

# Action Video Games Do Not Improve the Speed of Information Processing in Simple Perceptual Tasks

Don van Ravenzwaaij<sup>1</sup>, Wouter Boekel<sup>2</sup>, Birte U. Forstmann<sup>2</sup>, Roger Ratcliff<sup>3</sup>, and Eric-Jan Wagenmakers<sup>2</sup>

<sup>1</sup>University of Newcastle <sup>2</sup>University of Amsterdam <sup>3</sup>Ohio State University

Word count: 6,409

Correspondence concerning this article should be addressed to:

Don van Ravenzwaaij

University of Newcastle, Department of Psychology

University Drive, Aviation Building, room AVG11

Callaghan NSW 2308, Australia

Ph: (+61) 2-4921-5662

E-mail should be sent to don.vanravenzwaaij@newcastle.edu.au.

## Abstract

Previous research suggests that playing action video games improves performance on sensory, perceptual, and attentional tasks. For instance, Green, Pouget, and Bavelier (2010) used the diffusion model to decompose data from a motion detection task and estimate the contribution of several underlying psychological processes. Their analysis indicated that playing action video games leads to faster information processing, reduced response caution, and no difference in motor responding. Because perceptual learning is generally thought to be highly context-specific, this transfer from gaming is surprising and warrants corroborative evidence from a large-scale training study. We conducted two experiments in which participants practiced either an action video game or a cognitive game in five separate, supervised sessions. Prior to each session and following the last session, participants performed a perceptual discrimination task. In the second experiment we included a third condition in which no video games were played at all. Behavioral data and diffusion model parameters showed similar practice effects for the action gamers, the cognitive gamers, and the non-gamers and suggest that, in contrast to earlier reports, playing action video games does not improve the speed of information processing in simple perceptual tasks.

**Keywords:** Action Games, Diffusion Model, Probabilistic Inference, Moving Dots Task.

Video games are immensely popular: they are played in 67% of US households and keep the average gamer occupied for an estimated eight hours a week.<sup>1</sup> One of the most popular genres in video gaming is the so-called *action video game*. Typically, action video games revolve around violent battles in war-like situations. Although this type of video game has been developed for entertainment purposes only, a growing body of research suggests that playing these games improves performance on a wide range of perceptual and cognitive tasks (e.g., Green & Bavelier, 2012). This suggestion is surprising — the effects of perceptual learning tend to be highly context-dependent and therefore fail to generalize broadly (Fahle, 2005; but see Liu & Weinshall, 2000; Jeter, Doshier, Petrov, & Lu, 2009). The presence of transfer effects from action video game playing may have profound societal, financial, and logistic ramifications; for instance, action video game playing may moderate cognitive decline in the elderly, assist the recovery of stroke patients, or be part of special needs educational programs. These real-life ramifications mean it is incumbent on the field to critically assess the existing evidence for the benefit of action video game playing.

In a letter to *Nature*, Green and Bavelier (2003) discussed advantages of video game players (VGPs) over non-video game players (NVGPs) in a compatibility task, an enumeration task, a spatial attention task, and an attentional blink task. Furthermore, to rule out the possible confound of pre-existing differences between VGPs and NVGPs, Green and Bavelier (2003) conducted a training experiment in which NVGPs were required to play an action game, *Medal of Honor*, for one hour per day on ten consecutive days. As a control condition, another group of NVGPs played *Tetris*, a game requiring visuo-spatial skills instead. The authors concluded that compared to the control group, the NVGPs trained on the action game improved more on the enumeration task, the spatial attention task, and the attentional blink task.

A number of subsequent studies have examined the effect of training on action video games. For example, in a study by Li, Polat, Makous, and Bavelier (2009), playing action video games led to enhanced visual contrast sensitivity whereas playing a non-action control video game (henceforth referred to as cognitive game) did not. In addition, Green and Bavelier (2006) showed how training on action video games, compared to training on cognitive games, improves performance on multiple object tracking tasks.

A further study by Feng, Spence, and Pratt (2007) reported that pre-existing gender differences in spatial attention disappear after as little as ten hours of action video game training. Specifically, women benefitted more from training in action video games than did men. Participants who were trained on a cognitive game showed no such improvements.

Moreover, in a study by Schlickum, Hedman, Enochsson, Kjellin, and Felländer-Tsai (2009), it was reported that training on action video games improved performance of medical students on a virtual reality surgery task. Note though that the authors also found some benefits for training on a cognitive game.

---

<sup>1</sup>For these and other statistics on gaming see for instance <http://www.esrb.org/about/video-game-industry-statistics.jsp>.

Finally, video games also appear to benefit the elderly. For instance, a study by Drew and Waters (1986) showed that playing arcade video games improved manual dexterity, eye–hand coordination, RTs, and other perceptual–motor skills among residents of an apartment house for elderly citizens. Clark, Lanphear, and Riddick (1987) found that playing video games improved performance of elderly adults in a two–choice stimulus–response compatibility paradigm.

In sum, video gaming seems to improve performance on a range of different tasks. In an attempt to pinpoint the locus of the improvement, Green et al. (2010) conducted a study in which they compared performance of VGPs and NVGPs with the help of two mathematical decision making models: the diffusion model (Ratcliff, 1978; Ratcliff & McKoon, 2008) and a neural decision maker model (Beck et al., 2008). The authors found that VGPs outperformed NVGPs on a visual motion discrimination task and an auditory discrimination task. Furthermore, the authors concluded that the advantage of VGPs over NVGPs is caused by a higher rate of information processing, whereas VGPs had a lower response caution than NVGPs, and motor processing was unaffected.

At first glance, the evidence in favor of action gaming benefits seems compelling. However, Boot, Blakely, and Simons (2011) warn against confounds and pitfalls that most of the above studies fall prey to. Such pitfalls include overt recruiting (creating differing demand characteristics), unspecified recruiting methods, no tests of perceived similarity between tasks and games, and possible differential placebo effects. All studies without a training regimen furthermore suffer from the confound of possible pre–existing differences between VGPs and NVGPs in aptitude on perceptual learning tasks. The studies that did use a training paradigm often feature only two measurement occasions (i.e., prior to training and after all training), a coarse design in which any impact of video game playing on performance cannot be traced over time as it develops. Finally, many training studies do not supervise game–play, making it near impossible to confirm the extent to which participants have fulfilled their training requirements.

Aside from these issues, a number of studies fail to find any reliable benefit of playing video games (Boot, Kramer, Simons, Fabiani, & Gratton, 2008; Irons, Remington, & McLean, 2011; Murphy & Spencer, 2009). Perhaps most importantly, a recently published meta–analysis by Powers, Brooks, Aldrich, Palladino, and Alfieri (in press) provides a summary of effect sizes by game type. They report Cohen’s *ds* for the benefit of a range of performance measures for the following game types: action/violent = 0.22 [0.13 – 0.30], mimetic = 0.95 [0.66 – 1.23], non–action = 0.52 [0.31 – 0.73], and puzzle = 0.30 [0.16 – 0.45]. Thus, according to the meta–analysis, effect sizes for action video games are smaller than for any other type of video game. It appears, therefore, that the final verdict on the benefit of action video game playing is still pending.

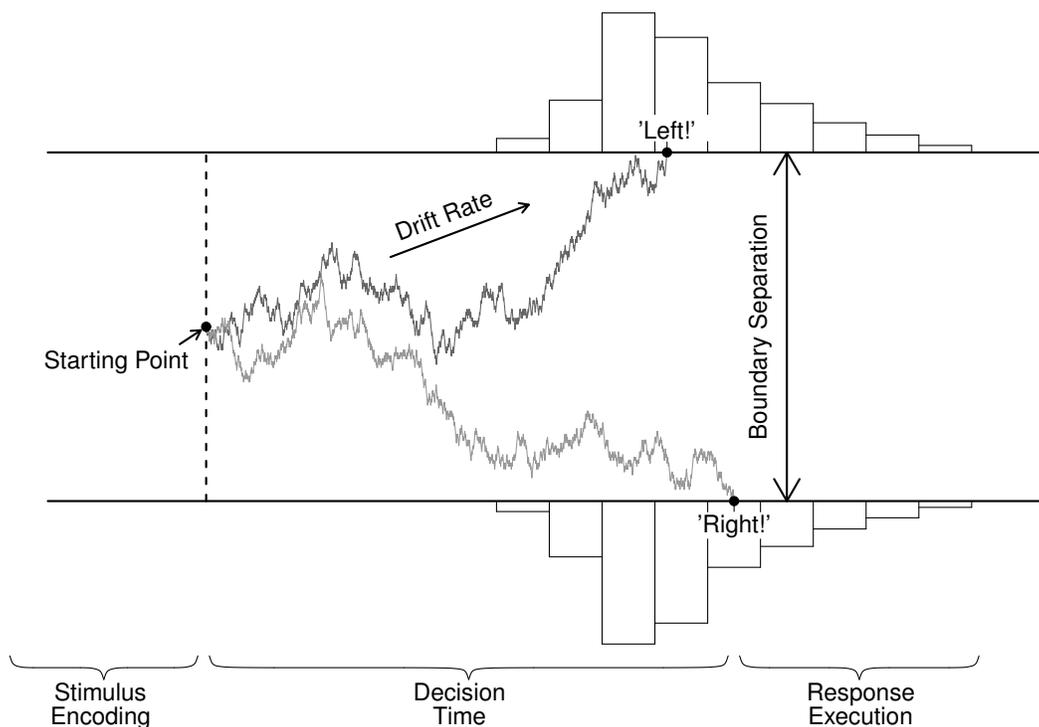
If action video games are to be used as a training method to improve perceptual and cognitive abilities, we need to be certain that they result in a tangible benefit; after all, we do not want to force grandmothers, stroke patients, and children with autism to spend their time shooting up aliens for nothing. Thus, the purpose of this study is to investigate the two claims made by Green et al. (2010): Does action video game playing improve performance on perceptual tasks? And if so, does this benefit reside in a higher rate of information processing?

In two experiments, we addressed these fundamental issues by administering a train-

ing design to two or three groups of randomly assigned participants. Two of the groups played video games under supervision; one group trained on an action video game and one group trained on a cognitive game. In the second experiment, an additional third group served as a control. During training, we repeatedly measured performance on a perceptual discrimination task. We analyze the behavioral data of this task but also examine the data through the lens of the diffusion model (Ratcliff, 1978). In the next section, we first introduce the diffusion model. Then, we will discuss Experiments 1 and 2 and conclude by discussing the general ramifications of our results.

### The Diffusion Model

In the diffusion model for speeded two-choice tasks (Ratcliff, 1978; Wagenmakers, 2009; van Ravenzwaaij & Oberauer, 2009), stimulus processing is conceptualized as the accumulation of noisy information over time. A response is initiated when the accumulated evidence reaches a predefined threshold (Figure 1).



*Figure 1.* The diffusion model and its parameters as applied to the moving dots task. Evidence accumulation begins at  $z$ , proceeds over time guided by drift rate  $v$ , and halts whenever the upper or the lower boundary is reached. Boundary separation  $a$  quantifies response caution. Observed RT is an additive combination of the time during which evidence is accumulated and non-decision time  $T_{er}$ .

The diffusion model assumes that the decision process starts at  $z$ , after which infor-

mation is accumulated with a signal-to-noise ratio that is governed by drift rate  $\xi$ , normally distributed over trials with mean  $v$  and standard deviation  $\eta$ .<sup>2</sup> Values of  $\xi$  near zero produce long RTs and high error rates. Boundary separation  $a$  determines the speed-accuracy tradeoff; lowering  $a$  leads to faster RTs at the cost of a higher error rate. Together, these parameters generate a distribution of decision times  $DT$ . The observed RT, however, also consists of stimulus-nonspecific components such as response preparation and motor execution, which together make up non-decision time  $T_{er}$ . The model assumes that  $T_{er}$  simply shifts the distribution of  $DT$ , such that  $RT = DT + T_{er}$  (Luce, 1986). The model specification is completed by including parameters that specify across-trial range in starting point,  $s_z$ , and non-decision time,  $s_t$  (Ratcliff & Tuerlinckx, 2002). Hence, the four key components of the diffusion model are (1) the speed of information processing, quantified by mean drift rate  $v$ ; (2) response caution, quantified by boundary separation  $a$ ; (3) a priori bias, quantified by starting point  $z$ ; and (4) mean non-decision time, quantified by  $T_{er}$ .

The diffusion model has been applied to a wide range of experimental paradigms, including perceptual discrimination, letter identification, lexical decision, recognition memory, and signal detection (e.g., Ratcliff, 1978; Ratcliff, Gomez, & McKoon, 2004; Ratcliff, Thapar, & McKoon, 2006; Klauer, Voss, Schmitz, & Teige-Mocigemba, 2007; Wagenmakers, Ratcliff, Gomez, & McKoon, 2008; van Ravenzwaaij, van der Maas, & Wagenmakers, 2011; Ratcliff, Thapar, & McKoon, 2010). Recently, the diffusion model has also been applied in clinical settings featuring sleep deprivation (Ratcliff & van Dongen, 2009), anxiety (White, Ratcliff, Vasey, & McKoon, 2010), and hypoglycemia (Geddes et al., 2010). The model has also been extensively applied in the neurosciences (Ratcliff, Hasegawa, Hasegawa, Smith, & Segraves, 2007; Philiastides, Ratcliff, & Sajda, 2006; Mulder, Wagenmakers, Ratcliff, Boekel, & Forstmann, 2012).

The advantages of a diffusion model analysis are twofold. First, the model takes into account entire RT distributions, both for correct and incorrect responses. This contrasts with a traditional analysis that considers only the mean RT for correct responses, and perhaps error rate, but ignores entirely the shape of the RT distributions and the speed of error responses. Second, the model allows researchers to decompose observed RTs and error rates into latent psychological processes such as processing speed and response caution. In the traditional analysis no attempt is made to explain the observed data by means of a psychologically plausible process model.

In their study on the effects of action video game playing on RT tasks, Green et al. (2010) report an increase in drift rate and a decrease in boundary separation for participants who practiced action games relative to participants who practiced cognitive games. Below we report two experiments that feature supervised video game play, repeated measurements of perceptual task performance, and a diffusion model decomposition of the data.

## Experiment 1

As a first test of the results of Green et al. (2010), we set out to compare NVGPs in an action game condition versus NVGPs in a cognitive game condition. Both groups were

<sup>2</sup>Mathematically, the change in evidence  $X$  is described by a stochastic differential equation  $dX(t) = \xi \cdot dt + s \cdot dW(t)$ , where  $s \cdot dW(t)$  represents the Wiener noise process with mean 0 and variance  $s^2 \cdot dt$ . Parameter  $s$  is a scaling parameter and is usually set to 0.1.

trained on video games for ten hours, divided equally over five separate sessions. Prior to every two hours of gaming and in a sixth and final behavioral session, participants performed a perceptual discrimination task. See Figure 2 for a graphical depiction of the events in all six sessions.

### *Participants*

Twenty students from the University of Amsterdam (18 women, 2 men), aged 18 to 25 (mean = 20.6, SD = 2.4), participated on six separate days in exchange for course credit or a monetary reward of 112 euros. Participants were screened for gaming experience. In order to qualify for participation, students could not play video games for more than two hours per week on average at the present and for no more than five hours per week on average during any time in their life. Participants were randomly assigned to either the “action” or the “cognitive” condition, under the restriction that each condition contain ten participants.

### *Materials*

*Video Games.* In the action condition, participants played *Unreal Tournament 2004*. This game is a so-called *first-person shooter*. The aim of the protagonist is to navigate a three-dimensional world using keyboard and mouse, shooting scores of enemies and avoiding to be shot himself. This game requires response speed, anticipation, and planning. In the cognitive condition, participants played *The Sims 2*. This game is a *strategy game*. The aim of the protagonist is to live a virtual life and manage its basic “requirements” such as fulfilling social obligations, obtaining food, and keeping a job. This game does not rely on response speed, or at least not to the extent that *Unreal Tournament 2004* does.

*Moving Dots Task.* Each of the six sessions featured a moving dots task (Ball & Sekuler, 1982; Newsome & Paré, 1988; Britten, Shadlen, Newsome, & Movshon, 1992) with two blocks of 200 trials each. On every trial, the stimulus consisted of 120 dots, 40 of which moved coherently and 80 of which moved randomly. After each 50-ms frame, the 40 coherently moving dots moved 1 pixel in the target direction. The other 80 dots were relocated randomly. On the subsequent frame, each dot might switch roles, with the constraint that there were always 40 dots moving coherently between a given set of frames. The moving dots stimulus gives the impression that the cloud of dots is systematically moving or turning in one direction, even though the cloud remains centered on the screen. Each dot consisted of 3 by 3 pixels, and the entire cloud of dots had a diameter of 250 pixels. Dots were randomly distributed over this pixel range. Participants indicated their response by pressing one of two buttons on an external device with their left or right index finger.

Immediately prior to each stimulus, a fixation cross was displayed for a random interval of 500, 800, 1000 or 1200 ms. Participants had 1500 ms to view the stimulus and give a response. The stimulus disappeared as soon as a response was made. If, for a given trial, the participant’s response was slower than 1000 ms, participants saw the message “te langzaam” (too slow) at the end of the trial. In contrast to Experiment 2, coherence level was not calibrated on an individual basis.

### Procedure

Upon entering, if it was the participant's first session, he or she received a general instruction about the procedure and signed an informed consent form. The participant then completed the moving dots task and a lexical decision task.<sup>3</sup> The moving dots task and the lexical decision task each took approximately 20 minutes to complete. Then, in all but the sixth and final session, the participant played the video game (action or cognitive, depending on the condition) for an hour. An experimenter was in the room while the participants played the action video games, had the computer screen in sight at all times, and was available for assistance in the unlikely event a participant got stuck (this rarely happened). Following the first hour of gaming, the participant had a break of 10 to 15 minutes, after which he or she played the video game for a second hour, completing the session. The first, second, third, fourth, and fifth session all took approximately three hours each. For the sixth session, the participant first completed the moving dots task and the lexical decision task, and then filled out a payment form. The sixth session took approximately one hour. In total, all sessions took approximately 16 hours per participant, 10 hours of which were spent gaming.

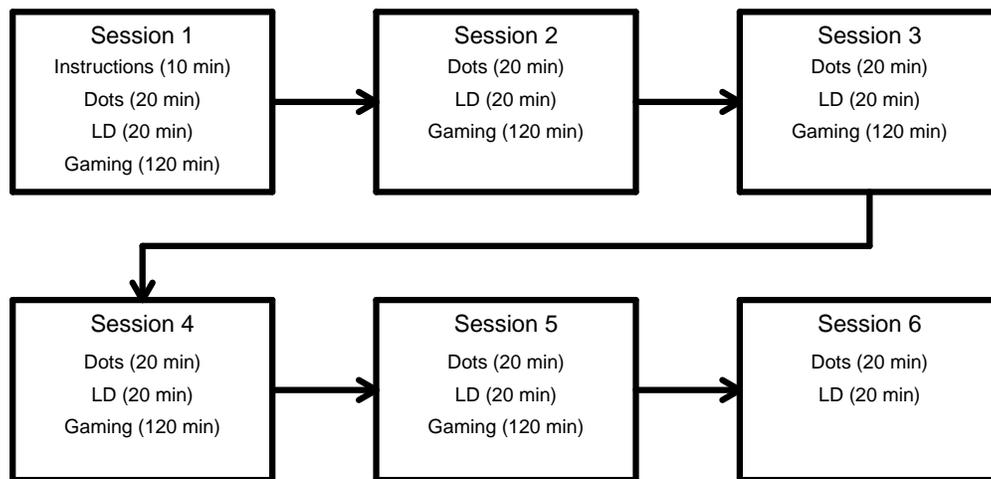


Figure 2. A flowchart of the design for Experiment 1. Dots = Moving Dots Task, LD = Lexical Decision Task. Supervised sessions took place on different days, spanning at most seven days.

<sup>3</sup>For consistency with Experiment 2, only the results for the moving dots task are reported. The results for the lexical decision task may be found in the online appendix, available at <http://www.donvanravenzwaaij.com/Papers.html>. Compared to the cognitive game, playing the action video game did not lead to improved performance on the lexical decision task.

## Results

The next three subsections present the behavioral results (mean RT and accuracy), the diffusion modeling decomposition, and the diffusion model fit.<sup>4</sup> For the remainder of the statistical analyses, we report not only conventional  $p$ -values but also Bayes factors (e.g., Jeffreys, 1961; Kass & Raftery, 1995; Hoijtink, Klugkist, & Boelen, 2008). Bayes factors represent “the primary tool used in Bayesian inference for hypothesis testing and model selection” (Berger, 2006, p. 378); in contrast to  $p$ -values, Bayes factors allow researchers to quantify evidence in favor of the null hypothesis vis-à-vis the alternative hypothesis. For instance, when the Bayes factor  $BF_{01} = 10$  the observed data are 10 times more likely to have occurred under  $H_0$  than under  $H_1$ . When  $BF_{01} = 1/5 = 0.20$  the observed data are 5 times more likely to have occurred under  $H_1$  than under  $H_0$ . In the following, Bayes factors for analysis of variance are based on the BIC approximation (e.g., Wagenmakers, 2007; Masson, 2011), and Bayes factors for  $t$ -tests are based on the default Bayesian  $t$ -test proposed by Rouder, Speckman, Sun, Morey, and Iverson (2009).

### *Behavioral Results*

One participant withdrew from the experiment after the first session and was replaced. For each participant, we excluded all RTs below 275 ms, as these were likely to be guesses. This led to the exclusion of 0.1% of all RTs and did not affect the results.

Figure 3 shows the within-subject effects for mean RT and accuracy. Across conditions, participants’ mean RTs shortened in subsequent sessions, as confirmed by the presence of a negative linear trend over sessions on mean RT ( $F(1, 96) = 53.0$ ,  $p < .001$ ,  $BF_{01} = 2.8 \cdot 10^{-9}$ ). Thus, practice on the moving dots task resulted in faster responding. Importantly, this session effect for mean RT did not interact with gaming condition ( $F(1, 96) = 0.26$ ,  $p > .05$ ,  $BF_{01} = 8.75$ ). From the first to the last session, the overall speedup in mean RT was 51 milliseconds for the action condition and 59 milliseconds for the cognitive condition.

In addition to speeding up, participants also made more mistakes in subsequent sessions across conditions (linear trend:  $F(1, 96) = 6.6$ ,  $p < .05$ ,  $BF_{01} = 0.37$ ). There was no evidence for an interaction between session and gaming condition for accuracy ( $F(1, 96) = 1.8$ ,  $p > .05$ ,  $BF_{01} = 3.86$ ).

In sum, practice on the moving dots task decreased mean RT for both the action and the cognitive condition. Playing the action video game did not result in better performance compared to playing the cognitive video game. Response accuracy decreased slightly over sessions, hinting at the possibility that participants became less cautious as they improved with practice (see also Dutilh, Wagenmakers, Vandekerckhove, & Tuerlinckx, 2009). In order to quantify the psychological factors that drive the observed effects we now turn to a diffusion model decomposition.

### *Diffusion Model Decomposition*

The diffusion model was fit to the data using the DMAT software package (Vandekerckhove & Tuerlinckx, 2007), which minimizes a negative multinomial log-likelihood function. Each participant was fit separately. We fixed starting point  $z$  to be half

<sup>4</sup>Data from both experiments is available at <http://www.donvanravenzwaaij.com/Papers.html>.

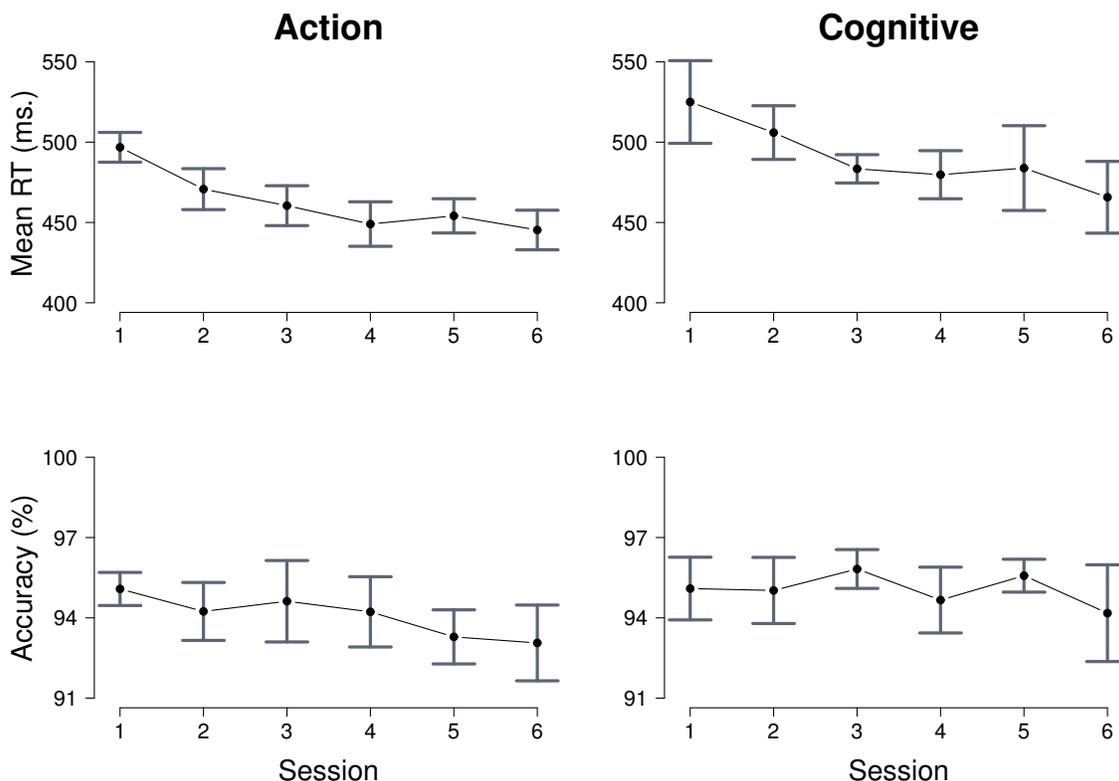


Figure 3. The within-subject effects of the action condition (top panels) and the cognitive condition (bottom panels) on mean RT (left panels) and response accuracy (right panels) for the moving dots task from Experiment 1. Error bars represent 95% confidence intervals.

of boundary separation  $a$ , as there was no reason to expect a bias for either the left or right direction in the moving dots stimuli. We estimated a separate mean drift rate  $v$ , boundary separation  $a$ , and non-decision time  $T_{er}$  for each session. Furthermore, we constrained the standard deviation of drift rate  $\eta$ , range of starting point  $s_z$ , and range of non-decision time  $s_t$  to be equal across sessions.

The diffusion model captured the error rates okay for the cognitive condition and somewhat poorly for the action condition. The RTs were captured well on average.<sup>5</sup> Figure 4 shows the within-subject effects for drift rate  $v$ , boundary separation  $a$ , and non-decision time  $T_{er}$ . Across conditions, participants processed information faster in subsequent sessions, but the effect leveled off for later sessions; this visual impression is confirmed by the presence of a positive linear trend over sessions for drift rate  $v$  ( $F(1, 96) = 24.7$ ,  $p < .001$ ,  $BF_{01} = 1.1 \cdot 10^{-4}$ ). From the first to the last session, the overall training effect on drift rate  $v$  was 0.13 for the action condition and 0.23 for the cognitive condition. Importantly, there was no interaction between session and gaming condition for drift rate ( $F(1, 96) = 0.77$ ,  $p > .05$ ,  $BF_{01} = 6.71$ ).

For boundary separation  $a$ , there was no evidence for a linear trend over sessions

<sup>5</sup>For details, see Figure 3 of the online appendix.

across conditions ( $F(1, 96) = 3.5, p > .05, BF_{01} = 1.71$ ) and there was no evidence for an interaction between session and gaming condition ( $F(1, 96) = 0.03, p > .05, BF_{01} = 9.82$ ). For non-decision time  $T_{er}$  there was also no evidence for the presence of a linear trend over sessions across conditions ( $F(1, 96) = 0.64, p > .05, BF_{01} = 7.16$ ) and no evidence for an interaction between session and gaming condition ( $F(1, 96) = 0.09, p > .05, BF_{01} = 9.59$ ).

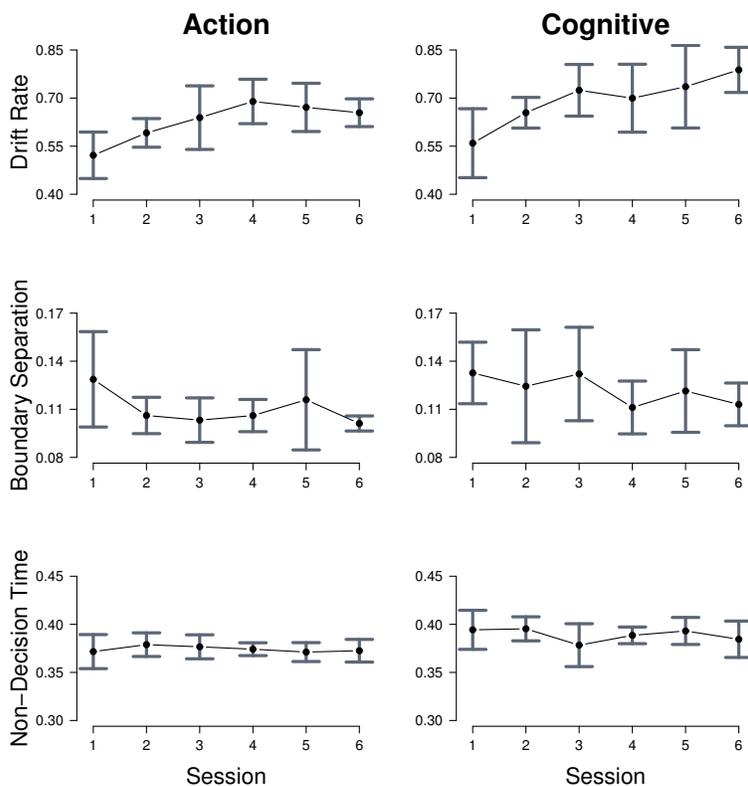


Figure 4. The within-subject effects of the action condition (top panels) and the cognitive condition (bottom panels) on drift rate  $v$  (left panels), boundary separation  $a$  (middle panels), and non-decision time  $T_{er}$  (right panels) for the moving dots task from Experiment 1. Error bars represent 95% confidence intervals.

In sum, practice on the moving dots task increased the rate of information processing. Response caution and non-decision time were unaffected. The practice-induced increase in the rate of information processing was unaffected by the type of game played. Hence, contrary to the results by Green et al. (2010), playing action video games did not yield an increased benefit on information processing.<sup>6</sup>

<sup>6</sup>The conclusions from the diffusion model parameters can only be relied upon, however, when the model provides a satisfactory fit to the data. In order to reassure the reader that the diffusion model gives a good description of the data, we present *model predictives* in the online appendix.

### Interim Conclusion

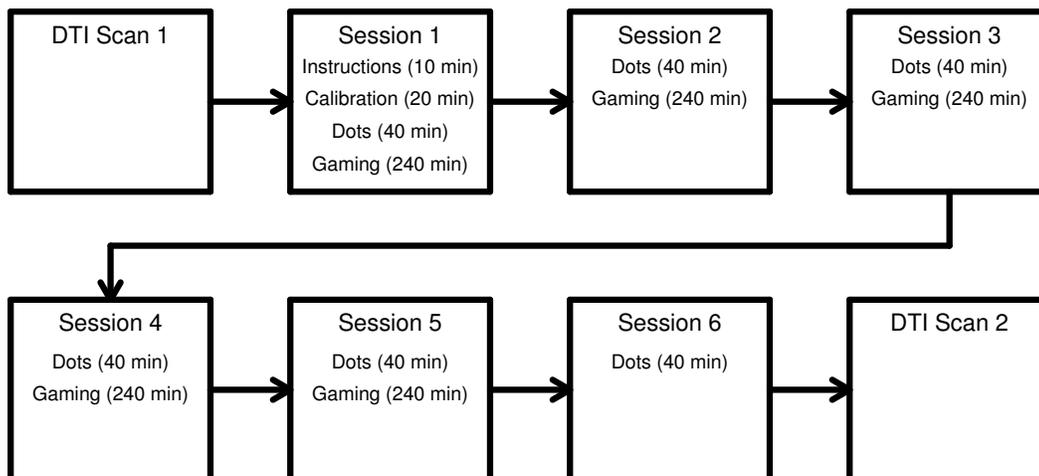
Experiment 1 showed no benefit of action video game playing, neither in the behavioral data nor in the diffusion model parameters. Hence, the results from Experiment 1 are at odds with the findings from Green et al. (2010). However, one may argue that we failed to find the effect because ten hours of game-play training are insufficient to elicit a reliable effect. Of course, earlier training studies also used ten hours of game-play (e.g., Green & Bavelier, 2003; Feng et al., 2007, see Powers et al., in press for a review). Moreover, Experiment 1 showed not even a hint of an effect, making its hypothetical appearance after additional training hours somewhat implausible. Nevertheless, we decided to conduct a second experiment in which we doubled the number of hours spent gaming, for both the action and the cognitive game condition. To verify that participants actually improved in the action video game, we monitored their skill level. As an additional safeguard, we also included a no-gaming condition which served as a baseline for both the action and the cognitive game condition. We also calibrated the moving dots task to each individual to produce response accuracies that were not at ceiling. In addition, we tested 45 participants instead of 20 and we increased the number of moving dots trials in each session from 400 in Experiment 1 to 1000 in Experiment 2. Finally, in an attempt to reduce the impact of potential confounds due to differential expectations (e.g., Boot et al., 2011), we told participants a believable cover story as to the goal of the experiment.

### Experiment 2

As a second test of the results of Green et al. (2010), we set out to compare NVGPs in an action game condition to NVGPs in a cognitive game condition and to NVGPs in a no-gaming condition. The action and cognitive groups were trained on video games for twenty hours, divided equally over five separate sessions. Prior to every four hours of gaming and in a sixth and final behavioral session, participants performed a perceptual discrimination task. In the first session, we calibrated the coherence level of the moving dots task to obtain comparable and off-ceiling response accuracies for each participant. On two separate days, one prior to the first session and one after the final session, participants underwent a diffusion tensor imaging scan. Both of these scans were compared to examine improvements with practice on the moving dots task. The results of these scans are unrelated to the video gaming and will be reported elsewhere. See Figure 5 for a graphical depiction of the events in all six sessions.

#### *Participants*

Forty-five students from the University of Amsterdam (19 women, 24 men), aged 17 to 24 (mean = 20, SD = 1.8), participated on six separate days in exchange for course credit and a monetary reward of 63 euros. Participants were screened for gaming experience. In order to qualify for participation, students could not play video games for more than one hour per week on average at the present and for no more than ten hours per week on average during any time in their life. Participants were randomly assigned to either the “action”, the “cognitive”, or the “control” condition, under the restriction that each condition contained fifteen participants. As pointed out by Boot et al. (2011), participants in the action condition may expect to improve on the experimental tasks, whereas participants



*Figure 5.* A flowchart of the design for Experiment 2. Calibration = Practice and calibration block for the moving dots task, Dots = Moving dots task. Supervised sessions took place on different days, spanning exactly seven days. In the control condition, participants terminated their session upon completion of the moving dots task.

in the cognitive condition may not. This confound has the potential to create spurious benefits from action video-gaming, and in order to attenuate its influence we (falsely) informed participants that they participated in an experiment that examined the effect of a perceptual task on their performance in video game playing.

### *Materials*

*Video Games.* See the section “Video Games” in Experiment 1 for details.

*Moving Dots Task.* For Experiment 2 we modified the moving dots task used in Experiment 1. Specifically, we calibrated task difficulty (i.e., coherence level) and increased the number of trials. For calibration purposes, each participant performed a practice block of 400 trials with stimuli of varying difficulty (i.e., 0%, 10%, 20%, 40%, and 80% coherence, for 80 trials each in a randomly interleaved order). The mean RTs and accuracy data from this practice block were then fit with the Palmer diffusion model where drift rate is constrained to be proportional to the coherence level (Palmer, Huk, & Shadlen, 2005). The psychometric curve predicted by the Palmer diffusion model was then used to determine for each participant the coherence level that corresponds to 75% accuracy. This coherence level was then fixed throughout all of the experimental blocks.

Participants had 2000 ms to view the stimulus and give a response. The stimulus disappeared as soon as a response was made. If, for a given trial, the participant’s response was slower than 2000 ms, participants saw the message “No Response” at the end of the trial. If the participant’s response was faster than 200 ms, participants saw the message

“Too Fast”. The fixation cross was present on screen at all times. The moving dots task took approximately 40 minutes to complete.

### *Procedure*

The procedure in Experiment 2 differed from that of Experiment 1 in a number of ways. Firstly, the initial session commenced with a 400-trial moving dots practice block that was used for the individual calibration of task difficulty. After completing this practice block participants had a five minute break, during which the experimenter set up the main task. Secondly, the moving dots task used in sessions one through six consisted of 1000 trials during which participants had three self-paced breaks after 250, 500, and 750 trials. Thirdly, in sessions one through five, participants played the video game for four hours instead of two (except in the control condition in which each session ended as soon as the participant completed the moving dots task). Participants had three self-paced breaks after each hour of gaming. In the action condition, the difficulty level was adapted to the ability of the participant.<sup>7</sup> Finally, Experiment 2 did not feature a lexical decision task.

## Results

There were no significant gender main effects or interactions with condition for mean RT or accuracy. In what follows, we have collapsed across gender in our presentation of the results. Figure 6 shows the improvement in difficulty level over consecutive games, averaged over participants. Note that levels range from 1, “Novice”, to 8, “Godlike”. The figure shows substantial training effects for the first 10 hours (up to approximately game number 50, depending on the individual), and a substantially reduced training effect on the action video game in the second 10 hours of gaming. This law of diminished returns is characteristic of virtually all human learning (e.g., Newell & Rosenbloom, 1981).

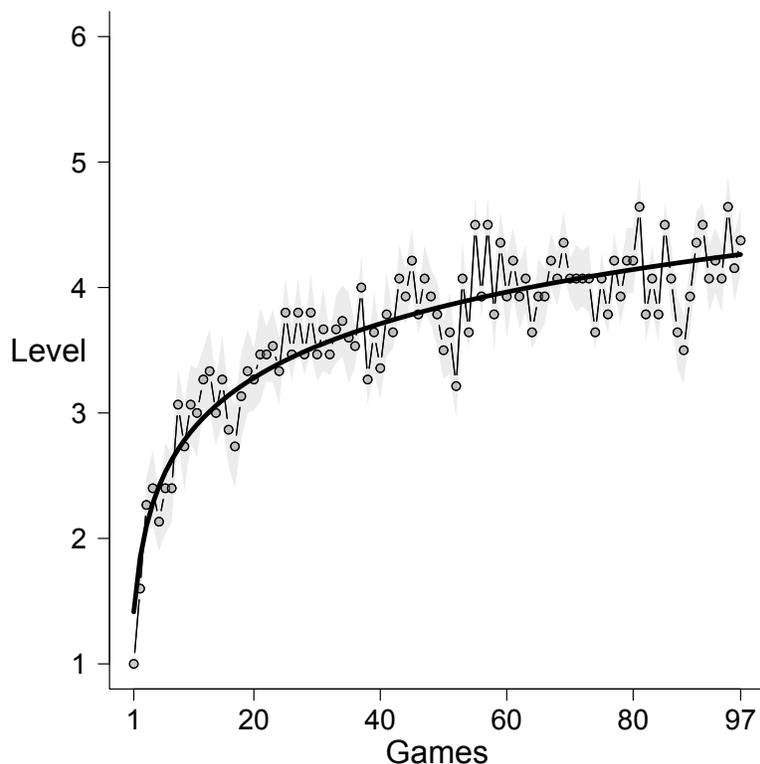
The next three subsections present the behavioral results (mean RT and accuracy), the diffusion modeling decomposition, and the diffusion model fit.

### *Behavioral Results*

One participant from the action condition and one participant from the cognitive condition withdrew their participation. For each participant, we excluded all RTs below 275 ms, as these were likely to be guesses. This led to the exclusion of 0.7% of all RTs and did not affect the results.

Figure 7 shows the within-subject effects for mean RT and accuracy. Across conditions, participants’ mean RTs shortened in subsequent sessions, as confirmed by the presence of a negative linear trend over sessions on mean RT ( $F(1, 209) = 60.5$ ,  $p < .001$ ,  $BF_{01} = 2.0 \cdot 10^{-11}$ ). Thus, practice on the moving dots task resulted in faster responding. Interestingly, this session effect for mean RT interacted with gaming condition ( $F(2, 209) = 6.8$ ,  $p < .01$ ,  $BF_{01} = 0.24$ ), but not in the expected direction: as the number of sessions increases, participants speeded up both in the cognitive game condition ( $t(13) = 5.0$ ,  $p < .001$ , 95% c.i. = [105.4 – 266.8],  $BF_{01} = 0.01$ ) and in the no-game condition ( $t(14)$

<sup>7</sup>Participants continuously played in “Deathmatch” mode. If the participant had more kills than any of the AI opponents after 10 minutes of play, the experimenter increased the difficulty level by one, and if the participant did not, the experimenter decreased the difficulty level by one.



*Figure 6.* Average skill level of participants over consecutive games (10 minutes per game). The dark line represents the best fitting linear model through the log-transformed skill level data, with the grey area representing the standard error of the mean data points. Most participants were unable to play more than 97 games, given the time required for starting up new games, bathroom breaks, etc.

= 3.9,  $p < .01$ , 95% c.i. = [52.1 – 176.5],  $BF_{01} = 0.04$ ) whereas they did not speed up in the action game condition ( $t(13) = 1.4$ ,  $p > .05$ , 95% c.i. = [-17.9 – 79.2],  $BF_{01} = 2.11$ ). From the first to the last session, the overall session effect on mean RT was 31 milliseconds for the action condition, 186 milliseconds for the cognitive condition, and 114 milliseconds for the no-game condition.

In contrast to Experiment 1, participants in Experiment 2 became more accurate in subsequent sessions across conditions; there was a significant positive linear trend over sessions for accuracy ( $F(1, 209) = 179.3$ ,  $p < .001$ ,  $BF_{01} = 1.8 \cdot 10^{-28}$ ). This discrepancy is most likely due to the individual calibration of task difficulty. There was no evidence for an interaction between session and gaming condition for accuracy ( $F(2, 209) = 0.97$ ,  $p > .05$ ,  $BF_{01} = 130.3$ ) – in fact, the Bayes factor suggests that there is strong evidence in favor of the absence of such an interaction.

In sum, practice on the moving dots task led to faster performance for the no-game condition and the cognitive game condition, but not for the action game condition. This result is at odds with that reported by Green et al. (2010). In order to quantify the

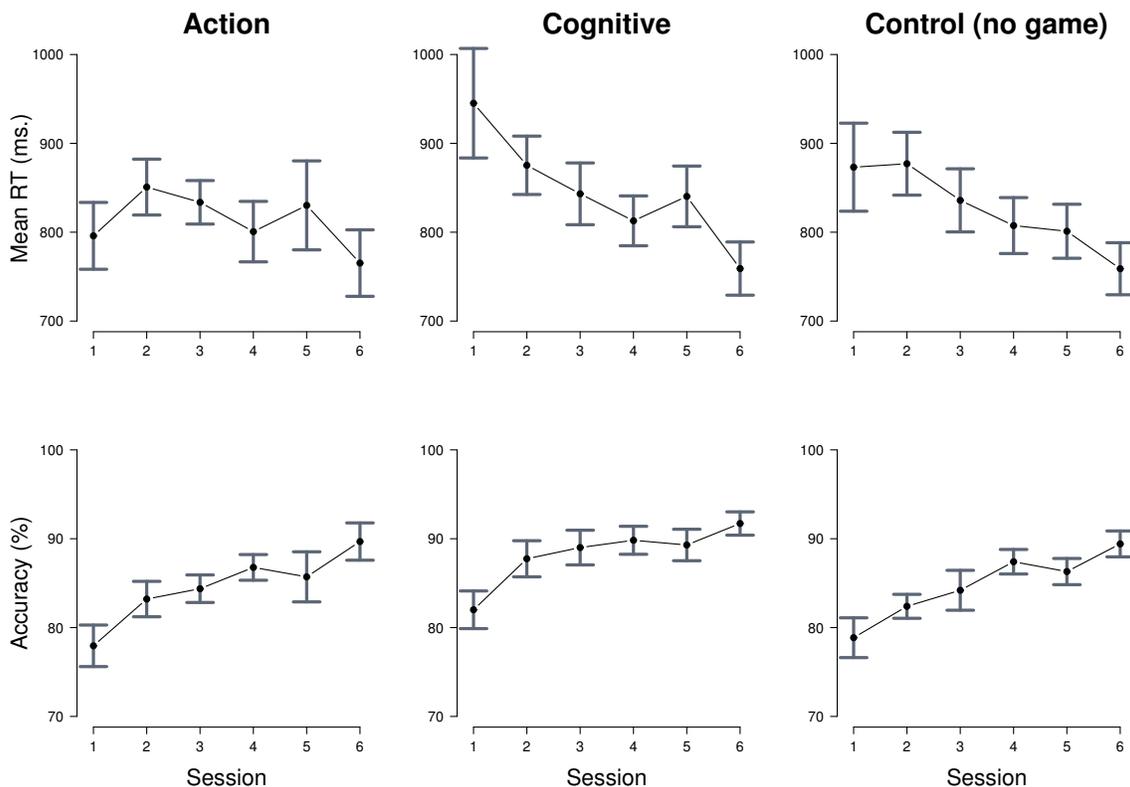


Figure 7. The within-subject effects of the action condition (left panels), the cognitive condition (middle panels), and the no-game conditions (right panels) on mean RT (top panels) and response accuracy (bottom panels) for the moving dots task from Experiment 2. Error bars represent 95% confidence intervals.

psychological factors that drive the observed effects we now turn to a diffusion model decomposition.

#### Diffusion Model Decomposition

The diffusion model analyses followed the outline provided under Experiment 1 above. The diffusion model captured the data well.<sup>8</sup> Figure 8 shows the within-subject effects for drift rate  $v$ , boundary separation  $a$ , and non-decision time  $T_{er}$ . Across conditions, participants processed information faster in subsequent sessions; this visual impression is confirmed by the presence of a positive linear trend over sessions for drift rate  $v$  ( $F(1, 209) = 201.1, p < .001, BF_{01} = 5.0 \cdot 10^{-31}$ ).<sup>9</sup> For all conditions the increase in drift rate was about 0.1. Importantly, there was no interaction between session and gaming condition for drift rate ( $F(2, 209) = 2.2, p > .05, BF_{01} = 23.5$ ).

For boundary separation  $a$ , there was no evidence for a linear trend over sessions across

<sup>8</sup>For the detailed model predictives, see Figure 4 of the online appendix.

<sup>9</sup>Note that the absolute size of the drift rates is lower than for Experiment 1, a phenomenon that reflects that the moving dots paradigm for Experiment 2 was more difficult in order to prevent ceiling effects.

conditions ( $F(1, 209) = 2.3$ ,  $p > .05$ ,  $BF_{01} = 4.50$ ). There was a significant interaction between session and game condition ( $F(2, 209) = 9.3$ ,  $p < .001$ ,  $BF_{01} = 0.02$ ); boundary separation  $a$  was increasing over sessions in the action condition ( $t(13) = 2.4$ ,  $p < .05$ , 95% c.i. =  $[0.002 - 0.028]$ ,  $BF_{01} = 0.51$ ), whereas boundary separation did not change in the cognitive condition ( $t(13) = 2.0$ ,  $p > .05$ , 95% c.i. =  $[-0.001 - 0.030]$ ,  $BF_{01} = 0.95$ ) nor in the no-game condition ( $t(14) = -1.8$ ,  $p > .05$ , 95% c.i. =  $[-0.039 - 0.003]$ ,  $BF_{01} = 1.21$ ). In other words, participants became more cautious in subsequent sessions for the action condition. Note that all Bayes factors for these specific comparisons suggest the evidence is ambiguous.

For non-decision time  $T_{er}$  there was no linear trend over sessions across conditions ( $F(1, 209) = 0.28$ ,  $p > .05$ ,  $BF_{01} = 12.76$ ) and no evidence for an interaction between session and gaming condition ( $F(2, 209) = 1.8$ ,  $p > .05$ ,  $BF_{01} = 33.95$ ).

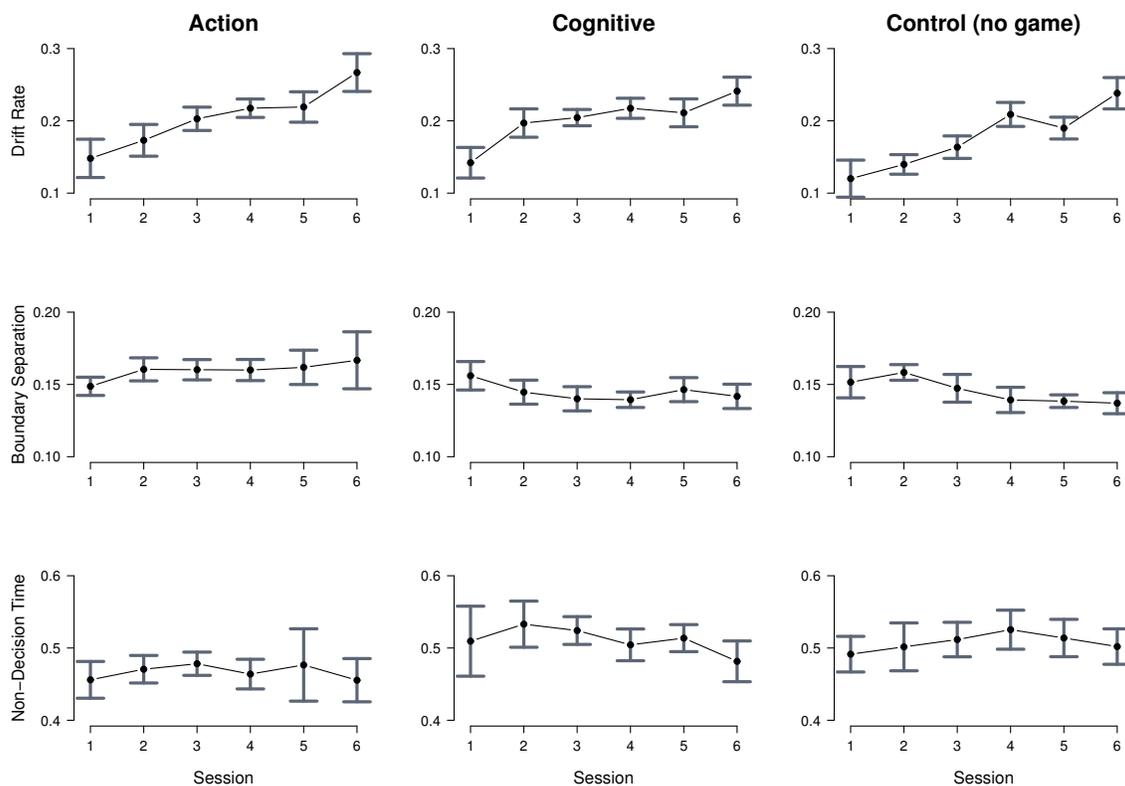


Figure 8. The within-subject effects of the action condition (left panels), the cognitive condition (middle panels), and the no-game conditions (right panels) on drift rate  $v$  (top panels), boundary separation  $a$  (middle panels), and non-decision time  $T_{er}$  (bottom panels) for the moving dots task from Experiment 2. Error bars represent 95% confidence intervals.

In sum, practicing the moving dots task led to benefits in terms of mean RT for the no-game and cognitive game conditions, but not for the action game conditions. In addition, practicing the moving dots task led to an increase in response accuracy for all three conditions. When viewed through the lens of the diffusion model, it became clear that

these practice effects were caused by an increase in the speed of information processing.

The benefits of practice were no greater for participants playing action video games than for participants playing cognitive games or for participants who did not play video games at all, a statement that holds true both for the behavioral measures and for the diffusion model drift rate parameters.

### Conclusion

Two training experiments showed that performance on a perceptual discrimination task improves with practice. This is hardly surprising. More noteworthy is that we failed to find any performance benefit for participants who played an action video game compared to participants who played a cognitive game or no game at all. Neither the behavioral data nor a diffusion model analysis revealed even a trace of a performance increment due to playing action video games. The Bayesian analyses supported the null hypothesis of equal performance for all gaming conditions.

So why do we fail to find a gaming-specific benefit on the moving dots task? Was the amount of hours playing video games insufficient? Recall that our participants played for 10 hours in Experiment 1 and for 20 hours in Experiment 2. The training study of Green et al. (2010), however, employed 50 hours of video game training; perhaps the difference is due to these last 30 hours?

There are multiple reasons why this alternative explanation is unlikely. First and foremost, a recent meta-analysis by Powers et al. (in press) concluded: “In true experiments, effect sizes were comparable across studies utilizing varying amounts of training (under 10 h vs. over 10 h), which suggests that learners quickly adapt cognitive processes to the design features of specific games, and may not need extensive practice to accrue training benefits.” Second, as underscored by Powers et al. (in press), the number of training hours in our two experiments is average and high, respectively. Third, the critique is valid only when the effect is of a rather special nature, namely: (1) it does not manifest itself at all during the first 20 hours, and then develops quite strongly over the next 30 hours; and (2) it is somehow different from other effects of gaming, because, in general, “length of training (...) failed to moderate the effects” (Powers et al., in press).

It may be argued that the repeated testing on the moving dots task produced such a strong practice effect that it swamped any improvements due to the gaming itself. We believe this interpretation of our results to be implausible. There are no indications of a ceiling effect in the data from our second experiment. Testing six times instead of two greatly increases the power to detect a beneficial gaming effect, should it exist (see section “Power Analysis Experiment 2” of the online appendix for a demonstration). On top of that, if the power of our design had been too low, the Bayesian analyses would have indicated ambivalent evidence. However, we found clear evidence in favor of the null hypothesis.

Nevertheless, one could argue that the gaming transfer effects are so fragile and specific that they are only present when the target task has received little practice. For instance, assume that test-retest benefits are of two kinds, (a) improvement in visuo-perceptual processing (unfolding over a long time scale); and (b) improvement in pressing response buttons (unfolding over a very short time scale). It could possibly be the case that the transfer effect from gaming influences only the peripheral process (b), and that process (b) is at ceiling relatively quickly.

Of course, this account is speculative, contradicts prior theorizing in which the benefit is thought to represent general facilitation of cognitive processing, and raises the question of the relevance of the transfer effect in practical application. So even if this alternative account is true, it would render the transfer effect of gaming rather uninteresting, as it may represent a temporary adjustment of a peripheral response process rather than a change in cognitive processing.

Our findings contradict the claim that action–video game playing selectively improves cognitive processing as argued by Green et al. (2010). It is possible that overt recruiting, unspecified recruiting methods, or the lack of supervised game–play has caused the results of Green et al. (2010) to be different from ours see e.g., Boot et al. (2011). However, our findings are consistent with the outcome of the recent meta–analysis by Powers et al. (in press), who concluded that “In true experiments, action/violent game training was no more effective than game training utilizing nonaction or puzzle games, but mimetic games showed large effects.” Powers et al. (in press) did conclude that there is a reliable transfer effect of gaming in general, and this is inconsistent with our result from Experiment 2 – our control group improved just as much as the two groups of gamers.

In order to resolve the contradictory findings it may be essential to engage in adversarial collaborations or at least state the analysis plan in advance of data collection (Chambers, 2013; Wagenmakers, Wetzels, Borsboom, van der Maas, & Kievit, 2012; Wolfe, 2013). Furthermore, we suggest that researchers who plan to study the effects of video game playing consider the following guidelines:

1. Despite the considerable effort required, conduct a training study on non–gamers rather than compare gamers to non–gamers. Training studies eliminate confounds due to pre-existing differences between gamers and non–gamers.

2. Design experiments with sufficient power. By testing many participants and including many trials for the psychological task at hand you increase the chances of confidently confirming or disconfirming the effects of interest.

3. Calibrate your psychological task on an individual basis to ensure error rates that are sizable and similar across participants. Training effects on response accuracy are easier to detect when accuracy is not near ceiling. A sufficient number of errors also benefits RT modeling with the diffusion model (e.g., compare the model predictives for Experiment 1 and Experiment 2).

4. Use Bayes factors to quantify evidence. Only by using Bayes factors can researchers quantify the evidence favoring the null hypothesis.

5. Use the diffusion model to decompose the behavioral effects in their underlying psychological processes such as speed of information processing, response caution, and non–decision time.

We find the claim enticing that people can boost their cognitive capacities by playing violent action video games. However, our result urge caution and suggest that before the video games find application, a series of purely confirmatory experiments is in order. We hope that our experiments will encourage such confirmatory work.

## References

- Ball, K., & Sekuler, R. (1982). A specific and enduring improvement in visual motion discrimination. *Science*, *218*, 697–698.
- Beck, J. M., Ma, W. J., Kiani, R., Hanks, T., Churchland, A. K., Roitman, J. D., et al. (2008). Probabilistic population codes for Bayesian decision making. *Neuron*, *60*, 1142–1152.
- Berger, J. O. (2006). Bayes factors. In S. Kotz, N. Balakrishnan, C. Read, B. Vidakovic, & N. L. Johnson (Eds.), *Encyclopedia of statistical sciences*, vol. 1 (2nd ed.) (pp. 378–386). Hoboken, NJ: Wiley.
- Boot, W. R., Blakely, D. P., & Simons, D. J. (2011). Do action video games improve perception and cognition? *Frontiers in Psychology*, *2*, 1–6.
- Boot, W. R., Kramer, A. F., Simons, D. J., Fabiani, M., & Gratton, G. (2008). The effects of video game playing on attention, memory, and executive control. *Acta Psychologica*, *129*, 387–398.
- Britten, K. H., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (1992). The analysis of visual motion: A comparison of neuronal and psychophysical performance. *Journal of Neuroscience*, *12*, 4745–4765.
- Chambers, C. D. (2013). Registered Reports: A new publishing initiative at Cortex. *Cortex*, *49*, 609–610.
- Clark, J. E., Lanphear, A. K., & Riddick, C. C. (1987). The effects of videogame playing on the response selection processing of elderly adults. *Journal of Gerontology*, *42*, 82–85.
- Drew, D., & Waters, J. (1986). Video games: Utilization of a novel strategy to improve perceptual motor skills and cognitive functioning in the non-institutionalized elderly. *Cognitive Rehabilitation*, *4*, 26–31.
- Dutilh, G., Wagenmakers, E.-J., Vandekerckhove, J., & Tuerlinckx, F. (2009). A diffusion model decomposition of the practice effect. *Psychonomic Bulletin & Review*, *16*, 1026–1036.
- Fahle, M. (2005). Perceptual learning: Specificity versus generalization. *Current Opinion in Neurobiology*, *15*, 154–160.
- Feng, J., Spence, I., & Pratt, J. (2007). Playing an action video game reduces gender differences in spatial cognition. *Psychological Science*, *18*, 850–855.
- Geddes, J., Ratcliff, R., Allerhand, M., Childers, R., Wright, R. J., Frier, B. M., et al. (2010). Modeling the effects of hypoglycemia on a two-choice task in adult humans. *Neuropsychology*, *24*, 652–660.
- Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature*, *423*, 534–537.
- Green, C. S., & Bavelier, D. (2006). Enumeration versus multiple object tracking: The case of action video game players. *Cognition*, *101*, 217–245.
- Green, C. S., & Bavelier, D. (2012). Learning, attentional control, and action video games. *Current Biology*, *22*, R197–R206.
- Green, C. S., Pouget, A., & Bavelier, D. (2010). Improved probabilistic inference as a general learning mechanism with action video games. *Current Biology*, *20*, 1573–1579.
- Hojtink, H., Klugkist, I., & Boelen, P. (2008). *Bayesian evaluation of informative hypotheses that are of practical value for social scientists*. New York: Springer.
- 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*, 224–231.
- Jeffreys, H. (1961). *Theory of probability*. Oxford, UK: Oxford University Press.
- Jeter, P. E., Doshier, B. A., Petrov, A., & Lu, Z.-L. (2009). Task precision at transfer determines specificity of perceptual learning. *Journal of Vision*, *9*, 1–13.
- Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*, 773–795.
- Klauer, K. C., Voss, A., Schmitz, F., & Teige-Mocigemba, S. (2007). Process components of

- the implicit association test: A diffusion-model analysis. *Journal of Personality and Social Psychology*, *93*, 353–368.
- Li, R., Polat, U., Makous, W., & Bavelier, D. (2009). Enhancing the contrast sensitivity function through action video game training. *Nature Neuroscience*, *12*, 549–551.
- Liu, Z., & Weinshall, D. (2000). Mechanisms of generalization in perceptual learning. *Vision Research*, *40*, 97–109.
- Luce, R. D. (1986). *Response times*. New York: Oxford University Press.
- Masson, M. E. J. (2011). A tutorial on a practical Bayesian alternative to null-hypothesis significance testing. *Behavior Research Methods*, *43*, 679–690.
- Mulder, M. J., Wagenmakers, E.-J., Ratcliff, R., Boekel, W., & Forstmann, B. U. (2012). Bias in the brain: A diffusion model analysis of prior probability and potential payoff. *Journal of Neuroscience*, *32*, 2335–2343.
- Murphy, K., & Spencer, A. (2009). Playing video games does not make for better visual attention skills. *Journal of Articles Supporting the Null Hypothesis*, *6*, 1–20.
- Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1–55). Hillsdale, NJ: Erlbaum.
- Newsome, W. T., & Paré, E. B. (1988). A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *The Journal of Neuroscience*, *8*, 2201–2211.
- Palmer, J., Huk, A. C., & Shadlen, M. N. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision. *Journal of Vision*, *5*, 376–404.
- Philiastides, M. G., Ratcliff, R., & Sajda, P. (2006). Neural representation of task difficulty and decision-making during perceptual categorization: A timing diagram. *Journal of Neuroscience*, *26*, 8965–8975.
- Powers, K. L., Brooks, P. J., Aldrich, N. J., Palladino, M. A., & Alfieri, L. (in press). Effects of video-game play on information processing: A meta-analytic investigation. *Psychonomic Bulletin & Review*.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59–108.
- Ratcliff, R., Gomez, P., & McKoon, G. (2004). Diffusion model account of lexical decision. *Psychological Review*, *111*, 159–182.
- Ratcliff, R., Hasegawa, Y. T., Hasegawa, Y. P., Smith, P. L., & Segraves, M. A. (2007). Dual diffusion model for single-cell recording data from the superior colliculus in a brightness-discrimination task. *Journal of Neurophysiology*, *97*, 1756–1774.
- Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, *20*, 873–922.
- Ratcliff, R., Thapar, A., & McKoon, G. (2006). Aging, practice, and perceptual tasks: A diffusion model analysis. *Psychology and Aging*, *21*, 353–371.
- Ratcliff, R., Thapar, A., & McKoon, G. (2010). Individual differences, aging, and IQ in two-choice tasks. *Cognitive Psychology*, *60*, 127–157.
- Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review*, *9*, 438–481.
- Ratcliff, R., & van Dongen, H. P. A. (2009). Sleep deprivation affects multiple distinct cognitive processes. *Psychonomic Bulletin & Review*, *16*, 742–751.
- Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t-tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*, 225–237.
- Schlickum, M. K., Hedman, L., Enochsson, L., Kjellin, A., & Felländer-Tsai, L. (2009). Systematic video game training in Surgical novices improves performance in virtual reality endoscopic surgical simulators: A prospective randomized study. *World Journal of Surgery*, *33*, 2360–2367.

- van Ravenzwaaij, D., & Oberauer, K. (2009). How to use the diffusion model: Parameter recovery of three methods: EZ, fast-dm, and DMAT. *Journal of Mathematical Psychology*, *53*, 463–473.
- van Ravenzwaaij, D., van der Maas, H. L. J., & Wagenmakers, E.-J. (2011). Does the Name–Race Implicit Association Test Measure Racial Prejudice? *Experimental Psychology*, *58*, 271–277.
- Vandekerckhove, J., & Tuerlinckx, F. (2007). Fitting the Ratcliff diffusion model to experimental data. *Psychonomic Bulletin & Review*, *14*, 1011–1026.
- Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of *p*-values. *Psychonomic Bulletin & Review*, *14*, 779–804.
- Wagenmakers, E.-J. (2009). Methodological and empirical developments for the Ratcliff diffusion model of response times and accuracy. *European Journal of Cognitive Psychology*, *21*, 641–671.
- Wagenmakers, E.-J., Ratcliff, R., Gomez, P., & McKoon, G. (2008). A diffusion model account of criterion shifts in the lexical decision task. *Journal of Memory and Language*, *58*, 140–159.
- Wagenmakers, E.-J., Wetzels, R., Borsboom, D., van der Maas, H. L. J., & Kievit, R. A. (2012). An agenda for purely confirmatory research. *Perspectives on Psychological Science*, *7*, 627–633.
- White, C. N., Ratcliff, R., Vasey, M. W., & McKoon, G. (2010). Using diffusion models to understand clinical disorders. *Journal of Mathematical Psychology*, *54*, 39–52.
- Wolfe, J. M. (2013). Registered reports and replications in Attention, Perception, & Psychophysics. *Attention, Perception, & Psychophysics*, *75*, 781–783.