Video Game Use and Cognitive Performance: Does It Vary with the Presence of Problematic Video Game Use?

Emily Collins, PhD, and Jonathan Freeman, PhD

Abstract

Action video game players have been found to outperform nonplayers on a variety of cognitive tasks. However, several failures to replicate these video game player advantages have indicated that this relationship may not be straightforward. Moreover, despite the discovery that problematic video game players do not appear to demonstrate the same superior performance as nonproblematic video game players in relation to multiple object tracking paradigms, this has not been investigated for other tasks. Consequently, this study compared gamers and nongamers in task switching ability, visual short-term memory, mental rotation, enumeration, and flanker interference, as well as investigated the influence of self-reported problematic video game use. A total of 66 participants completed the experiment, 26 of whom played action video games, including 20 problematic players. The results revealed no significant effect of playing action video games, nor any influence of problematic video game play. This indicates that the previously reported cognitive advantages in video game players may be restricted to specific task features or samples. Furthermore, problematic video game play may not have a detrimental effect on cognitive performance, although this is difficult to ascertain considering the lack of video game player advantage. More research is therefore sorely needed.

Introduction

As the technology underlying video games has advanced, the level of visual detail, the realism, and the cognitive demands on the player have drastically increased. Video games have also seen a huge rise in popularity, and they are now more widespread than ever before, resulting in a record number of individuals regularly exposing themselves to incredibly visually complex stimuli. It is well documented that visual systems can be altered by exposure to different visual environments, resulting in improved performance in related perceptual tasks. However, while in these contexts the improvements are specific to the trained task, a growing body of research has begun to suggest that action video games can train cognitive and perceptual skills that generalize to other paradigms and to untrained parameters. This has been argued to be due to the requirements of the game, such as the need to process and respond quickly to visual and auditory cues and the risk of punishments for not doing so. The transferable nature of these abilities may also be due to the flexibility in the kinds of training video games provide, as increasing the need for switching between task demands and the requirements for adjustments in the approach or skills used has been suggested to encourage a greater level of transfer to novel situations.

This has been well supported by the literature, which has demonstrated that action video game players as well as nonplayers trained using action video games have a greater attentional capacity, as measured by both flanker compatibility and enumeration paradigms, greater mental rotation, executive control (normally measured by task switching performance, for which participants are required to alternate quickly between two conflicting tasks), and visual short-term memory (VSTM), to name but a few. However, these discoveries have not gone undisputed, and several studies have reported there to be no significant differences between VGPs and NVGPs. This has been in regards to flanker interference, enumeration performance, and other measures that should be related to ones previously found to differentiate the two groups, such as visuospatial cognition and certain task switching paradigms. Many of these authors have argued that this could be due to video games only improving performance in certain tasks, and not the underlying processes they purportedly measure. This

Department of Psychology, Goldsmiths, University of London, London, United Kingdom.
may be particularly true for task switching performance, as Karle et al. found that VGs only outperformed NVGPs in paradigms with low cross task interference, meaning that the response keys for the two separate tasks had to be different. Moreover, research into mental rotation ability has been mixed in deciding whether males or females (or indeed both) stand to benefit from video game use. It is therefore clear that more research is sorely needed.

One factor that has been largely ignored in this area is the presence of problematic video game play (PVGP). This refers to the existence of addiction-like symptoms relating to video games, usually identified by meeting the criteria outlined by the Diagnostic Statistical Manual markers for addiction. These are salience, mood modification, tolerance, withdrawal, conflict, and relapse. There is substantial variation in estimated prevalence rates, although a recent meta-analysis suggested that the average percentage of those affected is around 6% of gamers, and PVGP has been argued to be distinct from casual or even excessive game play in terms of associations and effects. Evidence of this dissociation between problematic and nonproblematic game play has been generated by Sun et al. who assessed performance on a multiple object tracking (MOT) task. It was discovered that while participants who reported to have previously played video games problematically outperformed nonplayers, they also outperformed those who were currently experiencing PVGP. Current PVGPs did not significantly differ from the control group of nonplayers. This signifies that the presence of problematic play may in fact impede the development or the demonstration of such benefits, or that disruptions in related aspects of cognitive function increase the likelihood of the development of PVGP. However, the lack of research in this area makes it unclear whether this issue is restricted to MOT or if it generalizes to tasks relying on similar cognitive processes. Indirect evidence offers some indication that generalization is possible, as reduced gray matter volumes have been found in PVGPs compared to NPVGPs. The authors claim that as a result, problematic and nonproblematic groups differ in the ability of the left anterior cingulate to maintain attention and control executive functions, and this therefore suggests this might be reflected in differences in cognitive task performance.

From the reviewed literature, it is clear that there are many unexplained discrepancies indicating an urgent need for replication and further investigation into PVGP.

**Hypotheses**

Based on the above literature, it was hypothesized that NPVGPs would outperform NVGPs in terms of greater accuracy in all of the paradigms. Due to the previously suggested potential for PVGP either to restrict these advantages or to be a result of poorer cognitive performance, it was also predicted that PVGPs would differ from NVGPs in performance, demonstrating significantly lower accuracy across the tasks than NPVGPs. Overall, it was also predicted that VGs in general would also demonstrate superior accuracy than NVGs.

**Methodology**

**Participants**

The final data set consisted of 66 participants (15 female), aged from 18 to 47 years ($M=22.61$, $SD=4.48$), although some failed to specify their age. Participants were recruited through departmental e-mails, posters in university buildings, and a participation scheme for course credit. The vast majority were students (86.57%), with the remaining working full time (5.97%), part time (2.99%), self-employed (2.99%), or unemployed (1.49%). There were 20 NVGPs and 46 VGPs. The VG group was made up of 20 PVGPs (playing an average of 6.92 hours of action games per week, $SD=6.07$) and 26 NPVGs (playing an average of 6.86 hours per week, $SD=6.67$), as classified by a self-report measure described in more detail in the section “Questionnaire measures.” Highly engaged gamers were specifically targeted in order to increase the likelihood of attracting a sufficient number of PVGs for the analysis, resulting in a larger proportion of problematic gamers than would be expected in the normal population. Out of those who played games, the average score for the PVGP measure following recoding (meaning participants were awarded one point for meeting each criterion, with a maximum score of six) was 2.71 ($SD=2.01$).

Ages were similar across the groups ($M=21.93$, $SD=3.15$; $M=22$, $SD=3.4$; and $M=23.85$, $SD=6.10$ for NVGPs, NPVGs, and PVGPs respectively). The differences between the groups in terms of age were therefore not significant, $F(2,58)=1.16$, $p=0.32$. Although the groups did differ in the distributions of gender (NVGPs consisted of 10 females, NPVGs three, and the PVGP group of only two; $\chi^2=12.17$, $p<0.01$, $t$ tests revealed no significant differences when comparing males and females in all of the measures of cognitive performance, except for accuracy for eight items in the enumeration paradigm, $F(1,61)=8.01$, $p<0.05$, the reaction time based switch cost for the task switching paradigm $F(1,61)=7.57$, $p<0.01$, and accuracy for the second rotation in the mental rotation task, $F(1,61)=5.02$, $p<0.05$. While not ideal, all of these outcomes were analyzed in combination with other measures that did not differ between the genders (for instance, other rotations, alternative measures of performance), minimizing the potential influence of gender on the overall findings.

**Materials**

The cognitive tasks were the same as those used and created by Boot et al. with only minor alterations, which will be noted. All were run using E-Prime v1.2, and they were performed on a 17” Dell Vostro laptop with a refresh rate of 60Hz. Tasks are summarized in Figure 1, and the outcome measures of accuracy and reaction times constituted the dependent variables.

The enumeration paradigm involved the presentation of between one and nine white dots (diameter 0.25”) on a black background. Participants were required to respond according to how many dots they had viewed by pressing the corresponding keyboard key. The stimuli were positioned randomly, with the restriction that none could appear in the central location. A total of 32 practice trials were completed before 160 test trials.

The mental rotation task displayed two shapes that were either identical (with one rotated) or mirror images of each other, and participants had to respond using the keyboard according to which was the case. The stimuli were based on the shapes used in Tetris, and there were eight possible rotations ranging from 0° to 315°, except for the Z or backwards
Z shapes, which rotated a maximum of 135° due to issues with how they would appear at greater rotations. The shapes were shown 3° from the fixation point and measured approximately 2.4° × 2.4°. A small-scale pilot experiment indicated that only 10 rather than the original 30 practice trials were required before the 128 test trials.

The task switching paradigm required participants to judge whether a number displayed on the screen was high or low (above or below five) or odd or even, depending on the color of the background. Due to potential ambiguities in the required response, the number five was not presented. The task consisted of one block of each individual task (30 trials each) before 30 practice joint trials and then 160 test trials. The task for each trial was chosen at random.

In the VSTM task, participants had to decide whether a displayed item was present in the previous array by pressing the corresponding key on the computer keyboard. In one block, the items were hollow shapes; in the second, they were squares containing different colors; in the third, the items were a combination of the two (e.g., a yellow star, a blue square). Each block involved 4 practice and 68 experimental trials, making a total of 12 practice and 204 experimental trials.

For the flanker compatibility task, the aim was to indicate what direction the center arrow was pointing in by pressing the corresponding keyboard key, while ignoring the direction of the flanker arrows. In half of the trials, these flankers were pointing in the same direction as the target arrow (the compatible condition), and in the other half, the arrows were pointing in the opposite direction (the incompatible condition). After 20 practice trials, participants completed 100 experimental trials.

Questionnaire measures. In addition to the computer tasks, participants were also asked to complete an online questionnaire assessing video game use. This took the form of reporting how many hours per week video games of various genres were played.

PVGU was measured by the short version of the Game Addiction (GA) scale. This scale demonstrated good reliability (α = 0.823) and consisted of seven items. In line with previous classification systems, participants who agreed to four or more items were deemed problematic gamers.

Procedure

The online questionnaire was completed in the participants’ own time using EFS software (www.unipark.de). The computer tasks took 1 hour to complete, and the tasks were attempted in a random order, as dictated by a random order generator. This was to avoid potential order effects. Testing took place in university owned rooms across a number of different universities.

Results

Previous research has recommended a focus on accuracy data due to differences in the speed of motor responses between VGPs and NVGPs. Therefore, although reaction time data are presented, they are interpreted with caution.

Analyses of variance (ANOVAs) were used to investigate differences in performance between gamers and nongamers, as well as NPVGPs and PVGPs, with group membership entered as a between subjects variable. Greenhouse–Geisser corrections were applied where appropriate.

Problematic video game players

The data from each paradigm were analyzed separately, although for all, PVG status (NVGP, NPVGP, or PVGP) was entered as a between subjects variable.
Enumeration. For this task, the number of items (numerosity) was entered as a within subjects variable. There was a significant effect of numerosity, with lower accuracies, $F(1, 59) = 148.94, p < 0.01, \eta^2_p = 0.938$, and longer reaction time, $F(1, 58) = 120.04, p < 0.01, \eta^2_p = 0.674$, for higher numbers. However, there was no effect of group membership (PVGP, NPVGP, or NVGP; see Fig. 2).

Flanker compatibility. Trial type (whether the flankers were compatible or compatible) was entered as a within subjects variable. As predicted by previous research, there was a significant effect of trial type; participants were more accurate, $F(1, 60) = 41.063, p < 0.01, \eta^2_p = 0.406$, and demonstrated quicker responses, $F(1, 60) = 551.715, p < 0.01, \eta^2_p = 0.902$, for compatible over incompatible trials. There were no significant effects of PVGP status or any significant interactions (see Fig. 3).

In line with the analyses conducted in Irons et al., internal comparisons were also performed for each group individually using $t$ tests to investigate whether the groups differed in the effects of the trial type. The level of flanker interference in terms of reaction times and accuracy were also compared. However, the groups showed similar flanker compatibility effects, and so no significant differences were found.

Task-switching. The trial type (whether it was a switch or repeat trial) was entered as a within subjects variable and was found to have a significant effect on reaction times. This took the form of longer reaction times for switch over repeat trials, $F(1, 59) = 227.11, p < 0.01, \eta^2_p = 0.794$. There was, however, no effect on accuracy and no significant effect of PVGP status (see Fig. 4), nor any related interactions.

In accordance with the analyses conducted on the flanker compatibility data, internal comparisons of switch and repeat trials were also conducted for both groups individually. Once again, the groups showed the same pattern with switch costs evident in reaction times but not accuracy.

Mental rotation. The within subjects variable for this task was the degree of rotation, which showed a significant effect on accuracy $F(1, 59) = 15.936, p < 0.01, \eta^2_p = 0.213$, and reaction times $F(1, 59) = 66.744, p < 0.01, \eta^2_p = 0.527$; accuracy was lower and reaction times longer for the middle rotations over the lowest and highest rotations. Although there was no significant effect of group membership on accuracy, there was an effect on reaction time, $F(2, 60) = 4.247, p < 0.05, \eta^2_p = 0.124$. Pairwise comparisons demonstrated this difference to be between NVGPs and NPVGPs, with NPVGPs showing faster reaction times for the first ($t(21) = 2.071, p < 0.05, d = 0.894$), fifth ($t(22) = 2.133, p < 0.05, d = 0.914$), sixth ($t(20) = 2.085, p < 0.05, d = 0.944$), and seventh rotations.
The hypotheses were not upheld, and the present results would indicate that video game use does not influence cognitive performance. The lack of group differences remained, even when restricting the gaming sample to only “frequent” gamers (those playing 7 hours or more per week), indicating that a lack of gaming experience in this group was not responsible for the unexpected results. Moreover, PVGP was not found to influence performance, with PVGs not differing from NPGVs or NVPVs. This therefore runs counter to existing theories that video games are able to train these processes, as this should have been evidenced in superior performance in VGs when compared to NPGVs. This also contradicts the theory that the presence of PVGP either prevents the VGP advantage being demonstrated, or that disruptions in related aspects of cognitive function increase the likelihood of the development of PVGP. As all of the expected effects indicating normal performance were present in the data (e.g., the effects of rotation degree for the mental rotation paradigm, flanker compatibility, numerosity in the enumeration task, and trial type for the task switching measure), this was presumably not due to methodological issues in the tasks themselves. The reasons behind the finding that VGPVs do not generally differ from NPGVs therefore lie elsewhere.

Although unexpected due to the volume of research in favor of VGP advantages, this is not the only study to have failed to replicate some of these findings. For instance, similarly nonsignificant results have been generated in relation to enumeration,13 flanker compatibility,17 and task switching.19 This appears largely to support the notion that specific task features are important in generating significant group differences. In the case of flanker compatibility, Irons et al.17

**Table 2. Mean Accuracy Percentages Across Cognitive Tasks for Nongamers and Frequent Gamers**

<table>
<thead>
<tr>
<th>Cognitive Task</th>
<th>Nongamers</th>
<th>Frequent Gamers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enumeration</td>
<td>M 74.93</td>
<td>79.005</td>
</tr>
<tr>
<td></td>
<td>SE -2.37</td>
<td>-2.07</td>
</tr>
<tr>
<td>Flanker compatibility</td>
<td>M 96.05</td>
<td>96.19</td>
</tr>
<tr>
<td></td>
<td>SE -0.98</td>
<td>-0.82</td>
</tr>
<tr>
<td>Task switching</td>
<td>M 84.83</td>
<td>79.72</td>
</tr>
<tr>
<td></td>
<td>SE -2.81</td>
<td>-2.34</td>
</tr>
<tr>
<td>Mental rotation</td>
<td>M 87.93</td>
<td>91.82</td>
</tr>
<tr>
<td></td>
<td>SE -1.84</td>
<td>-1.53</td>
</tr>
<tr>
<td>VSTM</td>
<td>M 73.8</td>
<td>76.35</td>
</tr>
<tr>
<td></td>
<td>SE -2</td>
<td>-1.33</td>
</tr>
</tbody>
</table>

Note. For flanker compatibility and task switching paradigms, the means refer to accuracy in the incompatible and switch trials (respectively), as these were the trials in which a differences was expected. VSTM, visual short-term memory.

Visual short-term memory. The block type (color, shape, or a combination of the two) was entered as a within subjects variable. There was a significant effect of block type, indicating that participants were more accurate, \( F(1, 50) = 111.935, p < 0.01, \eta^2_p = 0.691 \), and faster for trials from the color only block, \( F(2, 51) = 10.478, p < 0.01, \eta^2_p = 0.170 \). Once again, however, there was no effect of PVGP status or any interactions (see Fig. 6).

**Table 1. Mean Accuracy Percentages Across Cognitive Tasks for Nongamers and Gamers, Irrespective of PVGP**

<table>
<thead>
<tr>
<th>Cognitive Task</th>
<th>Nongamers</th>
<th>Gamers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enumeration</td>
<td>M 74.927</td>
<td>79.005</td>
</tr>
<tr>
<td></td>
<td>SE 2.292</td>
<td>1.524</td>
</tr>
<tr>
<td>Flanker compatibility</td>
<td>M 94</td>
<td>91.227</td>
</tr>
<tr>
<td></td>
<td>SE 1.984</td>
<td>1.304</td>
</tr>
<tr>
<td>Task switching</td>
<td>M 85.721</td>
<td>81.383</td>
</tr>
<tr>
<td></td>
<td>SE 3.059</td>
<td>1.957</td>
</tr>
<tr>
<td>Mental rotation</td>
<td>M 87.928</td>
<td>90.994</td>
</tr>
<tr>
<td></td>
<td>SE 2.036</td>
<td>1.302</td>
</tr>
<tr>
<td>VSTM</td>
<td>M 73.797</td>
<td>76.353</td>
</tr>
<tr>
<td></td>
<td>SE 2.002</td>
<td>1.328</td>
</tr>
</tbody>
</table>

Note. For flanker compatibility and task switching paradigms, the means refer to accuracy in the incompatible and switch trials (respectively), as these were the trials in which a differences was expected. VSTM, visual short-term memory.

Visual short-term memory. The block type (color, shape, or a combination of the two) was entered as a within subjects variable. There was a significant effect of block type, indicating that participants were more accurate, \( F(1, 50) = 111.935, p < 0.01, \eta^2_p = 0.691 \), and faster for trials from the color only block, \( F(2, 51) = 10.478, p < 0.01, \eta^2_p = 0.170 \). Once again, however, there was no effect of PVGP status or any interactions (see Fig. 6).

**Table 1. Mean Accuracy Percentages Across Cognitive Tasks for Nongamers and Gamers, Irrespective of PVGP**

<table>
<thead>
<tr>
<th>Cognitive Task</th>
<th>Nongamers</th>
<th>Gamers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enumeration</td>
<td>M 74.927</td>
<td>79.005</td>
</tr>
<tr>
<td></td>
<td>SE 2.292</td>
<td>1.524</td>
</tr>
<tr>
<td>Flanker compatibility</td>
<td>M 94</td>
<td>91.227</td>
</tr>
<tr>
<td></td>
<td>SE 1.984</td>
<td>1.304</td>
</tr>
<tr>
<td>Task switching</td>
<td>M 85.721</td>
<td>81.383</td>
</tr>
<tr>
<td></td>
<td>SE 3.059</td>
<td>1.957</td>
</tr>
<tr>
<td>Mental rotation</td>
<td>M 87.928</td>
<td>90.994</td>
</tr>
<tr>
<td></td>
<td>SE 2.036</td>
<td>1.302</td>
</tr>
<tr>
<td>VSTM</td>
<td>M 73.797</td>
<td>76.353</td>
</tr>
<tr>
<td></td>
<td>SE 2.002</td>
<td>1.328</td>
</tr>
</tbody>
</table>

Note. For flanker compatibility and task switching paradigms, the means refer to accuracy in the incompatible and switch trials (respectively), as these were the trials in which a differences was expected. VSTM, visual short-term memory.
report that as group differences were able to be manipulated by aspects of the task design (albeit nonsignificantly), VGPs may only outperform NVGPs in specific paradigms with certain features. Similarly, Boot et al.\textsuperscript{13} also failed to replicate the VGP advantage in terms of enumeration performance using the same task as the present study. This once again suggests that paradigm-specific attributes may be responsible. The serial counting process argued to be superior in VGPs\textsuperscript{1} is also more influenced by the grouping and geometry of the presented items,\textsuperscript{26,27} indicating that subtle task variations may produce divergent results. In relation to task switching, Karle et al.\textsuperscript{19} found that although in one experiment VGPs did indeed show a smaller switch cost than NVGPs, this advantage was eradicated by increasing the task interference by keeping the response keys the same for both of the tasks. This would therefore provide further support for the notion that executive control is not improved in VGPs but that, possibly, superior performance in some task switching paradigms is in fact indicative of improved selective attention. More generally, this has implications for the theoretical standpoint that video games are capable of improving processes across all measurements of the outlined processes, as this would appear not to be the case, suggesting that VGPs may be improving performance through some mechanism other than the process being measured.

However, task design features cannot be responsible for all the discrepancies between the present study and previous literature. Although the use of two-dimensional stimuli in the mental rotation paradigm does not resemble the three-dimensional environments present in many video games, Boot et al.\textsuperscript{13} did report significantly superior performance in VGPs over NVGPs using the same two-dimensional paradigm, suggesting that this criticism may be unfounded. The same occurred for VSTM, and consequently, this could be symptomatic of variations within the sample, for instance in terms of the level of video game use. Alternatively, it may be that this is not a reliable and easily replicable finding, although more research would be needed to draw this conclusion.

More generally, it was predicted that PVGP would have a detrimental influence on cognitive performance, which was not supported. However, as no differences in accuracy between VGPs and NVGPs were found, it is difficult to interpret whether PVGPs and NPVGP should vary to a greater degree or not to be in line with this prediction. This is because the findings of Sun et al.\textsuperscript{21} suggested that the group differences were more a lack of cognitive advantage gained from video game play rather than a reduction in performance per se, something that is hard to assess in the absence of a demonstrable cognitive advantage. The present results, however, would indicate that based on performance on the paradigms used, PVGP is not an influential factor and is not associated with variations in reaction times or accuracy compared to either NPVGP or NVGP. This is in some ways a positive finding, as it highlights that the negative effects of PVGP may be restricted to the more interpersonal or social aspects of an individual’s life, and may not extend to cognitive processes.

However, it is worth noting that in the present sample, PVGP's and NPVGP's did not differ in regards to playing time. This may be an indication that the PVGP group were relatively unproblematic, especially as the mean number of hours played per week for this group was rather low (6.92 hours). That said, this remains a controversial point; while some rely solely on playing time for an indication of PVGP,\textsuperscript{28} or use this information to form a composite measure with a PVGP scale,\textsuperscript{29} most theories of PVGP argue that it is an entirely separate issue from excessive playing and therefore cannot be captured by measures of the extent of play.\textsuperscript{30} This is because an individual may develop a problematic relationship with games irrespective of the number of hours they spend playing.\textsuperscript{31} as the average PVGP score for the PVGP group would indicate in the present sample. Despite this argument, this does mean that these findings are unlikely to generalize to situations in which the extent of game play is used as a measure of PVGP.

Taken together, these findings suggest that the influence of video games on cognition is not as widespread as previous research may suggest, and that the VGP advantage may only be apparent in certain subpopulations or in relation to specific paradigms. This could be a result of divergent VGP categorizations, recruitment strategies,\textsuperscript{32} or an unknown variable that influences whether cognitive processes are improved by video game play. Previous studies have identified considerable individual differences, which goes some way to suggest the relationship between video games and cognition is not straightforward.\textsuperscript{33} The theoretical implications for this are therefore that the VGP advantage is not universal, calling into question whether the cognitive process itself is being trained by the video games, or whether it is in fact some lower level ability that dictates performance in some paradigms and not others. Conversely, it may only be evident in certain populations, and may rely on the presence of certain individual differences, making the possibility for video games to be actively used to improve these processes more restricted than may have been anticipated. Future research may therefore wish to attempt to circumvent some of the present limitations that are widespread in the literature, for instance by recruiting outside of university environments, not specifying that VGP and NVGP samples are being sought, and recruiting a larger number of participants in order to investigate whether group differences exist in certain demographics and not in others.

Author Disclosure Statement

No competing financial interests exist.

References


Address correspondence to:
Dr. Emily Collins
Goldsmiths
University of London
Department of Psychology
London
SE14 6NW
United Kingdom
E-mail: emilyimcollins@gmail.com