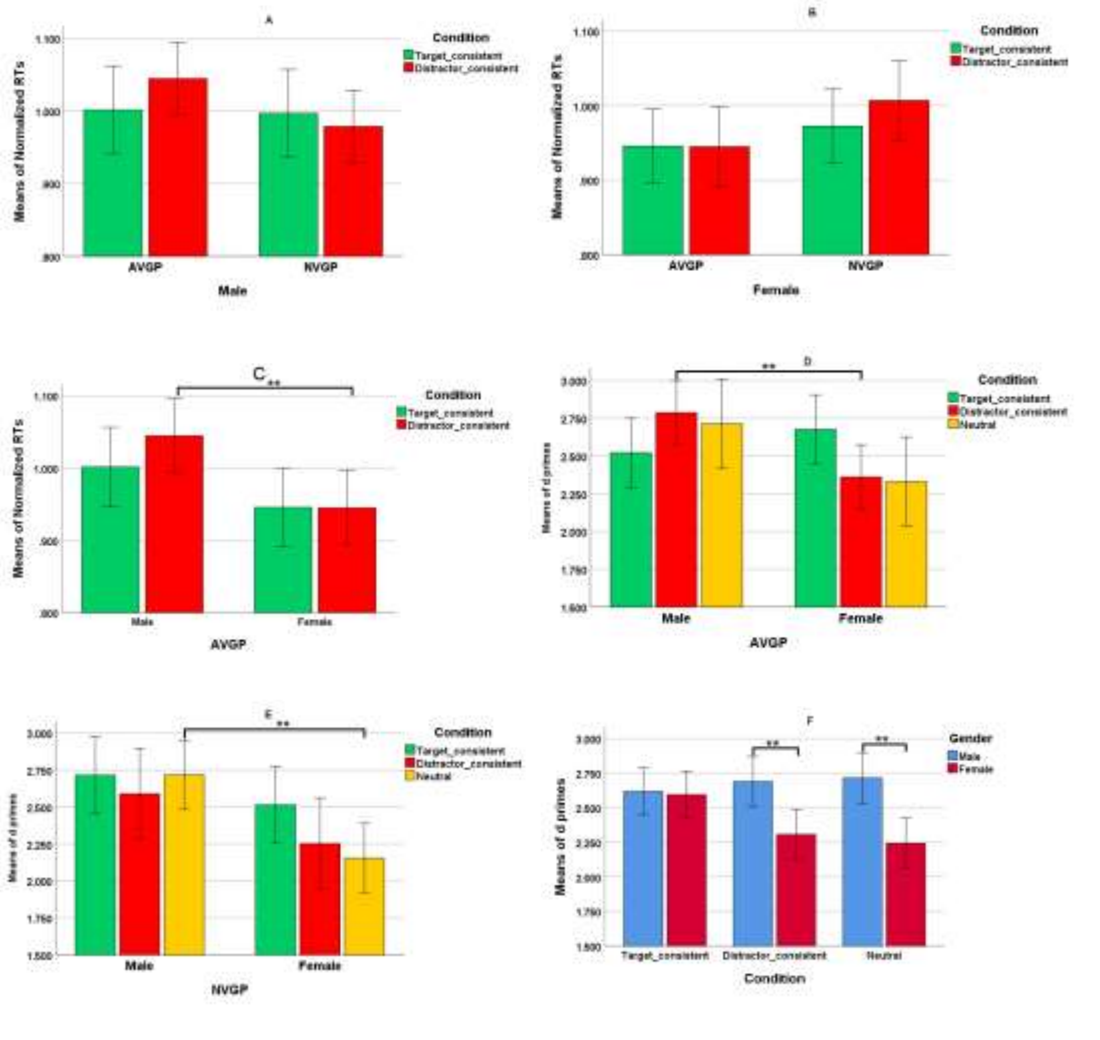

Figure 4: **(A)** Normalized reaction times in the target_consistent and distractor-consistent conditions were plotted for males in the two groups. Error bars indicate the standard error of means. **(B)** Normalized reaction times in the target_consistent and distractor-consistent conditions were plotted for females in the two groups. Error bars indicate the standard error of means. **(C)** Normalized reaction times in the target_consistent and distractor-consistent conditions were plotted for males and female AVGPs. **(D)** Visual search accuracy (d prime) for target_consistent, distractor-consistent, and neutral conditions and error rates were plotted for males and female AVGPs. **(E)** Visual search accuracy (d prime) for target_consistent, distractor-consistent, and neutral conditions and error rates were plotted for males and female NVGPs. **(F)** D primes were plotted for gender differences in the three conditions. Significant differences are indicated by Asterisks (*p-value<0.05, **p-value<0.01).

significant difference between male AVGPs and female AVGPs in the distractor-consistent condition ($t_{(28)}$ = 2.80, p = 0.009, Cohen's d=1.0) while there was not such this difference between male NVGPs and female NVGPs ($t_{(28)}$ = 1.47, p = 0.152). The result indicates that the interference effect for male AVGPs were greater than female AVGPs in way that they slowed down their reaction times in distractor-consistent condition more than female AVGPs.

By entering gender as factor in the repeated measures ANOVA gender*group*condition on *d'* scores, we found a significant interaction for condition and gender ($F_{(2, 55)}$ = 3.51, p = 0.037) and a significant main effect of gender ($F_{(1, 59)}$ = 11.34, p = 0.001). Follow up analyses showed significant differences between male AVGPs and female AVGPs in d prime for distractor-consistent condition ($t_{(28)}$ = 2.91, p = 0.007, Cohen's d=1.08) (Figure 4B) and between male NVGPs and female NVGPs in d prime for neutral condition ($t_{(28)}$ = 3.488, p = 0.002, Cohen's d=1.27) (Figure 4B). When we analyzed all accuracy data for gender difference, the differences between male and female in d prime of distractor-consistent condition ($t_{(58)}$ = 2.966, p = 0.004, Cohen's d=0.77) and neutral condition ($t_{(58)}$ = 3.667, p = 0.001, Cohen's d=0.97) were significant (Figure 4C). These findings suggest that males in general, provided more accurate responses in distractor-consistent and neutral condition than females.

Discussion

The first aim of the present study was to assess whether the visual search advantage demonstrated by AVGPs in comparison to NVGPs with simple stimuli in classic psychophysical tasks would generalize to a real-life multisensory complex situation. Using a visual search task in real-life scenes adapted from Kvasova & Soto-Faraco, 2019, with audio-visual congruent/incongruent cross-modal cues, we demonstrated that AVGPs can extend their advantage to real-life scenarios by a faster response time while they keep the accuracy rate equal to NVGPs. The advantage of AVGPs is consistent with previous studies ((Joseph D. Chisholm & Kingstone, 2012, 2015a; Donohue et al., 2010; Stewart, Martinez, Perdew, Green, & Moore, 2019) that have revealed that advantage of extensive action video game experience can transfer to naturalistic multisensory situations. While the mechanism of this generalization still remains debated, one explanation is that AVGPs may have a heightened speed of processing (e.g., M. W. Dye, C. S. Green, & D. Bavelier, 2009). An action video game requires rapid processing of audio-visual information and promotes action per unit of time, forcing players to make decisions and execute responses as fast as possible. Each speed processing which ends up with a speed accurate response provides an incentive for players while delays in processing often have severe consequences. Therefore, extensive experience with action video games may lead to more efficient visual processing and improve RTs across a range of unrelated tasks without sacrificing accuracy. Most past studies have compared AVGPs to NVGPs by measuring RTs have consistently shown that AVGPs are faster overall than NVGP (Castel et al., 2005; Greenfield et al., 1994; Nuyens, Kuss, Lopez-Fernandez, & Griffiths, 2019; Powers et al., 2013; Schubert et al., 2015; Stewart et al., 2019; Torner, Carbonell, & Castejón, 2019).

Another explanation for the advantage of AVGPs would be greater attentional control capacity, compared to NVGPs (e.g., C. S. Green & Bavelier, 2006), allowing them to direct their attention toward the spatial position of the visual stimulus more quickly and accurately, which in turn allows them to include other modality inputs when they are consistent with goal-directed visual search and exclude other irrelevant inputs from processing. It has been suggested that attentional control plasticity as a result of action video gameplay is a fundamental building block of cognitive enhancement in AVGPs (D. Bavelier & Green, 2019). Enhancement of various aspects of top-down attention as an effect of action video gameplay is seen in cross-sectional studies (Bediou et al., 2018; Joseph D. Chisholm & Kingstone, 2012, 2015b). Similarly, intervention studies mirrored the analysis of cross-sectional studies by showing that action video game training can radically alter visual attentional processing. The results of these studies suggest a causal role of action gaming on spatial cognition and top-down attention, since they rule out inherently enhanced attentional control in AVGPs (Bediou et al., 2018; Chiappe, Conger, Liao, Caldwell, & Vu, 2013; C. S. Green & Bavelier, 2012; S. Green & D. Bavelier, 2003). Furthermore, training studies make a distinction between action video games and other video game types such as mimetic games, sport games, simulation games or puzzle games because they could not find the same impact for other types of games (Cohen, Green, & Bavelier, 2007; Powers & Brooks, 2014; Powers et al., 2013).

Given the online cross-sectional design of the current study, it provides evidence that the advantages associated with action video game experience extend beyond typical laboratory tasks to more complex naturalistic contexts. Besides, it is important to note that a training component was not included in the study and a degree of caution must be advised when considering the causal relationship between action video game experience and improved visual search in real-life multisensory context. Therefore, future work assessing performance within multisensory paradigms pre- and post-action video game training will be important to further evaluate the relation between action video game experience and cognitive abilities in real life.

The second aim of the present study was to assess whether there is a cross-modal advantage for AVGPs than NVGPs. By entering the normalized data in the analysis of performance in the visual search task in real-life scenarios, the results suggest that AVGPs cannot benefit from cross-modal cues more than NVGPs, nor they are more resistant to distractors. These results consistent with the previous studies (Gao et al., 2018; Rupp et al., 2016) showing that there is no difference between AVGPs and NVGPs in the ability to integrate audio-visual stimuli or in the driving performance while distracted in a driving simulator, but AVGPs display greater ability in a discrimination (temporal order judgment) task for visual and auditory stimuli (Donohue et al., 2010) when the auditory stimulus comes after the visual. In the task used in our study, the sound presented slightly earlier (100 ms) than videos to capture attention and allow us to study object-based congruency/incongruency effect. The results indicate that although AVGPs are faster in overall visual search, they cannot benefit more from semantic audio-visual integration to direct their attention toward the target object than NVGPs. One explanation for this is that AVGPs and NVGPs may employ similar cognitive strategies when they need to process the information at the higher-level, semantic aspects. This result may suggest that the AVGP advantage may be based on early, low level processing stages. It has been shown that AVGPs benefit from low-level spatial and temporal factors in their visual (C. S. Green & Bavelier, 2006; Greenfield et al., 1994; Schubert et al., 2015; Wong & Chang, 2018; Xuemin & Bin, 2010), auditory (C. S. Green et al., 2010; Stewart et al., 2019), or audio-visual (Zhang, Tang, Moore, & Amitay, 2017) search tasks. A recent meta-analysis confirmed a smaller impact on higher cognitive performance like inhibition, and verbal cognition while there was not a significant impact on problem-solving in cross-sectional studies (Bediou et al., 2018). Future studies need to explore whether extensive action video game experience has a positive impact on high cognitive functions or it is limited to low-level perception, spatial cognition, and top-down attention.

From our exploratory analyses, we could reveal that semantic congruency between sounds and target objects speed up search latencies in comparison with neutral sounds while semantic incongruency in distractor-consistent condition did not produce any disadvantage with respect to the baseline (neutral condition). Hence, consistent cross-modal cues benefit performance, whilst incongruent cross-modal cues do not hinder performance to a measurable extent. This finding presents a replication of a previous study (Kvasova et al., 2019), indicating that semantic congruency effect can generalize beyond typical laboratory protocols and guide attention in a complex, multisensory environment.

The design of this study allows us to calculate false alarm rate and d prime for distractor-consistent condition which was a limitation in the Kvasova's study. In keeping with the conclusions of that study, there was no difference in d prime between distractor-consistent and neutral conditions. This indicates that the distractor effect neither slows down response time nor create a more impulsive response than neutral conditions. An open question here is that why semantic congruency can benefit performance but semantic incongruency does not hinder performance. There are two possibilities to answer this question. One plausible candidate seems to be that a target-inconsistent sound does not strongly distract participants from responding to the visual target when visual stimuli are highly informative. Yuval-Greenberg and Deouell (2009) studied the influence of visual on auditory processing and vice versa. They found an asymmetry in the extent to which one modality influenced the other. When participants were responding to an auditory stimulus, the congruency effect was greater for visual stimuli, by contrast, when they were responding to a visual stimulus, the congruency effect was smaller for auditory stimuli. It seems that in adults, auditory stimuli cannot confer a particular advantage for detecting visual stimuli. In this study, when the cue word is presented to participants, it activates a semantic network related to the target and creates an attentional template to direct participant's attention to other semantically congruent inputs from other modalities while inhibiting unrelated cross-modal information. If inputs come from other modalities that are incongruent with the target search and is easy to inhibit, it cannot enter in the processing to create a distractor effect. In contrast, if the inputs from other modalities are highly informative and are not easy to inhibit, they enter the processing and direct attention to creating a distractor effect. In our study, some auditory stimuli may not be informative enough to direct participant's attention and affect visual search because some of them are physically similar (e.g., the sound of the coin and keys) or semantically similar (e.g., sound of musical instruments). Thus, we cannot be sure that they were powerful enough to play a proper role of a distractor or neutral sound. To clarify this point, further studies should be run on the effect of high informative (hard to inhibit) versus low informative (easy to inhibit) congruent/incongruent cross-modal cues. In addition, it would be interesting to compare cross-modal congruency advantages between adults and children, as it can provide information on tracking developmental changes in semantic associations.

Our exploratory analyses on gender provide some further interesting results. Our result demonstrated female AVGPs were faster in distractor-consistent condition than male AVGPs, but male AVGPs provide more accurate responses in distractor-consistent and neutral conditions than female AVGPs. It means that the increased speed of processing noted in female AVGPs in distractor-consistent condition can be viewed as impulsive behavior, in which female AVGPs respond faster but make more anticipatory errors. Furthermore, the finding from the general analysis of data demonstrated that females provide more impulsive responses in distractor-consistent and neutral conditions than males. With the due cautiousness, given the exploratory nature of this analysis, this finding could be in line with previous studies indicating that there is a gender difference in selective attention and spatial cognition (Evans & Hampson, 2015; Halpern, 2013; Merritt et al., 2007; Posner & Marin, 2016; Stoet, 2017). If so, the gender difference in cognitive functions may partly justify known gender differences in game preferences. According to game industry surveys, the male and female game player ratio has become closer to equal and since the 2010s, women have been found to make up around half of all gamers (Harwell, 2014). However, only a quarter of female video game players are in favor of videogames to have violent content, instead, half of them more likely to play puzzle and strategy-based video games (Lopez-Fernandez, Williams, & Kuss, 2019). To answer this question why do women play action video games less than men, a recent meta-analysis review suggests a vast variety of factors from physical to motivational and psychosocial and well-being that need to be considered (Lopez-Fernandez, Williams, Griffiths, & Kuss, 2019).

Low prevalence of female AVGPs has caused a great gender imbalance in cross-sectional studies comparing AVGPs with NPVGs, which has been almost exclusively based on a male population. Moreover, a large proportion of the action video game training studies have done with small samples of gender-imbalanced participants (Cohen et al., 2007; C. S. Green & D. Bavelier, 2003; C. S. Green et al., 2010; S. Green & D. Bavelier, 2003; Adam C. Oei & Patterson, 2013). These studies typically demonstrate the benefits of AVG training in experimental groups but they did not compare the effect of this training on males and females. However, some studies showed that video game playing may decrease the gender differences in selective attention and spatial cognition (Feng, Spence, & Pratt, 2007; Okagaki & Frensch, 1994; Subrahmanyam & Greenfield, 1994). Here, we could use a gender-balanced sample of participants, and what is more, with equal number of males and females in both AVPG and NVPG groups.

It is suggested, from the finding reported here, that the advantage for AVGPs is constrained to males, and it is likely because female AVGPs provide more speed-accuracy trade-off (speeding up resulting in more mistakes) responses on realistic multisensory scenes. We can conclude that AVG experience in females may not result in enhanced attentional control to enabling more efficient suppression of source of noise or distraction as well as males and thus, their performance affected by cross-modal task-irrelevant information. Based on these findings, there may be differences in the "learning to learn" abilities between males and females as a result of action AVG playing. As Daphne Bavelier, Green, Pouget, and Schrater (2012) said, these abilities could be driven by changes in cognitive resources: attentional resources, selective attention, divided attention, sustained attention, resource allocation and rule learning. As we discussed above, there is a gender difference in these resources. In our task attentional resources typically called when a visual target needs to be detected in a scene full of distractors and noises.

In general, the present gender-related differences are the result of exploratory analyses with probably insufficient statistical power. Therefore, the conclusions must be tentative and more studies, with larger samples and with the adequate statistical should be performed to follow up this result. For now, one conclusion that can be drawn from this study is that potential gender differences in the effects of AVGP, neglected in most studies so far, should be addressed carefully in the future. It is suggested that more cross-sectional and intervention studies need to evaluate AVG playing effect on other resources (like rule learning) in male and female AVGPs separately.

Along with the limitations that we discussed earlier, another specific problem with the current online study is that because the experiment was presented in a browser, auditory stimuli and video clips quality, size, and luminance have been varied between participants. It is unclear if this variation has been similar between the groups (AVGPs vs. NVGPs and men vs, women). If not, such a group difference might influence the data. Although there is no reason to assume such group differences are likely to have occurred because the finding of the cross-modal semantic effect was very similar to the laboratory-based study of Kvasova et al. (2019). Another limitation worth mentioning here is that as it is an online study which was done in the computers, the differences between AVGPs and NVGPs may be as a result of motivational or strategic benefits that can arise from extensive video game playing. Since many games they play are on the computers, AVGPs may have been more motivated to perform well tasks in the computers. However, if the higher motivational states were underlying the AVGPs improved performance, we should expect to see a uniform improvement for all AVGPs over NVGPs. That is, regardless of gender, AVGPs should differ from NVGPs cross all conditions. However, not only there was not a significant difference for all conditions between the two groups, but also our effect revealed a clear gender asymmetry in semantic incongruency effect cross AVGPs. Thus, it seems unlikely that the currently observed effects were due to differences in motivation alone. However, more works in other natural contexts (e.g., virtual reality or other situations that do not involve sitting in front of a display) are needed to evaluate the extent to which the observed advantages can be generalized.

Conclusion

The results of the present investigation demonstrate that AVGP advantage in attention tasks can be extended to naturalistic multisensory situations. We also failed to demonstrate that AVGP experience creates any specific advantage to benefit more from cross-modal cues, or more resistance to distractors, compared than NVGPs. Our findings from exploratory analyses shed some light on semantic aspects of multisensory integration by generalizing (and confirm) previous laboratory findings on semantic congruency/incongruency effect on cross-modal interactions. Other findings from exploratory analyses with gender indicate that the advantage of AVG experience may be constrained to males. This tentative finding, if confirmed, may result from females not benefiting as much from enhanced attentional control to enable efficient suppression of noise or distraction compared to male AVPGs. This finding may place some constrain to the hypothesis of "learning to learn" abilities thought to emerge as a result of AVG playing.

References

Bavelier, D., & Green, C. S. (2019). Enhancing Attentional Control: Lessons from Action Video Games. *Neuron, 104*(1), 147-163. doi:10.1016/j.neuron.2019.09.031

Bavelier, D., Green, C. S., Pouget, A., & Schrater, P. (2012). Brain Plasticity Through the Life Span: Learning to Learn and Action Video Games. *35*(1), 391-416. doi:10.1146/annurev-neuro-060909-152832

Bediou, B., Adams, D. M., Mayer, R. E., Tipton, E., Green, C. S., & Bavelier, D. (2018). Meta-analysis of action video game impact on perceptual, attentional, and cognitive skills. *Psychol Bull, 144*(1), 77-110. doi:10.1037/bul0000130

Castel, A. D., Pratt, J., & Drummond, E. (2005). The effects of action video game experience on the time course of inhibition of return and the efficiency of visual search. *Acta Psychologica, 119*(2), 217-230.

Chen, Y.-C., & Spence, C. (2010). When hearing the bark helps to identify the dog: Semantically-congruent sounds modulate the identification of masked pictures. *Cognition, 114*(3), 389-404.

Chen, Y.-C., Yeh, S.-L., & Spence, C. (2011). Crossmodal constraints on human perceptual awareness: auditory semantic modulation of binocular rivalry. *Frontiers in psychology, 2*, 212.

Chiappe, D., Conger, M., Liao, J., Caldwell, J. L., & Vu, K. P. (2013). Improving multi-tasking ability through action videogames. *Appl Ergon, 44*(2), 278-284. doi:10.1016/j.apergo.2012.08.002

Chisholm, J. D., Hickey, C., Theeuwes, J., & Kingstone, A. (2010). Reduced attentional capture in action video game players. *Attention, Perception, & Psychophysics, 72*(3), 667-671.

Chisholm, J. D., & Kingstone, A. (2012). Improved top-down control reduces oculomotor capture: The case of action video game players. *Attention, Perception, & Psychophysics, 74*(2), 257-262. doi:10.3758/s13414-011-0253-0

Chisholm, J. D., & Kingstone, A. (2015a). Action video game players' visual search advantage extends to biologically relevant stimuli. *Acta Psychologica, 159*, 93-99. doi:https://doi.org/10.1016/j.actpsy.2015.06.001

Chisholm, J. D., & Kingstone, A. (2015b). Action video games and improved attentional control: Disentangling selection- and response-based processes. *Psychonomic Bulletin & Review, 22*(5), 1430-1436. doi:10.3758/s13423-015-0818-3

Clark, K., Fleck, M. S., & Mitroff, S. R. (2011). Enhanced change detection performance reveals improved strategy use in avid action video game players. *Acta Psychologica, 136*(1), 67-72.

Cohen, J. E., Green, C. S., & Bavelier, D. (2007). Training visual attention with video games: not all games are created equal. In H. F. O'Neil, H. O'Neil, & R. S. Perez (Eds.), *Computer games and team and individual learning* (pp. 205–227). Amsterdam, the Netherlands: Elsevier Group Inc.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci, 3*(3), 201-215. doi:10.1038/nrn755

Donohue, S. E., Woldorff, M. G., & Mitroff, S. R. (2010). Video game players show more precise multisensory temporal processing abilities. *Atten Percept Psychophys, 72*(4), 1120-1129. doi:10.3758/app.72.4.1120

Ducrocq, E., Wilson, M., Vine, S., & Derakshan, N. (2016). Training Attentional Control Improves Cognitive and Motor Task Performance. *Journal of Sport and Exercise Psychology, 38*, 1-36. doi:10.1123/jsep.2016-0052

Durlach, P. J., Kring, J. P., & Bowens, L. D. (2009). Effects of action video game experience on change detection. *Military Psychology, 21*(1), 24-39. doi:10.1080/08995600802565694

Dye, M. G., Green, C. S., & Bavelier, D. (2009). The development of attention skills in action video game players. *Neuropsychologia, 47*(8-9), 1780-1789.

Dye, M. W., Green, C. S., & Bavelier, D. (2009). Increasing Speed of Processing With Action Video Games. *Current directions in psychological science, 18*(6), 321-326. doi:10.1111/j.1467-8721.2009.01660.x

Dye, M. W. G., Green, C. S., & Bavelier, D. (2009). The development of attention skills in action video game players. *Neuropsychologia, 47*(8), 1780-1789. doi:https://doi.org/10.1016/j.neuropsychologia.2009.02.002

Engle, R. W. (2002). Working memory capacity as executive attention. *Current directions in psychological science, 11*(1), 19-23.

Evans, K. L., & Hampson, E. (2015). Sex-dependent effects on tasks assessing reinforcement learning and interference inhibition. *Frontiers in psychology, 6*, 1044. doi:10.3389/fpsyg.2015.01044

Feng, J., Spence, I., & Pratt, J. (2007). Playing an action video game reduces gender differences in spatial cognition. *Psychological science, 18*(10), 850-855. doi:10.1111/j.1467-9280.2007.01990.x

Föcker, J., Mortazavi, M., Khoe, W., Hillyard, S. A., & Bavelier, D. (2019). Neural correlates of enhanced visual attentional control in action video game players: An event-related potential study. *J. Cognitive Neuroscience, 31*(3), 377–389. doi:10.1162/jocn_a_01230

Gao, Y.-L., Zhang, S.-H., Zhang, Y., Wang, M.-X., Li, Y.-X., Wang, Z.-N., . . . Ye, C. (2018). *Action Video Games Influence on Audiovisual Integration in Visual Selective Attention Condition*. Paper presented at the International Conference on Medicine Sciences and Bioengineering, Suzhou, Jiangsu, China.

Green, C., Gorman, T., & Bavelier, D. (2016). Action Video-Game Training and Its Effects on Perception and Attentional Control. In (pp. 107-116).

Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature, 423*(6939), 534-537. doi:10.1038/nature01647

Green, C. S., & Bavelier, D. (2006). Effect of action video games on the spatial distribution of visuospatial attention. *Journal of Experimental Psychology: Human Perception and Performance, 32*(6), 1465-1478. doi:10.1037/0096-1523.32.6.1465

Green, C. S., & Bavelier, D. (2012). Learning, attentional control, and action video games. *Current Biology: Cell Press, 22*(6), R197-R206. doi:10.1016/j.cub.2012.02.012

Green, C. S., & Bavelier, D. (2015). Action video game training for cognitive enhancement. *Current Opinion in Behavioral Sciences, 4*, 103-108. doi:https://doi.org/10.1016/j.cobeha.2015.04.012

Green, C. S., Kattner, F., Eichenbaum, A., Bediou, B., Adams, D. M., Mayer, R. E., & Bavelier, D. (2017). Playing Some Video Games but Not Others Is Related to Cognitive Abilities: A Critique of Unsworth et al. (2015). *Psychol Sci, 28*(5), 679-682. doi:10.1177/0956797616644837

Green, C. S., Pouget, A., & Bavelier, D. (2010). Improved probabilistic inference as a general learning mechanism with action video games. *Curr Biol, 20*(17), 1573-1579. doi:10.1016/j.cub.2010.07.040

Green, S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature, 423*(6939), 534-537.

Greenfield, P. M., DeWinstanley, P., Kilpatrick, H., & Kaye, D. (1994). Action video games and informal education: Effects on strategies for dividing visual attention. *Journal of Applied Developmental Psychology, 15*(1), 105-123. doi:https://doi.org/10.1016/0193-3973(94)90008-6

Halpern, D. F. (2013). *Sex differences in cognitive abilities*: Psychology press.

Harwell, D. (2014, 17 Oct). More women play video games than boys, and other surprising facts lost in the mess of Gamergate. *The Washington Post*.

Hein, G., Doehrmann, O., Müller, N. G., Kaiser, J., Muckli, L., & Naumer, M. J. (2007). Object familiarity and semantic congruency modulate responses in cortical audiovisual integration areas. *Journal of Neuroscience, 27*(30), 7881-7887.

Hisam, A., Mashhadi, S. F., Faheem, M., Sohail, M., Ikhlaq, B., & Iqbal, I. (2018). Does playing video games effect cognitive abilities in Pakistani children? *Pak J Med Sci, 34*(6), 1507-1511. doi:10.12669/pjms.346.15532

Hsiao, J.-Y., Chen, Y.-C., Spence, C., & Yeh, S.-L. (2012). Assessing the effects of audiovisual semantic congruency on the perception of a bistable figure. *Consciousness and Cognition, 21*(2), 775-787.

Hubert-Wallander, B., Green, C. S., Sugarman, M., & Bavelier, D. (2011). Changes in search rate but not in the dynamics of exogenous attention in action videogame players. *Attention, Perception, & Psychophysics, 73*(8), 2399-2412.

Iordanescu, L., Guzman-Martinez, E., Grabowecky, M., & Suzuki, S. (2008). Characteristic sounds facilitate visual search. *Psychonomic Bulletin & Review, 15*(3), 548-554.

Irons, J. L., Remington, R. W., & McLean, J. P. (2011). Not so fast: Rethinking the effects of action video games on attentional capacity. *Australian Journal of Psychology, 63*(4), 224-231.

Kvasova, D., Garcia-Vernet, L., & Soto-Faraco, S. (2019). Characteristic Sounds Facilitate Object Search in Real-Life Scenes. *Frontiers in Psychology, 10*(2511). doi:10.3389/fpsyg.2019.02511

Kvasova, D., & Soto-Faraco, S. (2019). Not so automatic: Task relevance and perceptual load modulate cross-modal semantic congruence effects on spatial orienting. *bioRxiv*, 830679. doi:10.1101/830679

Laurienti, P. J., Wallace, M. T., Maldjian, J. A., Susi, C. M., Stein, B. E., & Burdette, J. H. (2003). Cross-modal sensory processing in the anterior cingulate and medial prefrontal cortices. *Human brain mapping, 19*(4), 213-223.

Li, R., Polat, U., Makous, W., & Bavelier, D. (2009). Enhancing the contrast sensitivity function through action video game training. *Nat Neurosci, 12*(5), 549-551. doi:10.1038/nn.2296

List, A., Iordanescu, L., Grabowecky, M., & Suzuki, S. (2014). Haptic guidance of overt visual attention. *Atten Percept Psychophys, 76*(8), 2221-2228. doi:10.3758/s13414-014-0696-1

Lopez-Fernandez, O., Williams, A. J., Griffiths, M. D., & Kuss, D. J. (2019). Female Gaming, Gaming Addiction, and the Role of Women Within Gaming Culture: A Narrative Literature Review. *Frontiers in psychiatry, 10*, 454-454. doi:10.3389/fpsyt.2019.00454

Lopez-Fernandez, O., Williams, A. J., & Kuss, D. J. (2019). Measuring Female Gaming: Gamer Profile, Predictors, Prevalence, and Characteristics From Psychological and Gender Perspectives. *10*(898). doi:10.3389/fpsyg.2019.00898

Mastroberardino, S., Santangelo, V., & Macaluso, E. (2015). Crossmodal semantic congruence can affect visuo-spatial processing and activity of the fronto-parietal attention networks. *Frontiers in Integrative Neuroscience, 9*(45). doi:10.3389/fnint.2015.00045

Merritt, P., Hirshman, E., Wharton, W., Stangl, B., Devlin, J., & Lenz, A. (2007). Evidence for gender differences in visual selective attention. *Personality and Individual Differences, 43*(3), 597-609. doi:https://doi.org/10.1016/j.paid.2007.01.016

Nuyens, F. M., Kuss, D. J., Lopez-Fernandez, O., & Griffiths, M. D. (2019). The Empirical Analysis of Non-problematic Video Gaming and Cognitive Skills: A Systematic Review. *International Journal of Mental Health and Addiction, 17*(2), 389-414. doi:10.1007/s11469-018-9946-0

Oei, A. C., & Patterson, M. D. (2013). Enhancing Cognition with Video Games: A Multiple Game Training Study. *PLOS ONE, 8*(3), e58546. doi:10.1371/journal.pone.0058546

Oei, A. C., & Patterson, M. D. (2015). Enhancing perceptual and attentional skills requires common demands between the action video games and transfer tasks. *Frontiers in psychology, 6*, 113. doi:10.3389/fpsyg.2015.00113

Okagaki, L., & Frensch, P. A. (1994). Effects of video game playing on measures of spatial performance: Gender effects in late adolescence. *Journal of Applied Developmental Psychology, 15*(1), 33-58. doi:https://doi.org/10.1016/0193-3973(94)90005-1

Pesquita, A., Brennan, A. A., Enns, J. T., & Soto-Faraco, S. (2013). Isolating shape from semantics in haptic-visual priming. *Exp Brain Res, 227*(3), 311-322. doi:10.1007/s00221-013-3489-1

Posner, M. I., & Marin, O. S. (2016). *Attention and performance XI*: Routledge.

Powers, K. L., & Brooks, P. J. (2014). Evaluating the Specificity of Effects of Video Game Training. In F. C. Blumberg (Ed.), *Learning by Playing: Video Gaming in Education* (pp. 302–330). Oxford, UK: Oxford University Press.

Powers, K. L., Brooks, P. J., Aldrich, N. J., Palladino, M. A., & Alfieri, L. (2013). Effects of video-game play on information processing: A meta-analytic investigation. *Psychonomic Bulletin & Review, 20*(6), 1055-1079. doi:10.3758/s13423-013-0418-z

Roberts, K. L., & Hall, D. A. (2008). Examining a supramodal network for conflict processing: a systematic review and novel functional magnetic resonance imaging data for related visual and auditory Stroop tasks. *Journal of cognitive neuroscience, 20*(6), 1063-1078.

Rupp, M. A., McConnell, D. S., & Smither, J. A. (2016). Examining the Relationship Between Action Video Game Experience and Performance in a Distracted Driving Task. *Current Psychology, 35*(4), 527-539. doi:10.1007/s12144-015-9318-x

Sauter, M., Draschkow, D., & Mack, W. J. B. S. (2020). Building, Hosting and Recruiting: A Brief Introduction to Running Behavioral Experiments Online. *10*(4), 251.

Schubert, T., Finke, K., Redel, P., Kluckow, S., Müller, H., & Strobach, T. (2015). Video game experience and its influence on visual attention parameters: An investigation using the framework of the Theory of Visual Attention (TVA). *Acta Psychologica, 157*, 200-214. doi:https://doi.org/10.1016/j.actpsy.2015.03.005

Soto-Faraco, S., Kvasova, D., Biau, E., Ikumi, N., Ruzzoli, M., Morís-Fernández, L., & Torralba, M. (2019). *Multisensory Interactions in the Real World*. In Elements in Perception. Retrieved from https://www.cambridge.org/core/elements/multisensory-interactions-in-the-real-world/549C5120543E2298E1295696FBC39F3E doi:DOI: 10.1017/9781108578738

Spagna, A., Mackie, M. A., & Fan, J. (2015). Supramodal executive control of attention. *Frontiers in psychology, 6*, 65. doi:10.3389/fpsyg.2015.00065

Stewart, H., Martinez, J., Perdew, A., Green, S., & Moore, D. (2019). *Auditory cognition and perception of action video game players*.

Stoet, G. (2017). Sex differences in the Simon task help to interpret sex differences in selective attention. *Psychological research, 81*(3), 571-581. doi:10.1007/s00426-016-0763-4

Strobach, T., Frensch, P. A., & Schubert, T. (2012). Video game practice optimizes executive control skills in dual-task and task switching situations. *Acta Psychol (Amst), 140*(1), 13-24. doi:10.1016/j.actpsy.2012.02.001

Subrahmanyam, K., & Greenfield, P. M. (1994). Effect of video game practice on spatial skills in girls and boys. *Journal of Applied Developmental Psychology, 15*(1), 13-32. doi:https://doi.org/10.1016/0193-3973(94)90004-3

Tomczak, M., Tomczak, E., Kleka, P., & Lew, R. (2014). Using power analysis to estimate appropriate sample size. *Trends in Sport Sciences, 21*(4), 195-206.

Torner, H. P., Carbonell, X., & Castejón, M. (2019). A comparative analysis of the processing speed between video game players and non-players. *Aloma: Revista de Psicologia, Ciències de l'Educació i de l'Esport, 37*(1).

Wong, N. H. L., & Chang, D. H. F. (2018). Attentional advantages in video-game experts are not related to perceptual tendencies. *Scientific Reports, 8*(1), 5528. doi:10.1038/s41598-018-23819-z

Xuemin, Z., & Bin, Y. (2010). Effects of action video game on spatial attention distribution in low and high perceptual load task. *Journal of Software, 5*(12), 1434-1441.

Yuval-Greenberg, S., & Deouell, L. Y. (2009). The dog’s meow: asymmetrical interaction in cross-modal object recognition. *Experimental Brain Research, 193*(4), 603-614. doi:10.1007/s00221-008-1664-6

Zhang, Y.-X., Tang, D.-L., Moore, D. R., & Amitay, S. (2017). Supramodal Enhancement of Auditory Perceptual and Cognitive Learning by Video Game Playing. *8*(1086). doi:10.3389/fpsyg.2017.01086

Statement of contribution

Mohammad Hamzeloo, Salvador Soto-Faraco, and Daria Kvasova contributed to the conception and design of the study. Daria Kvasova and Mohammad Hamzeloo prepared the stimuli. Mohammad Hamzeloo collected the data, performed the statistical analysis, and wrote the TFM. All authors contributed to the TFM revision and read and approved the final version of the TFM.

Appendix A

Demographic form

Age	
Gender* (Enter M for Male, F for female and O for others)	
Are you Right or Left-handed?* (Enter R if you are Right-handed or L if you are Left-handed)	
Do you have a normal or corrected-to-normal vision?* (Enter Y for Yes and N for No)	
From which country are you participating?* (Enter the name of your current country)	
What kind of browser are you using? (It is recommended to use Google Chrome or Firefox.	

Appendix B

Video Game Questionnaire – version November 2019

Video Game Playing Questionnaire – *DURING* THE PAST YEAR

Ss ID: ______________________________ Date: ______________

THIS QUESTION IS ABOUT YOUR GAMING ACTIVITIES DURING THE PAST 12 MONTHS. For each type of game, specify the number of hours played per week **when you played the most**. You will be given the choice between 6 different answers; Please choose one and only one: **never ;** **less than 1 hour ;** **between 1 and 3 hours ;** **between 3 and 5 hours ;** **between 5 and 10 hours ;** **more than 10 hours** *If your answer is at the cut-off, please select the lower option (e.g. if 3 hours per week, select 1 to 3)* *If you are playing Fortnite and/or Minecraft, please classify as shown:* Fortnite: Save the World or Battle Royale = First/Third Person Shooters Fortnite Creative = Puzzle Minecraft = Puzzle	**Hours/week**					
	never	less than 1 hour	between 1 and 3 hours	between 3 and 5 hours	between 5 and 10 hours	more than 10 hours
FIRST/THIRD PERSON SHOOTERS *(Call Of Duty, Halo, Battlefield, Half-Life, Overwatch, Counterstrike ...)*						
ACTION-RPG/ADVENTURE *(The Witcher, Mass Effect, Fallout 4, Skyrim, GTA, Assassin's Creed, Tomb Raider, The Last of Us, ...)*						
SPORTS/DRIVING *(Fifa, NHL, Mario Kart, Need for Speed, Forza, ...)*						
REAL-TIME STRATEGY/MOBA *(Starcraft, Warcraft I, II & III, DotA, Command & Conquer, League of Legends, Age of Empires, ...)*						
TURN-BASED/NON-ACTION ROLE-PLAYING/FANTASY *(World of Warcraft, Final Fantasy, Fable, Pokemon, Dragon Age, ...)*						
TURN-BASED STRATEGY/LIFE SIMULATION/PUZZLE *(Civilization, Hearthstone, The Sims, Restaurant Empire, Puzzle Quest, Bejeweled, Solitaire, Candy Crush, ...)*						
MUSIC GAMES *(Guitar Hero, DDR, Rock Band, ...)*						
OTHER : games that don't fit into any of the categories *(Phone games, Browser games, Fighting games, etc.)*						