The Dissertation Committee for Brian Daniel Glass Certifies that this is the approved version of the following dissertation:

Becoming a Gamer: Cognitive Effects of Real-Time Strategy Gaming

Committee:

W. Todd Maddox, Co-Supervisor

Bradley C. Love, Co-Supervisor

Alexander C. Huk

Risto Miikkulainen

David M. Schnyer

Becoming a Gamer: Cognitive Effects of Real-Time Strategy Gaming

by

Brian Daniel Glass, B. S.; M. A.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2012

Dedication

To my parents for teaching me that making games can be as fun as playing them, to my brother for countless hours of Duke Nukem 3D, to my teachers for raising a generation on Oregon Trail, to my wife for being a Where's Waldo for Nintendo savant and for letting me be her Player 2, and to my cat for having the ability to turn a cardboard box and a ping pong ball into a real-time strategy game.

Acknowledgements

I would like to thank Todd Maddox, Bradley Love, Marc Tomlinson, Devon Greer, Rusty Gomez, Reid Tissing, Kristen Crabtree, and all the members of the Maddox Lab for their help in making this project possible.

Becoming a Gamer: Cognitive Effects of Real-Time Strategy Gaming

Brian Daniel Glass, Ph. D.

The University of Texas at Austin, 2012

Co-Supervisor: W. Todd Maddox

Co-Supervisor: Bradley C. Love

Video gaming has become a major pastime in modern life, and it continues to accelerate in popularity. A recent wave of psychological research has demonstrated that core perceptual changes coincide with video game play. Video games incorporate highly complex and immersive experiences which invoke a range of psychological mechanisms. This complexity has led to intractability which precludes determining which specific attributes of video gaming lead to cognitive change. The current work represents a research initiative which uses real-time strategy (RTS) games to boost executive functioning. In order to establish a link between video game features, video game behavior, and cognitive changes, an attention-switching tests two different forms of the same RTS game. Additionally, a difficulty titration paradigm attenuates individual differences in gaming skill. Thus, this project represents a critical advancement over prior research in that aspects of the video game itself were controlled and used to experimentally examine resulting cognitive change. Participants completed a psychological task battery before and after video game training, as well as at a mid-test. The battery covered a range of cognitive abilities including long-term memory, working memory, several attention-related constructs, risk taking, visual search, task switching and multitasking. These tasks were divided into two groups depending on the level of executive functioning components associated with the task performance. This resulted in a group of executive tasks and a group of other tasks. Because the high-switching gaming condition involves control and maintenance over a larger spread of gaming situations, performance on the executive task cluster was expected to improve more for this condition relative to the low-switching gaming condition. To reduce the impact of practice effects and the peripheral aspects of video gaming in interpreting the results, the Sims group was used a control baseline. A meta-analytical Bayes factor technique was used to determine the strength of performance changes from pre-test to mid-test, post-test, and follow up. By post-test, there was evidence that RTS training in the high attention-switching condition had improved on executive functioning tasks but not on other tasks. These results provide further evidence that video game training leads to psychological benefits over time.

Table of Contents

List of Tables xi

List of Figures xiv

Chapter 1: Overview 1

Chapter 2: Background and Significance 6

Observational Research: Gamers vs. Non-Gamers 6

Experimental Research: Video Game Training 9

Action Game Training and Perceptual Processing Speed 9

Video Game Training Differences and Transfer Specificity 10

Non-Gaming Training and Skill Transfer 14

The Non-Gaming Versus Gaming Distinction 14

Training Executive Functioning 15

How Complex Decision Making Tasks May Lead to Generalized Learning 18

Chapter 3: Experimental Design 20

Video Game Training 20

Video Game Installation and Tutorials 23

Video Game Procedure 24

Task Battery Procedure 25

Task Switching 25

Information Filtering 27

The Attention Network Test (ANT) 27

The Multimedia Multitasking Index (MMI) 29

Multi-location Memory Task 30

Visual Search 32

Stroop Test 33

WAIS-IV Digit Span 34

Operating Span 34

Balloon Analog Risk Taking Task (BART) 35

Task Counterbalancing 36

Chapter 4: Analysis and Meta-Analysis Techniques 37

EZ-Diffusion Modeling 37

Sims as Control Baseline 39

Meta-analytical Bayes Factor 40

Chapter 5: Results 43

StarCraft Gaming Results 43

Task Battery Results 44

Participants 44

Meta-Analytical Bayes Factor 45

The Multimedia Multitasking Index (MMI) 46

Stroop 46

Attentional Network Test (ANT) 47

Task Switching 49

Multi-location Memory Task 50

Operating Span (Ospan) 51

Balloon Analog Risk Taking Task (BART) 51

Visual Search 52

Information Filtering 53

Wais-IV Digit Span 55

Follow-Up Session 55

Bayes Factor Meta-analysis 56

Meta-Analysis for Follow-Up Session 62

General Discussion 65

Future Directions 68

Feature Analysis 68

How Novices Become Experts 69

Extracting Feature Weights for Training Regimens 69

Real-time Predictive Analysis of Behavior in Complex Tasks ..69

RTS Games as Method for Improving Component Processes 70

Gaming Interventions for ADHD 70
Gaming Interventions for Complications in Aging 71
APPENDIX: EZ-Diffusion Parameters and Assumptions...................................74
Alternative EZ-Diffusion Parameters 74
EZ-Diffusion Modeling Assumption Verification..74
Attentional Network Test (ANT)...74
Task Switching...75
Multi-Location Memory Task..75
Visual Search 76
Information Filtering...76
References...83

List of Tables

Table 1: ANT difference scores from pre-test to post-test. Significance of t-tests is reported relative to the Sims condition. Standard deviations are reported in parentheses. 48

Table 2: Task Switching difference scores from pre-test to post-test. Significance of t-tests is reported relative to the Sims condition. Standard errors in parentheses...49

Table 3: Multi-Location Memory Task difference scores from pre-test to post-test. Significance of t-tests is reported relative to the Sims condition. Standard errors in parentheses 50

Table 4: Visual Search Task difference scores from pre-test to post-test. Significance of t-tests is reported relative to the Sims condition. Standard errors in parentheses 52

Table 5: Information Filtering difference scores from pre-test to post-test. Significance of t-tests is reported relative to the Sims condition. Standard errors in parentheses 53

Table 6: Individual *t*-values used to compute meta-analytic Bayes factors for the Executive Tasks. For ANT, Task Switching, and Multi-location memory, *t*-values for drift rate are shown. These values are calculated by first taking the pre-test to post-test difference scores for each condition, and then comparing the SC condition to the Sims condition using a *t*-test...57

Table 7: Individual *t*-values used to compute meta-analytic Bayes factors for the Other Tasks. For Visual Search and Information Filtering, *t*-values for drift rate, boundary separation, and non-decision time are shown. These values are calculated by first taking the pre-test to post-test difference scores for each condition, and then comparing the SC condition to the Sims condition using a *t*-test. 58

Table 8: Meta-analytic Bayes Factors for Executive Task group and Other Task group, for SC-LO, SC-HI, and SC-BOTH (combined) on mid-test versus pre-test, post-test minus pre-test, and post-test minus mid-test, versus the Sims control group 61

Table 9: Meta-analytic Bayes Factors for Executive Task group and Other Task group, for SC-HI versus SC-LO on mid-test versus pre-test, post-test minus pre-test, and post-test minus mid-test....................................62

Table 10: Meta-analytic Bayes Factors for Executive Task group and Other Task group, for SC conditions combined on mid-test, post-test and long term follow-up session. These results includes only those who returned for a follow-up session 64

Table 11: Meta-analytic Bayes Factors for Executive Task group and Other Task group, for SC-LO, SC-HI, and SC-BOTH (combined) on mid-test versus pre-test, post-test minus pre-test, and post-test minus mid-test, versus the Sims control group, with boundary separation parameter when EZ-Diffusion modeling was possible...78

Table 12: Meta-analytic Bayes Factors for Executive Task group and Other Task group, for SC-HI versus SC-LO on mid-test versus pre-test, post-test minus pre-test, and post-test minus mid-test, with boundary separation parameter when EZ-Diffusion modeling was possible.....................79

List of Figures

Figure 1: This is a screenshot from the action video game called Unreal Tournament. The objective is to find and kill opponents. 2

Figure 2: This is a screenshot from the real-time strategy game StarCraft. Here, various units are engaging in a battle with the goal of eliminating enemy armies. In the bottom right, a mini-map depicts the entire game map, with various units represented in their team color. 5

Figure 3: Two different StarCraft map types. The Interior-Enemy Map (left) features an arctic style graphics theme with centrally located opponents, while the Exterior-Enemy Map (right) features a jungle style graphics theme with opponents located on the map edge. 22

Figure 4: Difficulty level reached by map type for SC-HI and SC-LO. Error bars represent standard error. This reflects similar game difficulty and engagement between SC-HI and SC-LO. 44

Figure 5: Information Filtering Task performance change scores from pre-test to post-test by number of distractors: 0, 2, 4 or 6. 54

Figure 6: *t*-values for mid-test (after 20 hours), post-test (after 40 hours), and from mid-test to post-test (from 20 to 40 hours) for drift rate parameter (for ANT, Task Switching, Multi-location Memory, Visual Search and Information Filtering) and scores (for Stroop, Ospan, BART, and Digit Span) 59

Figure 7: *t*-values for mid-test (after 20 hours), post-test (after 40 hours), and from mid-test to post-test (from 20 to 40 hours) for drift rate parameter (for ANT, Task Switching, Multi-location Memory, Visual Search and Information Filtering) and scores (for Stroop, Ospan, BART, and Digit Span), for SC-HI versus SC-LO 60

Figure 8: *t*-values for mid-test (after 20 hours), post-test (after 40 hours), and follow-up session for drift rate parameter (for ANT, Task Switching, Multi-location Memory, Visual Search and Information Filtering) and scores (for Stroop, Ospan, BART, and Digit Span)..........................63

Figure 9: *t*-values for mid-test (after 20 hours) and post-test (after 40 hours) for boundary separation parameter (for ANT, Task Switching, Multi-location Memory, Visual Search and Information Filtering) and scores (for Stroop, Ospan, BART, and Digit Span). Higher boundary separation values represent less caution. 80

Figure 10: *t*-values for mid-test (after 20 hours) and post-test (after 40 hours) for non-decision time parameter (for ANT, Task Switching, Multi-location Memory, Visual Search and Information Filtering) and scores (for Stroop, Ospan, BART, and Digit Span)..81

Figure 11: *t*-values for mid-test (after 20 hours), post-test (after 40 hours), and follow-up session for boundary separation and non-decision time parameter (for ANT, Task Switching, Multi-location Memory, Visual Search and Information Filtering) and scores (for Stroop, Ospan, BART, and Digit Span) 82

Chapter 1: Overview

Video gaming has become a major pastime in modern life, and it continues to accelerate in popularity. In 2008, 72% of the general population and 97% of teenagers aged 12-17 reported playing video games (Entertainment Software Association, 2010). In teenagers, 99% of boys and 94% of girls report playing video games, and females are the fastest growing demographic (Lenhart, Kahne, Middaugh, Macgill, Evans, & Vitak, 2008; Rideout, Roberts, & Foehr, 2005). Video games are being played more frequently and in more locations: 50% of all teens reported playing "yesterday", and 60% of all teens play video games on portable devices. There is also remarkable diversity in the types of video games played; the top 5 most frequently played games are a music game, an action game, a sports game, solitaire, and a dancing game (which requires actual dancing).

A recent wave of psychological research has demonstrated that core perceptual changes coincide with video game play (Green & Bavelier, 2007). Both observational and experimental studies reveal that action video gaming is associated with generalized perceptual enhancements (Dye, Green, & Bavelier, 2009; Castel, Pratt, & Drummond, 2005). Action video games are first person shooter games that incorporate fast paced kill-or-be-killed situations (Figure 1). Video games incorporate highly complex, dynamic, and immersive experiences which invoke a range of psychological mechanisms. This complexity has led to intractability which precludes determining which specific attributes of video gaming lead to cognitive change. The current work represents a research initiative which uses a different form of video game, real-time strategy (RTS) games, to boost executive functioning. In order to establish a link between video game features, video game behavior, and cognitive changes, an attention-switching manipulation is employed by testing two different forms of the same RTS game. Additionally, a difficulty titration paradigm attenuates individual differences in gaming skill. Thus, this project represents a critical advancement over prior research in that aspects of the video game itself were controlled and used to experimentally examine resulting cognitive change.

Figure 1: This is a screenshot from the action video game called Unreal Tournament. The objective is to find and kill opponents.

The attention-switching manipulation involved two different forms of the same RTS game. In one condition, game maps were developed to involve game situations localized to two main map areas. In this condition, the participant controls one friendly base in combat against one computer controlled enemy base. Although game action spreads throughout the map through the course of game play, the two-base starting condition results in fewer map locations which require attention and action during the course of a game. In the other condition, game maps were developed to involve a broadened set of game activity locations. In this condition, the participant controls two friendly bases versus two enemy bases. The game maps are twice as large, and game play takes place in more locations spread across the map. Importantly, attentional load is controlled between the conditions with the use of a difficulty titration procedure which hones in a 50% win rate. This ensures that the difference between the conditions is on attention-switching as opposed to level of engagement or task difficulty.

Participants completed a psychological task battery before and after video game training, as well as at a mid-test. The battery was designed to cover a range of cognitive abilities including long-term memory, working memory, several attention-related constructs, risk taking, visual search, task switching and multitasking. These tasks were divided into two groups depending on the level of executive functioning components associated with the task performance. This resulted in a group of executive tasks and a group of other tasks. Because the high-switching gaming condition involves control and maintenance over a larger spread of gaming situations, performance on the executive task cluster was expected to improve more for this condition relative to the low-switching gaming condition.

A meta-analytical Bayesian analysis was used to test the effect of video game training on these two task groups. Task battery performance from pre-test to mid-test and post-test was contrasted with a control group trained on The Sims 2, a life simulator game which does not require rapid shifts of attention (Li, Polat, Scalzo, & Bavelier, 2010). This condition is a game playing control condition in that participants are engrossed in a computer game for an equivalent amount of time, but without the rigorous attention-shifting involved in RTS.

The hypothesis that RTS gaming would lead to enhanced performance on executive functioning tasks was based on three main points. First, prior demonstrations of action video game training leading to changes in perceptual performance are evidence that video game training in general has the potential to lead to real psychological change. Second, RTS video gaming may require fewer speeded visual/perceptual requirements than action games and more attentional requirements than action games (Basak, Boot, Voss, & Kramer, 2008). Third, attentional processes have been shown to have high levels of plasticity (Posner & Rothbart, 2007) although previous attention training paradigms are not as immersive, enjoyable, and widely practiced as video gaming. An RTS gaming experience requires the participant to attend to a variety of evolving sub-situations and sub-tasks which combine to make up one real-time gaming environment (Figure 2).

Another possibility is that the RTS game versus life simulator game distinction may map on to a distinction made between complex tasks with specified versus nonspecified goals (Burns and Vollmeyer, 2002; Osman, 2010). Tasks with specified goals lead to less generalized transfer, and vice versa for tasks with nonspecified goals. It may be the case that RTS is more similar to an open ended complex task in that a rich variety of strategies can be discovered and tested, whereas life simulator game play involves more specified action routes, is less complex, and does not require rapid attention switching.

Figure 2: This is a screenshot from the real-time strategy game StarCraft. Here, various units are engaging in a battle with the goal of eliminating enemy armies. In the bottom right, a mini-map depicts the entire game map, with various units represented in their team color.

Chapter 2: Background and Significance

Observational Research: Gamers vs. Non-Gamers

A popular distinction in research of video gaming is made between video game players (VGP) and non-video game players (NVGP). While this distinction is convenient for observational studies, it is non-experimental in nature. These observational studies do not use random assignment and controlled training to explore the effects of video gaming. Still, many intriguing psychological differences have been found between these two groups. The vast majority of these observational studies used reported play of action video games to distinguish between VGP and NVGP. Generally, findings indicate that VGP have enhanced perceptual abilities characterized by faster reaction times in speeded-choice tasks.

Studies have found perceptual performance differences between VGP and NVGP. Castel, Pratt, and Drummond (2005) tested action VGP and NVGP in a visual search task involving the identification of a target item in an array of distractor items. They found that action gamers demonstrated increased performance in terms of inhibition of return and search efficiency. VGP had lower reaction times (RT) in speeded-choice RT tasks than NVGP (Dye, Green, & Bavelier, 2009a). Importantly, accuracy was equivalently high for both groups. The tasks included spatial cuing, inhibition of return, visual search, and the Simon task. Donohue, Woldorff, and Mitroff (2010) tested action gamers and non-gamers in tests of temporal processing. Gamers were more precise in judging the simultaneity of audiovisual stimuli. Also, gamers were more accurate in judging temporal order of stimuli presentation.

Prior research has also found differences in attentional breadth between VGP and NVGP. Dye, Green, and Bavelier (2009b) also compared action gamers and non-gamers in the Attentional Network Test (ANT). Action gamers demonstrated a larger flanker compatibility effect, indicating slower reaction times when distracting (i.e., incompatible) flankers are present on a trial. The authors interpreted this performance deficit as a form of enhanced processing in that action gamers utilized a larger attentional window, which incorporated more flankers for processing. Sungur & Boduroglu (2012) tested action game players and non-gamers in three perceptual tasks: a color identification task, a multiple-object tracking task, and an attentional breadth task (assessing field of view). Action gamers were more accurate than non-gamers in recalling the color of briefly presented items. Likewise, action gamers were able to track objects more accurately and in higher numbers than non-gamers. Action gamers also demonstrated a wider field of view in the attentional breadth task. This result is in agreement with Dye, Green, and Bavelier (2009b) who interpreted detriments in flanker compatibility effects as an indication of increased attentional breadth.

Studies have also found differences in motor performance between VGP and NVGP. Granek, Gorbet, and Sergio (2010) compared the BOLD response of action gamers vs. non-gamers (self-reported) in a visuo-motor task using fMRI. The experiment contrasted standard and non-standard motor movement tasks. In non-standard motor tasks, response options were linked to incongruent motor movements. For example, responding "left" by moving a joystick to the right. They found that during motor planning (they were unable to assess BOLD signal during motor actions due to motion artifacts) NVGP relied more on a parietal-prefrontal network than VGP, who relied more on recruiting prefrontal cortical areas. Prefrontal cortical areas are associated with well learned motor patterns, while the parietal-prefrontal network is thought to subserve the learning of new motor patterns. The authors interpret this finding as a generalization of practiced visuo-motor skills to a novel visuo-motor task (although such incongruent response pairings do exist in video gaming, such as airplane-style flight controls which navigate upwards by moving a joystick downwards). This result demonstrates a difference in the motor planning networks of self-reported action-video gamers. Rosser, Lynch, Cuddihy, Gentile, Klonsky, and Merrell (2007) investigated video gaming habits among medical students participating in a didactic surgical skills training program. They determined the medical students with past and current video game play experience demonstrated better laparoscopic and suturing performance. Video game playing medical students completed the course faster, committed fewer surgical errors, and achieved higher test scores, relative to non-gaming students

Additionally, Colzato, van den Wildenberg, Zmigrod, and Hommel (2012) found that action gamers were faster and more accurate in n-back tasks (1-back and 2-back versions), supporting the idea that there are group difference between those who play action games and those who do not.

There is the possibility that anatomical differences underlie gamers and non-gamers, other than changes which occurred due to video gaming history. Erickson, et al. (2010) found that dorsal striatum volume correlated with the performance level of acquisition for Space Fortress. The dorsal striatum has previously been mediating both early and later stages of learning, as well as cognitive flexibility (see: Atallah, Lopez-Paniagua, Rudy, & O'Reilly, 2007). Still, due to the observational nature of this study, it is an open question whether there are anatomical differences between VGP and NVGP, or whether video gaming caused these structural changes.

While there is substantial evidence that action VGP differ from NVGP in their perceptual, motor, and perhaps attentional attributes, these studies demonstrate observational differences between lifelong action game players and non-players. Observational research may suffer from self-selection biases, in that those who become

VGP differ in important ways from NVGP. Further research has been carried out to examine potential causal links between video gaming and cognitive performance.

EXPERIMENTAL RESEARCH: VIDEO GAME TRAINING

Action Game Training and Perceptual Processing Speed

While observational studies are an important first step, they beg the question as to whether these changes also come about in NVGP who engage in video game play. Green and Bavelier (2008) outlined a framework for how training-induced learning can lead to cognitive changes due to the brain's considerable plasticity. They point out that most psychological training procedures are specific to one cognitive skill. Thus, many researchers have been unable to demonstrate generalized benefits from long term training (Karni & Bertini, 1997; Seidler, 2004). Green and Bavelier (2008) identify video games as a new form of cognitive task which may tap a large range of psychological abilities, thus leading to generalized cognitive improvement. Thus, it may be possible that our experiment may lead to cognitive changes beyond the hypothesized changes in attention-related mechanisms.

To test whether action video games led to faster or more efficient visual information processing, Li, Polat, Makous, and Bavelier (2009) used a contrast sensitivity procedure to compare VGP and NVGP. They found that after training to become a VGP from a NVGP, individuals showed better contrast sensitivity. This low level perceptual paradigm indicates that video game playing leads to faster visual information processing. Although action video games appear to lead to faster visual information processing, it may also be the case that increases in attentional control or attentional resources may result from video game playing, as evidenced by the many speeded RT studies outlined in Dye, Green, and Bavelier (2009a).

In an attempt to determine gender differences and sustained long-term effects of action game training, Feng, Spence, and Pratt (2007) found 10 hours of action game training to reduce gender differences in a mental rotation task and a useful field of view task. Before training, males demonstrated higher accuracy in the tasks relative to females. After 10 hours of action game training, gender differences were no longer statistically significant. At a 5-month follow up, participants continued to perform better than their pre-test baseline (i.e., the learning persisted), however the gender difference had returned. In essence, 10 hours of action game training elevated performance and reduced gender differences, although the gender difference reappeared at a 5-month follow up.

Sanchez (2012) sampled an unrestricted population (i.e., regardless of previous gaming habits) to test whether a short action game session would lead to enhanced ability to read and answer questions about a science topic. Participants either played a word game or an action game for 25 minutes. Next, they completed mental rotation tasks as well as a science topic comprehension test regarding plate tectonics. Those in the action game condition demonstrated enhanced rotation performance as well as higher scores on science topic questions involving causal concepts. In summary, these results indicate that action video game training, even when very brief, can lead to changes such as faster speed of processing in speeded RT tasks. There is also the potential that short bursts of action video gaming lead to improvements in causal reasoning.

Video Game Training Differences and Transfer Specificity

Video games have been used with the purpose of teaching a specific skill set. As an example, Masson, Bub, and Lalonde (2011) tested whether video gaming can be used to teach principles of object motion. Middle school aged children who had yet to receive didactic training regarding Newtonian principles of physical motion were (or were not) subjected to six sessions of 1 hour game play. The video game, Enigmo, involves many different projectile and trajectory scenarios in a simulated 3D and fluid gaming environment (all with Newtonian principles). A test of the understanding of Newtonian motion before and after gaming revealed little or no enhanced understanding of physical motion. In fact, a 30 minute didactic session increased understanding substantially, with no interaction as to the gaming condition. While this represents one use of video game training—to teach a specific lesson—the focus of the current research is broader. The present investigation is to determine whether basic properties of attention can be altered in a generalized fashion.

While action video games were used in the tasks in the previous sections, two prior studies did involve RTS gaming. In a group of older adults ($Mean_{Age} = 69.1$, $SD_{Age} = 6.1$), Basak, et al. (2008) demonstrated that 23.5 hours of RTS gaming led to increased performance in a task switching experiment relative to a control condition that did not play video games. This improvement correlated with task performance. Unlike previous action video game results, Basak, et al. (2008) found improved visual perceptual skills on one of four tests. They found increased accuracy in a mental rotation task after RTS training, but not in functional field of view, attentional blink, nor enumeration (identifying the number of targets in a briefly presented field). This is a good indicator that RTS gaming, like action video gaming, can lead to cognitive changes. The results also suggest that RTS video gaming may train attentional processes as opposed to visual perceptual processes. Still, it is not known how RTS gaming affects healthy young adults, and whether there are longer term effects.

Boot, Kramer, Simons, Fabiani, and Gratton (2008) exposed non-gamers to 21 hours of gaming in an action game (Medal of Honor), RTS game (Rise of Nations), and a puzzle game (Tetris). They included a comprehensive task battery that covered various forms of attention, memory, and executive control. Their participants played video games four times per week in 1.5 hour sessions in the laboratory, resulting in 21 hours of game play over roughly 25 days. The researchers found no differences in task performance relative to two control groups comprised of either gaming experts or non-gamers. Moreover, the authors did not replicate the findings of Green and Bavelier (2006) who found faster reaction times on speeded reaction time tasks. The researchers report that this may have been due to two factors: differences in task procedures and differences in gaming schedule. Boot, et al. (2008) emphasized accuracy in their task instructions to normalize general differences in reaction times between gamers and non-gamers. Also, their gaming schedule required 1.5 hours 4 times per week, as opposed to 1 hour daily. In sum, these results indicate that video game type and playing characteristics may be important for driving psychological change.

Boot, et al. (2010) tested generalized skill transfer that may result from playing Space Fortress, a video game with heavy action components that requires some basic planning and monitoring (arguably much less complex than RTS gaming). The task battery consisted of tests of visual processing speed, selective attention, executive control, memory, and real-life simulations such as a flight simulator. The authors found that game training contributed mainly to improvement on tasks which were most similar to Space Fortress (i.e., Sternberg memory task, dual-task manual tracking, and flight simulator performance). These results indicate that generalized skill transfer does not always occur, and may be highly dependent on video gaming types.

Maclin, Mathewson, Low, Boot, Kramer, Fabiani, and Gratton (2011) used EEG to evaluate changes in attentional resource allocation that occurs as a result of training in Space Fortress. The testing procedure involved playing Space Fortress while concurrently responding to an oddball task paradigm. Behaviorally, the results indicated that with practice, participants were able to achieve high levels of performance on both the video game and the oddball task, simultaneously. ERP analysis revealed that game-only and oddball-only performance was associated with alpha suppression during task events. With practice, game-plus-oddball performance resulted in magnified alpha power during task events, reducing a performance tradeoff detected in early stages of game-plus-oddball learning. Thus, attentional changes can occur with practice to allow better concurrent task performance and attention allocation. It is possible that video games that are more complex than Space Fortress might represent in itself a variety of concurrent tasks, similar to the task procedure in Maclin, et al. (2011).

Bailey, West, & Anderson (2010) made a distinction between proactive and reactive cognitive control. Proactive control is the optimization of task preparation for an upcoming trial, while reactive control represents "just-in-time" adjustments made within-trial prior to response. They found that VGPs, relative to NVGPs, demonstrated attenuated proactive control and equivalent reactive control. The inclusion criteria for VGP required specific gaming experience on four popular action games. This result coincides with recent findings that associate certain forms of media exposure with cognitive detriments. First, Swing (2008) found a positive correlation between video game experience and a self-report scale of attention-deficit/hyperactivity disorder (ADHD) symptoms. This self-report questionnaire has been shown to relate to clinical symptomatology (Kessler, et al., 2005). However, no distinctions were made between types of video games played. Second, an investigation of media multitasking revealed that self-reported high multitaskers exhibited poorer task switching performance and were more susceptible to distractor interference in an information filtering task (Ophir, Nass, & Wagner, 2009). Again, this study did not distinguish between types of video games, but rather measured media multitasking by querying how often a participant overlaps various forms of media intake (e.g., music, phone calls, web surfing, video games, etc.) This pattern of results indicates there may be negative aspects associated with video game play and/or high media intake in general.

In summary, although action game training has been shown to lead to boosts in perceptual performance, it is unclear how robust video game training is to differences in training procedures and schedules (Boot, Kramer, Simons, Fabiani, & Gratton, 2008) and how generally video game training transfers to other skills (Masson, Bub, & Lalonde, 2011). Moreover, there may be some indication that there is a negative relationship between action video gaming and cognitive control.

NON-GAMING TRAINING AND SKILL TRANSFER

The Non-Gaming Versus Gaming Distinction

It is important to consider the mechanisms involved in long-term practice, and how this training can lead to skill transfer. The following section offers a review of non-gaming training paradigms which have led, with variable success, to generalized skill transfer. There are critical distinctions to be made between non-gaming training procedures and video game training.

First, non-gaming training procedures are often psychological tests which have been re-purposed for use as a training exercise. In other words, "training" regimens involve participants simply taking psychological tests over and over again, followed by post-testing in other tests. This format has found its way into the design of clinical training regimens which have been criticized as boring and not engaging enough for sustained use (Halperin & Healey, 2011; Prins, Dovis, Ponsioen, Brink, & van der Oord, 2011). Second, video gaming is a human activity found in everyday life, *in natura*. Thus, it is of basic research value to determine the potential psychological effects of video gaming. The attractiveness of video game training as a research tool is that, unlike other activities of daily life, video gaming has a quantifiable structure that can be controlled, analyzed, and repeated in a way that most other activities can not (e.g., sailing, painting). Third, video gaming, like other everyday activities, tap many different cognitive processes simultaneously. In order to determine whether training in very complex and immersive tasks lead to generalized transfer, it is necessary to engage a range of cognitive processes. Otherwise, the experimental paradigm reverts to the training of specific skills.

Training Executive Functioning

Posner and Rothbart (2007) point out that although many psychological domains may be subject to high specificity in training and learning, attentional networks may be an exception. The reason for this is that brain areas which subserve attentional performance and regulation are inherently designed to have widespread connections. This broad influence is required for effective attentional resource allocation, since different brain areas must be coordinated to tune attentional requirements. The authors indicate that further work is necessary to determine exactly how general and long-lasting attentional training can be. In order for general skill transfer to be possible, however, it must first be shown that attentional related training does lead to neural changes. Posner, Sheese, Odludas, and Tong (2006) demonstrate this by providing evidence of neural changes that occur from attention training. They found that training on the ANT led to increased connectivity between dorsal anterior cingulate cortex and lateral prefrontal cortex. This finding dovetails with the previously mentioned result that VGP demonstrate increased prefrontal activity when performing novel motor tasks. Thus, given the ability to train attentional related brain areas, and the widespread influence of these areas throughout the brain, it is possible that attention training can lead to generalized transfer to other forms of attention.

To support this notion, Jaeggi, Studer-Luethi, Buschkuehl, Su, Jonides, and Perrig (2010) have demonstrated that n-back task training can lead to enhanced fluid intelligence in matrix reasoning tasks. The authors found that training on both dual and single n-back task led to better performance on Raven's Advanced Progressive Matrices. It could be argued that attentional mechanisms are highly involved in the n-back task, and thus these results represent further evidence that attention training leads to generalized skill transfer. For example, 10 hours of n-back training has been shown to lead to expanded attentional breadth (Verhaeghen, Cerella, & Basak, 2004).

Dahlin, Nyberg, Backman, and Neely (2008) trained young adults for 5 weeks on a battery of inter-related working memory tasks (Miyake et al., 2000). These tasks involved viewing serially presented items until a pause occurred, at which point the participant responded with the previous 4 items viewed. Thus, success required the continuous updating of the previous 4 items. Tasks varied by stimuli type (e.g., letters, numbers, colors), although the core task remained the same. Post-test transfer tasks tested mental speed, working memory, episodic memory, verbal fluency and reasoning. There were little or no transfer effects, outside a 3-back test which is similar in nature to the training task. Taken together, these n-back findings demonstrate that simply changing the stimuli type may not be enough to allow narrow training tasks to result in generalized far transfer effects, although near transfer effects may be possible.

Chen and Morrison (2010) developed an adaptive complex working memory task (CWM) to examine the generalizability of EF training. Verbal and spatial trials were interweaved to form the CWM. On a verbal trial, a verbal decision task (word/non-word) was interweaved with items to be stored (single letters). On a spatial trial, a spatial decision task (symmetric/asymmetric) was interweaved with locations to be stored (points on a grid). Thus, the task was more complex than the similar WM tasks (e.g., Ospan) in that two different variations were concurrently administered. The researchers found that 4 weeks of training led to enhanced performance in the Nelson–Denny reading comprehension test as well was the Stroop task.

To determine whether learning to overcome distractors leads to transfer to different sorts of stimuli, Kelley and Yantis (2009) utilized an information filtering paradigm. They determined that the level of variability in training distractors drove generalizability of the transfer effect on other types of stimuli. When training was specific to one sort of distractor condition, learning occurred rapidly but was specific to the training stimulus type. When training varied across different sorts of distractor conditions, learning was slower but transferred more generally. This indicates that generalized learning may be enhanced by variability in practice.

Karbach and Kay (2009) examined executive control training by training young adults on task switching and testing their near and far transfer abilities. Near transfer tasks included task switching tasks with different stimuli sets. Far transfer tasks included Stroop, verbal and spatial working memory, and a measure of fluid intelligence. Task switch training (relative to a single task training control group) led to slight increases in near and far transfer tasks. Single task training led to decreases in Stroop performance on post-test, and slight or no improvement on the other tasks. Still, it is not clear whether single task training is an appropriate control task for training, since a repeated and very highly discriminable identification regimen appears to be somewhat repetitive. Nevertheless, this training regimen appears to have somewhat of a near and far transfer effect.

How Complex Decision Making Tasks May Lead to Generalized Learning

Kalyuga (2008) offers an overview of how a cognitive load perspective might inform transferable knowledge and skills. The author posits that in order for deep transferable knowledge to be gleaned from a training paradigm, the level of cognitive load must be germane to the task at hand. If task demands are overwhelming, there is no room for cognitive resources to acquire "higher order knowledge structures" that will lead to generalized learning. However, if cognitive load is too low then generalized learning also does not occur due to a lack of even specific skill acquisition. As evidence, Merrienboer, Schuurman, de Croock, and Paas (2002) varied cognitive load in a training exercise designed to teach participants principles of causal reasoning in complex systems. The training environment was a graphical simulation of a factory with several interacting parts and functions. The transfer environment was a textual programming language scenario, and problems (e.g., inappropriate for-loop statements) were to be found. When cognitive load was high (due to contextual interference), transfer performance was best. This pattern of results suggest that, when gaming difficulty is properly controlled between participants (by using a difficulty titration procedure), RTS training has a level of engagement that is germane to the task, thus leading to generalizibility in transfer.

In a review of complex decision-making environments, Osman (2010) details how tasks with specific goals might differ from those with nonspecific goals in transferability of acquired skills. Tasks with specific goals, or conventional goals, the participant can reduce the distance to the end-state using clear means. In tasks with nonspecific goals, the participant is unconstrained to explore the strategy space, and must determine for him or herself how to achieve the goal end-state (Sweller, 1989). Researchers have found that nonspecific goals lead to better knowledge transfer to different domains. While Sweller (1989) attributes this to cognitive load, Burns and Vollmeyer (2002) propose that this result is due to the attentional demands of the task. They posit that while nonspecific goals lead to focusing on the rule space as well as the instance space (how the rule is applied), specific goals lead to focusing just on the instance space. With the rule space already defined by the task, specific goals do not give leeway for learning how to search for rules. Therefore, it may be the case that video games with different types of goal specifications might lead to different levels or forms of skill transfer. For example, RTS game play with higher levels of attention switching may lead to increased generalizability and transfer.

In an investigation of abstract versus concrete symbolism in learning and transfer, Sloutsky, Kaminski, and Heckler (2005) varied training-to-transfer regimens by "relative concreteness". They found that tasks shrouded in concrete symbolism hindered transference to abstract domains, and vice versa. The authors conclude that abstractness in training have benefits for learning and transfer. This informs hypotheses regarding RTS game training in that video game training may represent highly abstract principles of attentional control. For example, learning to effectively switch between various simultaneously evolving game situations may transfer to concrete psychological tasks in a way that training on repetitive and specific psychological tests does not lead to transfer. Goldstone and Sakamoto (2003) demonstrated that tasks involving complex adaptive systems which share superficial elements do not lead to inter-task transfer as well as tasks that do not. This indicates that training on highly unique and abstract complex systems may lead to good skill transfer. It is likely that RTS video game training shares elements with this experimental paradigm in that successful RTS game play requires the player to master highly abstract principles in order to tame a complex adaptive system.

Chapter 3: Experimental Design

Participants were drawn from the University of Texas at Austin subject pool and screened for video game experience. Only those with <2 hours per week of video game experience were included. See Results for more participant details and demographics. Participants completed the task battery in three sessions: pre-test, mid-test, and post-test. Between pre-test and post-test, participants completed 20 hours of video gaming on their own laptop. Between mid-test and post-test, participants completed a final 20 hours of video gaming for a total of 40 hours of gaming. The inability for Boot, Kramer, Simons, Fabiani, and Gratton (2008) to detect changes in attentional processes as a result of RTS gaming after 21 hours is inspiration for mid- and post-testing. Increased training time may be critical for RTS training to bring about cognitive change. There were three video gaming conditions: High attention-switching StarCraft (SC-HI), Low attention-switching StarCraft (SC-LO), and a control condition, The Sims (Sims). The RTS conditions featured a highly modified version of RTS, as described below.

VIDEO GAME TRAINING

Each participant engaged in 40 total hours of video gaming. This video game play occurred outside the lab, on the participants' own laptops. Cognitive testing occurred over three sessions in the laboratory. Cognitive pre-testing occurs first followed by 20 hours of gaming, a mid-training-test, then an additional 20 hours of gaming, and finally a post-test. Participants were given 7 weeks (49 days) to complete the study, with the average completion occurring after 43.7 days (SD = 6.24).

The StarCraft conditions features alternating game maps. One map features a battle involving central starting locations while the other features a battle involving starting locations along the map edges (see Figure 3). A titration procedure is used to maintain engagement and interest. If a participant wins a game, the next game is more challenging (i.e., the enemy has more resources to use). If a participant loses a game, the next game is less challenging (i.e., the enemy has fewer resources to use). These titrations are separate for each map type. The Low attention-switching StarCraft condition consists of two alternating maps, both of which contain one friendly base and one enemy base. The High attention-switching StarCraft condition consists of two alternating maps, both of which contain one friendly base and two enemy bases. The two maps are also differentiated by surface features, one map with an arctic theme and the other with a jungle theme. In addition to surface feature differences, the arctic map has two enemy bases that are separate but centrally located (the Interior-Enemy Map), while the jungle map has two enemy bases that are located on opposite corners of the map (the Exterior-Enemy Map). This distinction results in differing game play and the requirement of different maneuvers and tactics. For example, in the Interior-Enemy map, the opponent is easier to surround but more resistant to siege. In the Exterior-Enemy map, the opposite is true.

Figure 3: Two different StarCraft map types. The Interior-Enemy Map (left) features an arctic style graphics theme with centrally located opponents, while the Exterior-Enemy Map (right) features a jungle style graphics theme with opponents located on the map edge.

In addition to the titration procedure, two critical game modifications were developed. First, the mini-map was modified to no longer issue alerts of distal game circumstances. The mini-map is permanently visible at the bottom corner of the StarCraft gaming interface and offers an overview of the full game world. Normally, visual and audio signals alert the player to situations that happen throughout the world, such as the creation of a new game unit or an attack from the enemy. In order to emphasize the switching and attentional maintenance of the gaming experience, these alerts were removed. This required the player to monitor and remember game play situations throughout the map, as the player could no longer rely on automatic alerts. Second, players were only allowed to use the mouse and not the keyboard during game play. StarCraft features many keyboard shortcuts, such as creating automatic waypoints and unit groupings that facilitate game play and limit the level of attention switching required to defeat the computer controlled enemy. Key presses resulted in a bothersome audio tone and visual message to appear on the screen which that reminded the player that key presses were not allowed and being recorded. The count of keyboard presses was recorded for each game and participants universally complied with the instruction to not use the keyboard.

The Sims operates as a control condition. Game play in this life simulation game consists of managing the lives of simulated characters that populate a virtual town. The game requires a low level of attentional multitasking and there are no specified or rigorous goal states to achieve as in StarCraft.

Participants are chosen from a screening procedure which included information about the frequency of video game per week play in the past year, as well as per week in a time when they played video games the most. Information about which games are played is also included. Those who self report <2 hours of video game play per week in the past year and at the time of highest frequency are invited to participate. Due to the very low proportion of males who qualify, only females are included in the present study.

For analysis and correlation with cognitive testing results, those in the StarCraft conditions are measured by their gaming success. For example, the proportion of games won and the difficulty level achieved are measures of RTS gaming performance.

Video Game Installation and Tutorials

After completion of the pre-test, each participant brought their personal laptop to the experimenter for video game installation. The Windows operating system was installed on a drive partition if the participant’s computer was an Apple. To streamline the video game running procedure, we developed a special computer program to automatically run the video game program with a singe click. This program also recorded the cumulative time of each gaming session, and did not allow the participant to pass the first 20 hours without the experimenter entering a special password, thus ensuring participants did not exceed 20 hours of gaming before the mid-test.

After game installation, the experimenter described the video game and instructed the participant to complete tutorials before beginning video gaming. Existing tutorials from the game developers were used, as they taught the essentials of game play. Participants in the StarCraft conditions were additionally provided with a game information packet which described game play in further detail. All participants were informed that our special computer program was recording their gaming and keeping track of their time.

For StarCraft, the computer program recorded hundreds of features from each gaming unit (like complex chess pieces) every 250msec. These features included information such as whether the unit was near friendly or enemy units, current proportion of health remaining, time since last under attack, etc. Importantly, the computer program also recorded which gaming units the participant was currently selecting. This allows for future analysis of in-game performance. For The Sims, saved game information was recorded which carries information about the characters and neighborhood built by the participant during game play.

Video Game Procedure

For StarCraft, each gaming session consisted of a single match versus a computer opponent. Two game maps were used and the computer program automatically set up the game so that the maps alternated between gaming sessions. Game lengths ranged between 10 minutes and 80 minutes. Alternating maps allows for varied game play to keep the participant engaged across sessions, while also offering a cross-validation paradigm for future data analysis projects that will analyze the feature and selection information of the gaming units. In addition to the variation in game maps, a titration procedure was developed to adapt the difficulty of the computer opponent to the success of the participant. Depending on if the participant won or lost the previous game, the following game would either be one level more or less difficult, respectively. Fifteen levels existed for each map, with higher levels involving more difficult computer opponents (represented by offering the computer opponent more starting resources). The highest level attained was 13 by one participant, but see below for more details.

For The Sims, each gaming session consisted of monitoring and controlling a single virtual household in a virtual neighborhood. Each household consisted of various family members, as characters (called Sims) would often die or be born. The participant controlled various domestic activities such as buying furniture, finding a job, planting a garden, and interacting with neighbors. All participants were instructed to not play more than three hours of video games per day, and to refrain from playing the games in small time increments.

TASK BATTERY PROCEDURE

In this section, the individual task procedures are detailed. Additionally, the section includes the justification for which tasks are incorporated into the Executive task group or the Other task group.

Task Switching

The task switching paradigm (Monsell, 2003) is a dual stimulus-classification procedure during which the participant must switch back and forth between two different task modes. During one mode, the participant identifies a letter as a vowel or consonant. During the other mode, the participant identifies a number as even or odd. After training on both modes, the participant switches between the two task modes. Switch costs are calculated by the reaction time difference between items which followed an item from the same task mode versus items which occurred after a switch in task mode (Ophir, Nass, & Wagner, 2009).

On each trial, participants were cued for 200msec with "NUMBER" or "LETTER", depending on the trial type. The stimuli consisted of a number/letter pair (e.g., "4n" or "n4"). On number trials, the participant indicated whether the number was even or odd. On letter trials, the participant indicated whether the letter was a vowel or consonant. Letters were chosen from the set {a, e, i, o, u, p, k, n, s}; Numbers were chosen from the set {2, 4, 6, 8, 3, 5, 7, 9}. The interval between cue and stimulus onset was 226msec, while the intertrial interval was 950msec. First, the participant completed three practice sets: letter task only, number task only, and switching practice. Next, the participant completed four 80 trial test blocks. Each block had a 40% switch rate, and all contained the same number of 1-, 2-, 3-, and 4-streak sequences. The task takes approximately 20 minutes to complete.

Task Switching was determined to be a member of the Executive task group due to previous results that indicate attentional control processes underlie task switching performance. Rubinstein, Meyer, & Evans (2001) demonstrated that increases in task switching complexity had an additive effective on performance decline. This pattern supports that a planning and control stage precedes response action in task switching. This supports Rogers and Monsell’s (1995) investigation of lateralized readiness potential in task switching which indicated preparatory motor responses and response conflict preceded response action in task switching.

Information Filtering

The information filtering task (Vogel, McCollough, & Machizawa, 2005) involves the detection of change (or lack of change) in sequentially presented scenes. Participants were shown scenes consisting of red and blue rectangles with differing orientation for 100msec. They were instructed to ignore the blue rectangles and focus only on the red rectangles. The scene was removed for 900msec and then replaced with a new scene. Participants were asked to respond within 2000msec whether one of the red rectangles had been rotated in the new scene. The target red rectangle was either not rotated or rotated by 45 degrees clockwise or counterclockwise. Thus, the participant must filter the scene information by ignoring blue bars and detecting whether red bars have changed. The intertrial interval was set to 200msec. Scenes included 0, 2, 4 or 6 blue distractors, and 2, 4, 6, or 8 red distractors. The rectangles did not overlap (nor were within one rotation of overlapping) and were distributed evenly across the scene. The total number of rectangles never exceeded 8, thus 10 combinations were possible. There were an equal number of combinations shown, trial order was randomized across subjects, and the number of change and no-change trials was equivalent. There are 256 trials, and the task takes approximately 25 minutes to complete.

The information filtering was determined to be a member of the Other task group due to its heavy perceptual compnents. Kelley and Yantis (2009) determine that simpler information filtering paradigms (as in the present incarnation) are more reliant on early perceptual processes that are more disrupted by changes in task procedure or heterogeneity of distractor items.

The Attention Network Test (ANT)

The ANT is a flanker identification task with various trial types that can be contrasted to measure different forms of attention (Fan, McCandliss, Sommer, Raz, &

Posner, 2002; Posner, Sheese, Odludas, & Tong, 2006). On any given trial, a fixation cross is presented (prefixation phase), followed by a cue, followed by a fixation phase, followed by a target item flanked by two items on the left and another two on the right. The target is always an arrow pointing either left or right, while the flankers are either all arrows (directional flankers) or lines (neutral flankers). The cue is a fixation cross that is sometimes accompanied by an asterisk either at fixation, above, below, or both above and below. Following the cue, the target item and flankers are presented either above the fixation cross or below.

Task timing was such that the prefixation phase varies randomly between 400 and 1600ms. The cue remains visible for 100ms and the fixation phase lasts 400ms. Next, the target items are displayed in a response-terminated fashion. Finally, an intertrial interval lasts for 3500ms minus the prefixation phase duration and minus the recorded reaction time for the present trial. The experiment consists of a 24-trial practice block, followed by two experimental blocks of 96 trials.

Three types of cues are used: no cue (fixation with no asterisk), center cue (asterisk at fixation), double cue (asterisks above and below fixation), or spatial cue (asterisk above or below fixation). Three types of flankers are used: congruent flankers (arrows in the same direction as target), incongruent flankers (arrows in the opposite direction as target), or neutral (lines with no directionality). Trial variation in cues and flankers allows for three critical contrasts: alerting (no cue minus double cue trials), orienting (center cue minus spatial cue), and executive control (incongruent flanker minus congruent flanker). These contrasts can be made in terms of reaction time as well as accuracy.

The ANT was chosen as a measure for the Executive task group due to evidence from the literature that ANT performance measures the efficiency of distinct neural attentional subsystems (Raz & Buhle, 2006). These subsystems consist of alerting attention (norepinepherine network), orienting attention (superior parietal lobe, temporal parietal junction, and frontal eye fields), and executive attention (dorsal anterior cingulated and lateral prefrontal cortex; Fan, McCandliss, Fossella, Flombaum, & Posner, 2005). While there is substantial independence between these networks, it is unsurprising that interactions have been found. For example, the alerting network has been shown to inhibit the executive network while the orienting network activates and enhances the executive network. In summary, the ANT is a robust measurement of the efficiency and effectiveness of different forms of attentional abilities.

The Multimedia Multitasking Index (MMI)

The MMI is an index of simultaneous media consumption developed by Ophir, Nass, and Wagner (2009). The participant is presented with a matrix of possible media combinations and self reports who often he or she simultaneously engages in both of the media types. These pairwise reports and weighted with self reported time spent engaging in each medium, and then summed into one score. The developers of the measure found that MMI correlates negatively with the above mentioned information filtering and task switching tasks. In other words, those who tend to multitask more often tend to also be worse at task switching and information filtering.

The questionnaire has two phases: 1) Hours spent on each medium, 2) Pairwise multitasking frequency:

1) The participant indicated how many hours per week she spends engaging in the following media types: Print media, Television, Computer-based video (e.g., YouTube), Music, Non-music audio, Video or computer games, Telephone and mobile phone voice

calls, Instant messaging, Text messaging, Email, Web surfing, and Other computer-based applications (e.g., word processing).

2) A matrix was displayed of each pairwise combination, and the participant responded for each combination how often she simultaneously engages in both forms of media (e.g., Telephone while Web surfing) -- The options were "Most of the time", "Some of the time", "A little of the time", and "Never". Rows and columns were blacked out for media types which the participant responded 0 hours/week in part one. The questionnaire takes approximately 5 to 10 minutes to complete.

For scoring purposes, the matrix responses were assigned numerical values of 1, 0.67, 0.33, and 0. Then, a subscore for each media type was calculated by multiplying the matrix response by the number of hours, and dividing this product by the total number of hours. These subscores were then summed into one final MMI score.

Multi-location Memory Task

A novel memory task was developed for the purpose of the present investigation. We found that the above task switching paradigm did not require retention of the state of the previous task when switching. Other researchers have attempted to tackle this problem using a primary task which is interrupted by a secondary task, throughout which the state of the primary task must be retained (Charron & Koechlin, 2010). Both tasks were the same and involved spelling the word ‘tablet’ one letter at a time. We expanded on this idea by developing a paradigm involving math operations. For this task, one of two screen locations is active for any given trial. After a single initialization trial during which two green zeros are shown at both location, the participant is shown a white number at one of the locations. The participant responds whether the sum of this white number and the previous green number shown at that location is odd or even, and then a new green number is shown at that location. Thus, the participant must retain the number at the previous location in order to do well. Switch costs of accuracy and reaction time are possible to calculate. Also, it is possible to sort the switch costs by the length of the previous streak.

The participant completed a 30-trial practice block followed by a 100-trial testing block. Before the first trial of each block, two green zeros are shown simultaneously for 0.5 seconds, one at the left location and one at the right location. Then, the zeros are removed and the block begins. Throughout the block, each location is indicated by a white rectangle. On each trial, one of the two locations is cued by the white rectangle becoming thicker for 0.5 seconds. Then, a white number is shown at the cued location (picked randomly from 1 to 9). The participant then indicated whether the sum of the current white number and the green number to be previously shown at that location was even or odd. The white number was cleared from the screen upon response. After 250msec, a new green number was shown at the current location for 0.5 seconds. The next trial began after this 0.5sec delay. The location was selected randomly with equal probability on each trial, although the maximum streak length was set to 10. This is exceedingly rare, since a streak of 10 occurs on only 1 out of 2^10 trials. In the practice block, corrective feedback and detailed instructions were displayed for the first 10 trials to ensure the participant understood the instructions. The entire task tasks approximately 15 minutes to complete. Despite the memory component involved in this task, due to the face value similarity to task switching, this multi-location memory task was a member of the Executive task group.

Visual Search

The visual search task (Castel, Pratt, & Drummond 2005) involves the identification of a target in an array of items. There are two trial types: hard and easy. Easy trials have fewer items (4), and hard trials have more (12). All items were represented by white bars on 75% gray disks against a 25% gray background. The target item, if present, had a vertical white bar. Non-target items had white bars that were rotated 45 degrees clockwise or counterclockwise. There were an equal number of target and non-target trials, as well as an equal number of easy and hard trials, with the ordering determined randomly. Items were located randomly on the vertices of an imaginary six by six array (36 possible locations) with vertices spaced 100 pixels apart horizontally and vertically. Disks were 40 pixels in diameter; bars were 6 pixels in width. On each trial, a fixation cross was displayed for 500msec, followed by the response terminated item array. Corrective feedback was displayed for 1sec, followed by a 500msec inter-trial interval. The task consisted of two blocks, each containing 5 practice trials and 96 test trials. The task took approximately 15 minutes to complete.

Visual search was deemed to be a member of the Other task group due to the heavy reliance on early perceptual components. VanRullen and Thorpe (2001) indicate that simple visual search occurs early in visual processing time course. While Wolfe (2003) reviews many ways in which memory and attention can modulate visual search, very simple serial search is highly dependent on set size and distractor complexity. Because the current task involves very simple targets that “pop out” from the distractors, there is little indication that attentional or higher level modulation will have an impact on basic task performance.

Stroop Test

The Stroop Test involves three verbal subtests during which the experimenter counts the number of correct and error responses within 45 seconds. First, the participant reads a list of words which include “Red”, “Blue” and “Green” in random order and printed in black ink. Second, the participant is presented with a list of X’s printed in red, blue or green ink and must say the color of the X’s aloud. Lastly, the participant is presented with a list of the three color words in which every word is printed in a different color ink than indicated by the printed word itself (e.g., “Red” printed in green ink). For each subtest, the experimenter instructs the participant to read the printed items aloud (or say the color of the printed items) as quickly as possible. The experimenter counts the number of completed items along with the number of errors made by the participant in 45sec. The scores from the three subtests are combined using standardized procedures to create an interference score. This score has been shown to correlate with frontal lobe functioning, speed of processing, and selective attention (Howieson, Lezak, & Loring, 2004). The task takes approximately 4 minutes to complete.

The Stroop interference paradigm is a gold standard in attentional research, and has been shown to be dependent on attentional systems. Bush, et al. (1999) demonstrated Stroop deficiency in attention-related disorders such as ADHD. Carter, et al. (1997) investigated anterior cingulate gyrus dysfunction in schizophrenia and determined Stroop performance to be highly correlated with executive control dysfunction. Kane and Engle (2003) offer evidence that Stroop performance is highly dependent on goal neglect and response competition, which are aspects of cognitive control and attention. In short, the Stroop task is a key member of the Executive task group.

WAIS-IV Digit Span

The WAIS-IV digit span task involves remembering and repeating back sequences of numbers ranging in length from 2 to 9, and both forward and backward. For each length, the participant attempts to recall two sequences. If the participant recalls at least one of the sequences correctly, then the experimenter proceeds to the next sequence. This repeats until the participant is unable to recall either sequence. A combined accuracy z-score is calculated using standard procedures (Wechsler, 2008). The task takes approximately 4 minutes to complete.

Digit span has been shown to improve with practice and training of working memory components (Olesen, Westerberg, & Klingberg, 2004). However, the memory component distinction between RTS gaming and the Sims control game does not inspire a specific hypothesis that RTS gaming will increase digit span performance relative to the Sims. Instead, the role of attentional components are hypothesized to be more distinctive between RTS gaming and the Sims. For this reason, as well as the fact that digit span has no classical resemblance to measures of cognitive control or complex decision making, the WAIS-IV digit span was included in the Other task group.

Operating Span

Operating span (Ospan; Turner & Engle, 1989) is a metric of working memory which involves the retention of words while performing concurrent mathematical problems. It has been shown to correlate with a wide yet somewhat sporadic range of psychological functioning and outcomes such as early onset Alzheimer’s, stress coping, and alcohol consumption (Unsworth, Heitz, Schrock, & Engle, 2005).

In the automated Ospan, participants used the mouse to respond to math equations and identify strings of memorized letters. On each trial, a letter is shown on the screen for 800msec. Then, an unsolved math operation is shown. The participant was instructed to click the mouse once she solved the operation. The operation was cleared from the screen and replaced by a number. The participant signaled whether the number was the correct solution to the operation by clicking on buttons labeled "True" or "False". Next, a new letter is shown. After a determined set of letters had been shown, the participant was presented with a matrix of letters and was instructed to recall the previous string of letters. The experiment includes three sets of each string size from 3 to 7, resulting in 75 trials. The Ospan score was calculated using only strings which were successfully recalled without error. The total number of trials from these successful sets was summed into one final score. The task takes approximately 10 minutes to complete.

Although the Ospan task has a large memory component, it has been demonstrated that Ospan performance is associated with the working memory component of executive control (Brewin & Beaton, 2001; Turner & Engle, 1989). Kane and Engle (2002) argue that working memory span is intimately involved as a basic component of executive attention, and correlates with attentional control. For this reason, Ospan was included in the Executive task group.

Balloon Analog Risk Taking Task (BART)

The Balloon Analog Risk Taking Task (BART) involves decisions regarding when to cash in on monotonically increasing rewards for which the risk of losing the entire reward increases as well (Lejuez, et al., 2002). The scenario is depicted graphically with an ever expanding balloon that pops when it reaches a predetermined size drawn from a normal distribution around a mean size (unknown to the participant). The optimal strategy (i.e., the optimal ration of risk to reward) converges on a pop rate of 50%. Risky strategies will have higher pop rates, while conservative strategies will have lower pop rates. Pop rate, as well as other available task metrics, have been found to correlate with sensation seeking, impulsivity, constraint deficiencies, and self reported risky behavior such as addictive, health, and safety risks (Hunt, Hopko, Bare, Lejuez, & Robinson, 2005).

The experiment consists of 10 balloons, with each balloon ending when either the balloon pops or the participant decides to cash in on the reward. At the beginning of each balloon, the temporary amount of reward collected returns to 0, while the cumulative amount of collected reward remains on the screen. On each trial, the participant chooses to either pump the balloon to cash in on the reward. On the first trial, if the participant chooses to pump the balloon, it will explode with a probability of 1/128. The probability of explosion is 1/127 on the second trial, 1/126 on the third trial, etc. On the 128th trial, the balloon always explodes. Thus, the average breakpoint was at 64 pumps. The task took approximately 8 minutes to complete.

While BART performance has been linked to constructs such as risk taking and impulsivity, there is little indication that BART performance is indicative of core executive control processes. There is no strong hypothesis to indicate that RTS game training may increase or decrease impulsivity relative to the Sims condition. For this reason, it is included in the Other task group.

TASK COUNTERBALANCING

Tasks were counterbalanced using a Latin square, with the task ordering maintained on each session. The MMI was only used on pre-test (1st session) and follow-up (4th session). In total, nine different orderings were used.

Chapter 4: Analysis and Meta-Analysis Techniques

EZ-DIFFUSION MODELING

Computerized tasks with two-choice response characteristics and for which reaction time and accuracy are available lend themselves to diffusion modeling (Ratcliff & Rouder, 1998; Vandekerckhove & Tuerlincky, 2007). These models are used to determine whether performance changes in speeded reaction time tasks represent true cognitive enhancements or simply speed-accuracy tradeoffs. Because it is common in prior research in video game training to draw conclusions about performance in speeded RT tasks, it is important to rigorously examine potential speed-accuracy tradeoffs.

Wagenmakers, van der Maas, & Grasman (2007) provide a simple yet robust diffusion model. This model, the EZ-diffusion model, simulates trials as a scenario in which evidence accumulates over time in the direction of one of two response boundaries. Responses are made when the evidence marker reaches one of the two response bounds. The EZ-diffusion model takes into account the mean and variance of response times for correct responses as well as response accuracy to produce the essential parameters of diffusion modeling: quality of information (i.e., drift rate), response conservativeness (i.e., boundary separation), and nondecision time. Speed-accuracy tradeoffs are reflected by differences in boundary separation. Smaller boundary separations lead to faster yet more inaccurate responses. Increased drift rates signify better integration of decision making factors without a speed-accuracy tradeoff. Nondecision time is modeled as a delay in the commencement of the evidence accumulation process and thus signifies differences in reaction time that are not due to changes in drift rate nor bound separation distance.

The EZ-diffusion model is a simplification of the Ratcliff diffusion model in that certain assumptions are made. Thus, in order to appropriately apply the EZ-diffusion model, the experimental task and empirically collected data should satisfy these assumptions. First, the two choices in the two-choice paradigm must be chosen with equal frequency. This precludes the use of biased response rewards or uneven correct response categories. Second, the overall reaction time distribution must be skewed right. Nearly all speeded reaction time experiment results are skewed right due to occasional longer reaction times being more frequent than very short reaction times. To verify this, the reaction time distribution for each experiment was subject to D'Agostino's K-squared test, a statistical test of skewness proposed by D’Agostino (1970) and recommended by Wagenmakers, van der Maas, & Grasman (2007). Third, the RT distributions for correct and error responses should be from the same distribution. This is violated in the majority of speeded reaction time experiments, and so the authors verified the robustness of the EZ-Diffusion to this particular misspecification using simulations. Their results indicate that this misspecification is not fatal in that it simply causes the model to globally underestimate parameter values, leaving ordinal comparisons intact.

In order to verify that the ANT results satisfy the assumptions made by the simplified EZ-Diffusion model, 1) the Bernoulli trial characteristics of the frequency of the two response key choices were compared using a binomial test., 2) the overall RT distribution, (after removal of outliers ±2 SD from the mean), were tested to be skewed from a normal distribution with D'Agostino's test of normality, and 3) the RT distributions for correct and error responses were determined to not be from different distributions using the Kolmogorov–Smirnov test.

Lastly, accuracies of exactly 1 should be specially handled. For accuracies of exactly 1, edge-correction should be applied to convert the accuracy using Equation 1

(Macmillan & Creelman, 2004), where n is the number of trials. These assumptions are considered for each task and reported in the Appendix.

$$A_{New} = 1 - \frac{1}{2n} \tag{1}$$

The tasks for which EZ-Diffusion parameters are calculated and presented are ANT, Task Switching, Memory Location Switching, Visual Search, and Information Filtering. Drift rate is the primary measure of interest, although the boundary separation parameter and non-decision time parameter are reported in the Appendix.

SIMS AS CONTROL BASELINE

Repeated administration of cognitive tasks often results in practice effects, which have been shown to occur over retest intervals of minutes, days, and weeks (Collie, Maruff, Darby, & McStephen, 2002; Hausknecht, Halpert, Di Paolo, & Gerrard, 2006). In clinical studies, these often provide major obstacles for acquiring a proper control group for longitudinal studies (McCaffrey, Ortega, & Haase, 1993). To overcome this problem, some have gone as far as to develop specialized tasks that are stable over multiple administrations (e.g., Falleti, Maruff, Collie, & Darby, 2006).

The present study has the benefit of random assignment, avoiding many complications associated with clinical or neurological conditions interacting with test practice. Furthermore, the Sims condition is used as a control baseline condition against which SC-LO and SC-HI are compared. Performance change from pre-test to all three retest sessions (mid-test, post-test, and long-term follow-up) is calculated as a difference score. Comparisons are then made between the SC condition and Sims. In this way, retest effects are subtracted from the SC condition performance metrics.

Using a control baseline is important because performance has been shown to improve on retest across a wide range of cognitive tasks. Bartels, Wegrzyn, Wiedl, Ackermann, and Ehrenreich (2010) demonstrated that practice effects are most pronounced in the first few sessions and then stabilize over time. Attention, learning and memory, executive functions, and motor functions all improved monotonically over three retests within three months of first test performance. At six and 12 months, performance stabilized or dropped slightly but non-significantly. The pattern was similar although much less pronounced for language and visuospatial functioning. These results underscore the importance of not simply comparing performance before and after treatment, but rather comparing changes in performance against a control condition. For our purposes, the Sims represents a critical control group.

META-ANALYTICAL BAYES FACTOR

The goal of the analyses is to determine whether post-test differences in the two groups of tasks (executive tasks and other tasks) are substantial. To accomplish this, results from the various experiments are combined using a meta-analytic approach. While each task is analyzed and reported separately below, a meta-analytic analysis is appropriate to determine cognitive effects on a range of tasks.

A meta-analytic Bayes factor approach was used to determine task performance across the two task groups (Rouder & Morey, 2011). This approach calculates a single Bayes factor from the results of multiple independent experiments. To do so, the critical dependent measure from each task compared against the control Sims group to generate a paired t-value. The t-values within one group of tasks are combined to calculate a single meta-analytic Bayes factor.

The critical contribution of the meta-analytic Bayes factor is the ability to detect performance differences for a group of experiments that may not be substantial or significant within a single experiment. As Rouder and Morey (2011) explain, when a series of experiments with modest yet directional agreement are considered individually in a vacuum, they would each be dismissed as mustering unsubstantial support. However, when the group of experiments is taken as a whole, then modest yet directional agreement across the experiments amounts to a very strong case.

An individual Bayes factor is a ratio between the likelihood function of two models given some set of data. Thus, the Bayes factor can be used as an alternative to classical hypothesis testing in that no Gaussian or other distributions are referenced or assumed. Instead, the ratio of the probability of an alternative model is pitted against the probability of a null model, given some set of data. Unlike likelihood-ratio tests which use maximum likelihood parameter estimation, Bayes factors are calculated by integrating over the parameter space (Goodman, 1999; Wasserman, 2000). Equation 2 is the basic form of an individual Bayes factor.

$$BF = \frac{\Pr(D \mid M_1)}{\Pr(D \mid M_2)} = \frac{\int \Pr(\Theta_1 \mid M_1)\Pr(D \mid \Theta_1, M_1)d\Theta_1}{\int \Pr(\Theta_2 \mid M_2)\Pr(D \mid \Theta_2, M_2)d\Theta_2} \qquad (2)$$

Where D represents the empirical data, A represents the alternative hypothesis, N represents the null hypothesis, and θ represents the parameter set. Although one might suspect that multiple Bayes factors can be combined via multiplication, since both the numerator and denominator of the equation represent probabilities, this is not the case. This can be shown by randomly drawing several simulated results from a single underlying distribution. The individual Bayes factors will be small and often lower than 1, thus their product will likely be small despite all experiments being driven by the same underlying distribution. Pooling the results from this simulation results in a large Bayes factor, however experiments generally cannot (and should not) be pooled due to practical differences (e.g., one should not simply concatenate reaction time results from two different experimental procedures). Thus, Rouder and Morey (2011) offer a meta-analytic

Bayes factor which unifies distinct experimental results via the common language of the *t*-statistic.

The meta-analytic extension for the Bayes factor has a one-tail and two-tail form, depending on whether a Cauchy or half-Cauchy distribution is used as a prior on effect size. In the present study, we have no expectation given prior research that the Sims condition will outperform the SC-HI and SC-LO groups on post-test, thus we seek to contrast what we hypothesize to be an effect (RTS game training) versus no effect (Sims game training). For this reason, the half-Cauchy distribution is used.

Similar to classical, frequentist approaches to model comparisons, there is no quantitative threshold for determining whether the alternative model is favored over the null. Instead, Jeffreys (1961) offered a qualitative interpretation which has become the norm (Good, 1992; Aitkin, 1991): ratios less than 1:1 support the null hypothesis, 1:1 to 3:1 are insubstantial, 3:1 to 10:1 are substantial, 10:1 to 30:1 are strong, 30:1 to 100:1 are very strong, and ratios greater than 100:1 are decisive.

Chapter 5: Results

STARCRAFT GAMING RESULTS

Two procedural mechanisms were put in place to keep participants engaged in StarCraft gaming and limit within- and between-condition variation in gaming difficulty. First, games alternated between two different gaming maps. Second, a difficulty titration procedure was used. When a participant won a game, the following game featured an opponent that was slightly more difficult to defeat. Likewise, losing a game caused the opponent in the following game to be easier to defeat.

The titration procedure had the expected effect on the proportion of games won. On average, those in the SC-LO condition won 42.6% of their games (SD = 8.8%). Those in the SC-HI condition won 43.0% of their games (SD = 8.7%). This demonstrates that the conditions did not differ by gaming successful.

Final game level reached (ranging from 1 to 15) for SC-LO was 4.17 (SD = 2.52) in the Interior-Enemy Map, and 2.50 (SD = 1.25) for Exterior-Enemy Map. Final game level reached for SC-HI was 3.09 (SD = 1.44) in Map 1, and 2.83 (SD = 1.15) in Map 2. Figure 4 illustrates the level reached over time for SC-HI and SC-LO.

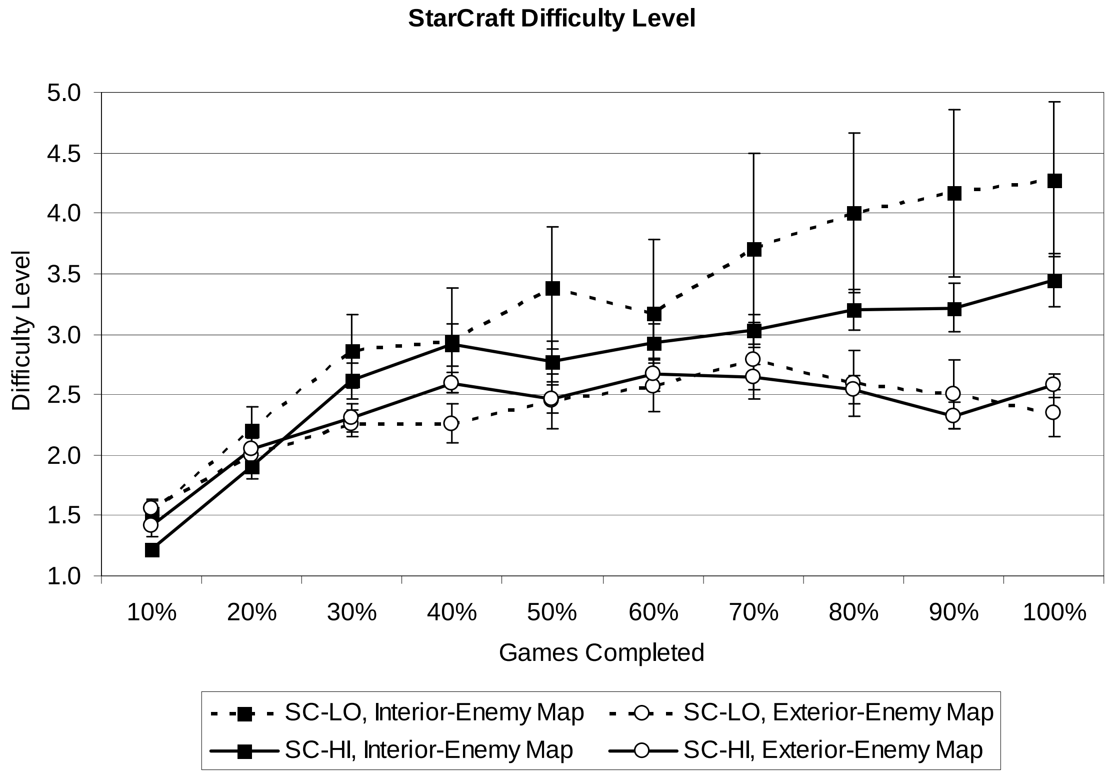

Figure 4: Difficulty level reached by map type for SC-HI and SC-LO. Error bars represent standard error. This reflects similar game difficulty and engagement between SC-HI and SC-LO.

TASK BATTERY RESULTS

Participants

Potential participants were screened with a web-based questionnaire regarding their current and past video gaming habits. The principle items of the questionnaire were, “In the last year, how many hours per week do you tend to play video games?” and “Prior to the last year, think back to the period in which you most frequently played video games. How many hours per week did you tend to play during that period?” Participants who reported 2 hours or less of video game play per week qualified for inclusion in the study. The average hours of current weekly game play was 0.38 (SD = 0.55). Participants were also asked to indicate the genres of video games they currently play, and no included participants endorsed real-time strategy games.

Due to the inability to secure a tenable source of qualified male participants, inclusion was restricted to females. The median video gaming hours per week was 8 hours for male respondents and 1 hour for female respondents. Out of 933 respondents, 22 males qualified for participation versus 207 qualified females.

Qualified participants were randomly assigned to one of the three conditions: SC-HI (n = 20), SC-LO (n = 26), or Sims (n = 26). The mean age was 20.4 years (SD = 1.1) for the SC-HI condition, 20.3 years (SD = 1.1) for the SC-LO condition, and 19.9 years (SD = 0.8) for the Sims condition. All participants had normal or corrected to normal vision, and none were colorblind. The Institutional Review Board of the University of Texas at Austin approved the study and informed consent was obtained from all participants.

Meta-Analytical Bayes Factor

As described above, the main focus of the analysis is on the meta-analysis of two main clusters of tasks: Executive tasks and Other tasks. To accomplish this, we calculated *t*-statistics to compare the RTS gaming conditions with the Sims control condition. Pre-test to post-test differences for critical dependent measures were calculated for each condition, and then compared against the Sims control condition with a *t*-statistic. This statistic was then combined within task cluster to generate a single Bayes factor. The following results represent specific task analysis, for completeness, but the critical meta-analytical results are presented below in the section titled Bayes Factor Meta-Analysis.

The Multimedia Multitasking Index (MMI)

MMI multitasking scores were calculated for each participant. Higher MMI scores signify greater amounts of time simultaneously engaged in more than one form of media (Ophir, Nass, & Wagner, 2009). The MMI score was 4.71 (SD = 1.57) for SC-HI, 4.31 (SD = 1.64) for SC-LO, and 4.40 (SD = 1.51) for the Sims condition. A one-way analysis of variance revealed no significant difference between these groups, $F(2, 100) = 0.63$, $p = 0.53$. The results also matched Ophir, Nass, and Wagner (2009) who reported an overall mean of 4.38 (SD = 1.52). In their interpretation of low and high multitaskers as those one standard deviation below and above the mean, respectively, the current group of participants would be considered to report a moderate level of multitasking.

Stroop

Stroop interference scores were tabulated and normalized according to previously established procedures (Botvinick, et al., 2001). The resulting Z-scores indicate performance relative to previously established population norms. As in most standardized tests, positive Z-scores represent better performance relative to population mean, and vice versa. The Stroop Z-scores at pre-test were 0.75 (SE = 0.12) for SC-HI, 0.78 (SE = 0.14) for SC-LO, and 0.87 (SE = 0.20) for the Sims. A one-way ANOVA revealed no differences between these groups on pre-test, $F(2, 51) = 0.18$, $p = 0.83$. The Z-score change on post-test was 0.56 (SE = 0.19) for SC-HI, 0.24 (SE = 0.15) for SC-LO, and 0.04 (SE = 0.14) for the Sims. A t-test revealed a significant difference between SC-HI and the Sims, $t(40) = 2.24$, $p = 0.03$. The other pairwise t-test results were non-significant, $p > 0.2$. A one-way ANOVA revealed a marginally significant difference between group means for the difference score, $F(2, 56) = 2.83$, $p = 0.07$.

Attentional Network Test (ANT)

Beyond the standard measures of reaction time and accuracy, the Attentional Network Test (ANT; Fan, et al., 2005; Raz & Buhle, 2006) offers three measures of attention: an alerting score (maintaining attentional readiness), an orienting score (selective attention), and an executive control score (resolving conflicting actions). The alerting score is calculated by subtracting center-cue trials from no-cue trials. The orienting score is calculated by subtracting spatial-cue trials from center-cue trials. The executive control score is calculated by subtracting congruent trials from incongruent trials. Table 1 reports the post-test difference scores for SC-HI, SC-LO, and the Sims.

	SC-HI	SC-LO	Sims
Median RT (ms)	-80.41 (18.81)	-46.14 (24.43)	-61.87 (13.31)
Accuracy	0.10 (0.03)*	0.06 (0.04)	0.01 (0.01)
Alerting Score (RT ms)	35.21 (17.14)	7.14 (12.71)	21.45 (11.70)
Orienting Score (RT ms)	-27.44 (13.68)‡	-13.82 (10.29)	-0.26 (7.67)
Executive Score (RT ms)	-75.53 (25.83)*	-51.00 (26.38)	-17.32 (11.31)
Alerting Score (Accuracy)	-0.01 (0.01)	-0.03 (0.02)	-0.02 (0.02)
Orienting Score (Accuracy)	0.01 (0.01)	0.03 (0.02)	-0.01 (0.01)
Executive Score (Accuracy)	0.28 (0.09)*‡	0.17 (0.11)	0.02 (0.03)
Log(Drift Rate)	-4.65 (-4.82)†	-4.89 (-4.75)	-5.25 (-5.26)
Boundary Separation	-0.04 (0.54)	-0.16 (0.48)	-1.10 (0.39)
Nondecision Time	-19.64 (8.20)	-12.82 (13.03)	-14.71 (11.18)

*$p < 0.05$, relative to the Sims condition; †$p < 0.06$, relative to the Sims condition
‡ Significant difference of absolute pre-test score: $p < 0.05$, relative to the Sims condition

Table 1: ANT difference scores from pre-test to post-test. Significance of t-tests is reported relative to the Sims condition. Standard deviations are reported in parentheses.

Task Switching

Performance on classic task switching was measured by calculating overall RT and accuracy, as well for trials that switched between a vowel/consonant letter identification task and an odd/even number identification task. Switch costs are calculated by subtracting non-switch trials from switch trials, for both RT and accuracy. Table 2 reports post-test performance change from pre-test for overall RT and accuracy as well as the costs of switching.

	SC-HI	SC-LO	Sims
Median RT (ms)	-199.59 (54.38)	-117.31 (31.58)	-172.42 (36.91)
Accuracy	0.012 (0.016)†	0.015 (0.020)	-0.021 (0.007)
RT Switch Cost (ms)	-51.33 (25.60)	4.65 (51.01)	-51.28 (29.09)
Accuracy Switch Cost	-0.024 (0.012)	0.013 (0.007)	-0.008 (0.010)
Log(Drift Rate)	-6.51 (-7.23)	-6.12 (-7.01)	-6.63 (-7.55)
Boundary Separation	-1.12 (0.61)	-0.98 (0.79)	-2.34 (0.55)
Nondecision Time	-94.64 (24.84)	-46.08 (19.32)	-39.9 (19.90)

*$p < 0.05$, relative to the Sims condition; †$p < 0.06$, relative to the Sims condition

Table 2: Task Switching difference scores from pre-test to post-test. Significance of t-tests is reported relative to the Sims condition. Standard errors in parentheses

Multi-location Memory Task

Performance on the Multi-location memory task was characterized by overall RT and accuracy, as well as RT and accuracy on location-switch trials relative to location-remain trials. Switch costs are calculated by subtracting location-remain trials from location-switch trials, for both RT and accuracy. Table 3 reports post-test performance change from pre-test for overall RT and accuracy as well as the costs of switching locations.

	SC-HI	SC-LO	Sims
Median RT (ms)	-44.86 (109.04)*‡	-227.76 (106.49)	-374.61 (91.27)
Accuracy	0.034 (0.032)	0.007 (0.033)	-0.020 (0.025)
RT Switch Cost (ms)	104.32 (122.12)*	-71.14 (90.93)	-247.49 (92.54)
Accuracy Switch Cost	0.088 (0.046)	0.081 (0.027)	0.043 (0.028)
Log(Drift Rate)	-6.68 (-7.16)	-6.69 (-7.19)	-7.23 (-7.51)
Boundary Separation	0.28 (0.88)	-0.80 (0.86)	-1.91 (0.84)
Nondecision Time	-46.71 (44.28)	-116.43 (43.79)	-121.35 (34.24)

*$p < 0.05$, relative to the Sims condition
‡ Significant difference of absolute pre-test score: $p < 0.05$, relative to the Sims condition

Table 3: Multi-Location Memory Task difference scores from pre-test to post-test. Significance of t-tests is reported relative to the Sims condition. Standard errors in parentheses

Operating Span (Ospan)

The experiment involves 75 trials, with each trial representing the memorization of a string of letters with length between 3 and 7, inclusive. The Ospan score was calculated using only strings which were successfully recalled without error. The total number of trials from these successful sets was summed into one final score. The pre-test scores were 44.3 (SE = 3.9) for SC-HI, 40.6 (SE = 4.1) for SC-LO, and 51.9 (SE = 3.8) for Sims. A one-way ANOVA revealed no significant differences between the conditions, $F(2, 54) = 1.96$, $p = 0.15$. The post-test difference scores from pre-test were 5.7 (SE = 2.6) for SC-HI, 6.7 (SE = 4.7) for SC-LO, and 4.1 (SE = 3.5) for Sims. A one-way ANOVA revealed no significant differences between the conditions for post-test difference scores, $F(2, 54) = 0.13$, $p = 0.88$.

Balloon Analog Risk Taking Task (BART)

In the BART, the optimal strategy (i.e., the optimal ration of risk to reward) converges on a pump rate of 50%, or 64 pumps administered. Number of pumps is calculated by taking the mean number of pumps per balloon trial. At pre-test, number of pumps administered was 27.4 (SD = 8.70) for SC-HI, 35.0 (SD = 9.22) for SC-LO, and 28.7 (SD = 12.59) for Sims. This is on par with prior results which found pump rates of 25.0 (SD = 9.6) for female participants, although the present results are less variable. There was a significant difference between SC-LO and SC-HI at pre-test, $t(33) = -2.53$, $p < 0.05$. However, the groups remained stable from pre-test to post-test. The post-test difference score for pumps administered was 3.47 (SE= 2.35) for SC-HI, 3.54 (SE= 2.18) for SC-LO, and 4.39 (SE= 2.59) for Sims. Pairwise t-tests identified no significant differences between the groups (*p*s > 0.8). A one-way ANOVA revealed no differences between the conditions for post-test change of pumps administered, $F(2, 53) = 0.05$, $p = 0.96$.

Visual Search

Performance on the visual search task is characterized by overall RT and accuracy, as well as for two trial types: hard and easy. Easy trials have fewer items (4), and hard trials have more (12). Table 4 reports median RT and accuracy for hard trials, easy trials, and overall.

	SC-HI	SC-LO	Sims
Median RT (ms)	-191.5 (30.8) ‡	-129.6 (20.9)	-148.4 (20.7)
Accuracy	-0.006 (0.01)	0.005 (0.01)	-0.001 (0.01)
Easy Trials (RT)	-125.9 (16.0)	-96.74 (16.0)	-103.5 (15.7)
Hard Trials (RT)	-285.1 (61.6) ‡	-208.3 (41.6)	-252.5 (38.4)
Easy Trials (Accuracy)	-0.008 (0.01)	0.010 (0.01)	-0.002 (0.01)
Hard Trials (Accuracy)	-0.004 (0.02)	-0.001 (0.01)	-0.001 (0.01)
Log(Drift Rate)	-5.78 (-7.32)	-5.45 (-7.11)	-5.53 (-7.21)
Boundary Separation	-1.63 (0.40)	-1.67 (0.35)	-1.61 (0.24)
Nondecision Time	-68.54 (14.4)	-28.45 (16.2)	-48.75 (15.48)

*$p < 0.05$, relative to the Sims condition
‡ Significant difference of absolute pre-test score: $p < 0.05$, relative to the Sims condition

Table 4: Visual Search Task difference scores from pre-test to post-test. Significance of t-tests is reported relative to the Sims condition. Standard errors in parentheses

Information Filtering

In the information filtering task, trials varied by the number of distractor items. Trials had either 0, 2, 4 or 6 distractors. Overall RT and accuracy were calculated, as well as RT and accuracy for the four distractor types. Additionally, the EZ-Diffusion model was used to calculate drift rates, boundary separation distances, and nondecision times overall and for the four distractor types. Post-test difference scores relative to pre-test results for trials are reported in Table 5. These differences scores are illustrated as a function of number of distractors in Figure 5.

	SC-HI	SC-LO	Sims
Median RT (ms)	-63.87 (25.96)	-54.97 (21.18)	-66.03 (31.04)
Accuracy	0.005 (0.018)	0.016 (0.032)	0.020 (0.017)
Log(Drift Rate)	-7.85 (-7.08)	-6.72 (-6.87)	-6.50 (-7.13)
Boundary Separation	-0.15 (0.25)	0.22 (0.26)	-0.41 (0.25)
Nondecision Time	-53.60 (16.46)	-59.11 (14.77)	-39.38 (22.27)

Table 5: Information Filtering difference scores from pre-test to post-test. Significance of t-tests is reported relative to the Sims condition. Standard errors in parentheses

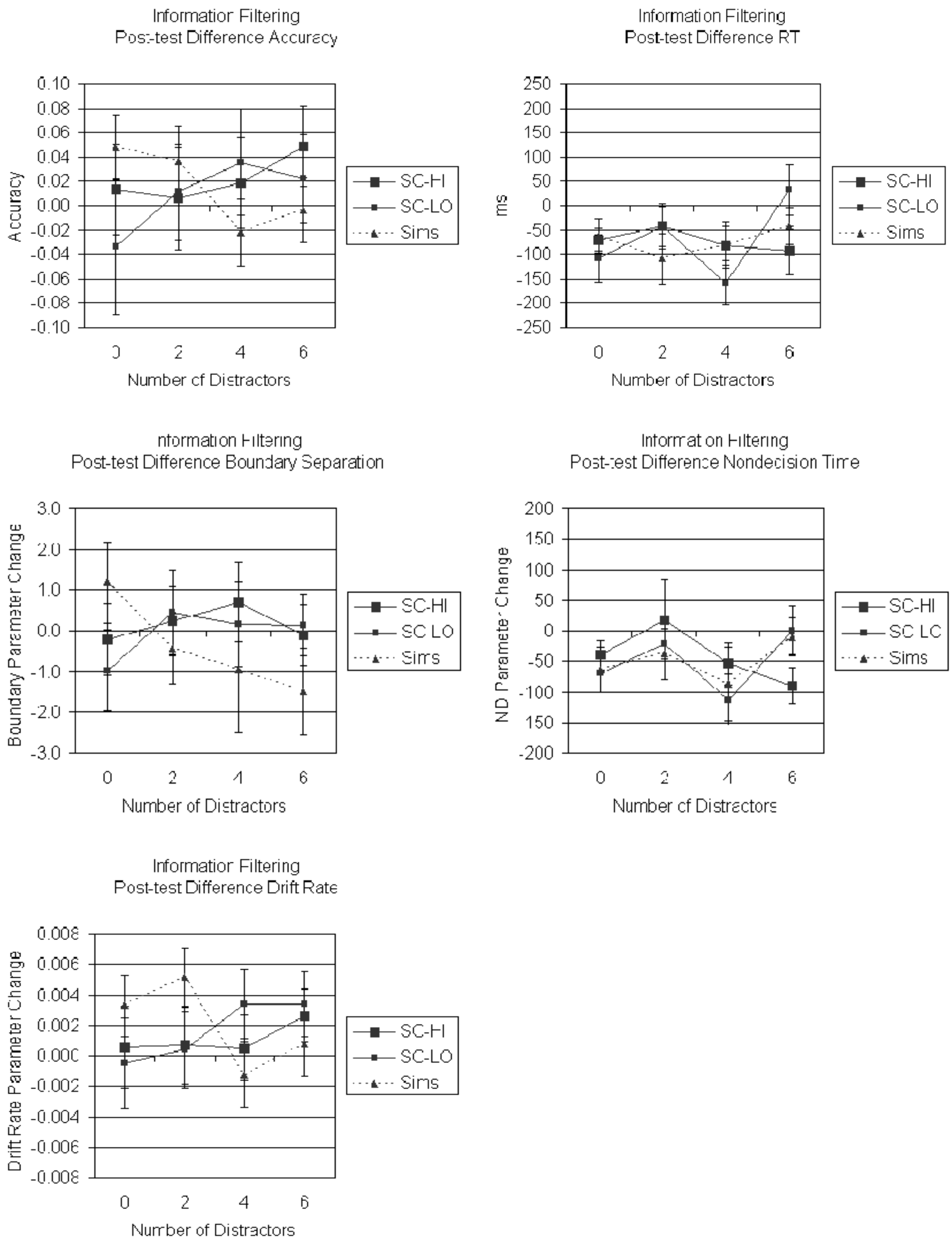

Figure 5: Information Filtering Task performance change scores from pre-test to post-test by number of distractors: 0, 2, 4 or 6.

Wais-IV Digit Span

Digit span Z-scores were calculated using standard procedures (Wechsler, 2008). Forward and backward span performance was measured by the number of number strings recalled correctly (two possible per string length) as well as the long string length recalled correctly. These scores are compared against a population norm, resulting in a Z-score. The pre-test Z-score was 0.10 (SE = 0.17) for SC-HI, 0.18 (SE = 0.22) for SC-LO, and 0.31 (SE = 0.18) for Sims. A one-way ANOVA revealed no group differences on pre-test, $F(2, 56) = 0.32$, $p = 0.73$. Post-test difference relative to pre-test was 0.52 (SE = 0.14) for SC-HI, 0.18 (SE = 0.18) for SC-LO, and 0.56 (SE = 0.20) for Sims. A one-way ANOVA revealed no group differences for the difference scores, $F(2, 56) = 1.35$, $p = 0.27$.

FOLLOW-UP SESSION

To assess potential long-term effects, participants were given the opportunity to return for a follow-up session. During the follow-up session, participants completed the full task battery and completed a questionnaire about their current video gaming habits. The follow-up session occurred an average of 145 days after the completion of the post-test. The follow-up group consisted of 12 (of 22) participants from SC-HI, 12 (of 19) participants from SC-LO, and 10 (of 22) participants from Sims. Participants were compensated $50 for completion of the follow-up session.

Following their participation in the follow-up task battery, participants were given a questionnaire about their gaming habits. Those in the SC-HI condition reported playing 1.18 hours per week (SD = 2.75), versus 0.23 (SD = 0.41) at pre-test (although this was driven by one participant who reported playing 9 hours per week, up from 0). Those in the SC-Lo condition reported playing 0.59 hours per week (SD = 1.44), versus 0.24 (SD = 0.41) at pre-test. Those in the Sims condition reported playing 0.38 hours per week (SD

= 0.74), versus 0.46 (SD = 0.60) at pre-test. A one-way ANOVA revealed no differences between the groups on the difference between reported follow-up and pre-test gaming hours per week, $F(2,30) = 0.76$, $p = 0.48$.

Participants were asked which video games they currently play, if any. No participants reported playing RTS games. For action video games, 18% of SC-HI endorsed, 17% of SC-LO endorsed, and 25% of Sims endorsed. For all other types of video games, 27% of SC-HI endorsed, 25% of SC-LO endorsed, and 25% of Sims endorsed. Again, the mean increase in reported video gaming from pre-test to follow-up was 25.2 minutes per week (SD = 115.4 minutes).

BAYES FACTOR META-ANALYSIS

To determine the two task battery groups (executive tasks and other tasks) we computed meta-analytic Bayes factors for each cluster of tasks. The executive task group was comprised of Stroop, ANT, Task Switching, Multi-location Memory, and Ospan. The other group was comprised of BART, Visual Search, Information Filtering, and Digit Span. Because the meta-analytic approach involves one dependent measure per experiment, we used drift rate parameters when available (boundary separation and non-decision time parameters are reported in the Appendix). This metric combines speed and accuracy into one parameter. Tables 6 and 7 report the *t*-value for the post-test change in performance relative to Sims condition.

For Stroop and Digit Span, Z-scores are used in lieu of drift rate because 1) calculating drift rates is not feasible, and 2) standardized procedures for normalized scoring are well established for these tasks. For Stroop, the color-word interference Z-score is used. For Digit Span, the forward-backwards combined Z-score is used. For Ospan, the performance score is used. For BART, mean balloon pumps per trial is used.

The *t*-values for post-test difference from pre-test, mid-test difference from pre-test, and post-test difference from mid-test are illustrated in Figure 6. The *t*-values for SC-HI compared against SC-LO (as opposed to comparing against the Sims condition) are illustrated in Figure 7.

	SC-HI	SC-LO	SC-BOTH
Stroop	2.84*	1.32	3.06*
ANT	2.24*	0.86	2.28*
Task Switching	0.23	0.98	0.92
Multi-location Memory	0.68	0.69	0.98
Ospan	0.61	0.56	0.81

*$p < 0.05$

Table 6: Individual *t*-values used to compute meta-analytic Bayes factors for the Executive Tasks. For ANT, Task Switching, and Multi-location memory, *t*-values for drift rate are shown. These values are calculated by first taking the pre-test to post-test difference scores for each condition, and then comparing the SC condition to the Sims condition using a *t*-test.

	SC-HI	SC-LO	SC-BOTH
BART	0.61	-0.39	-0.56
Visual Search	-1.29	0.42	-0.58
Information Filtering	-1.34	-0.29	-1.00
WAIS-IV Digit Span	-0.25	-2.66*	-1.76

Table 7: Individual *t*-values used to compute meta-analytic Bayes factors for the Other Tasks. For Visual Search and Information Filtering, *t*-values for drift rate, boundary separation, and non-decision time are shown. These values are calculated by first taking the pre-test to post-test difference scores for each condition, and then comparing the SC condition to the Sims condition using a *t*-test.

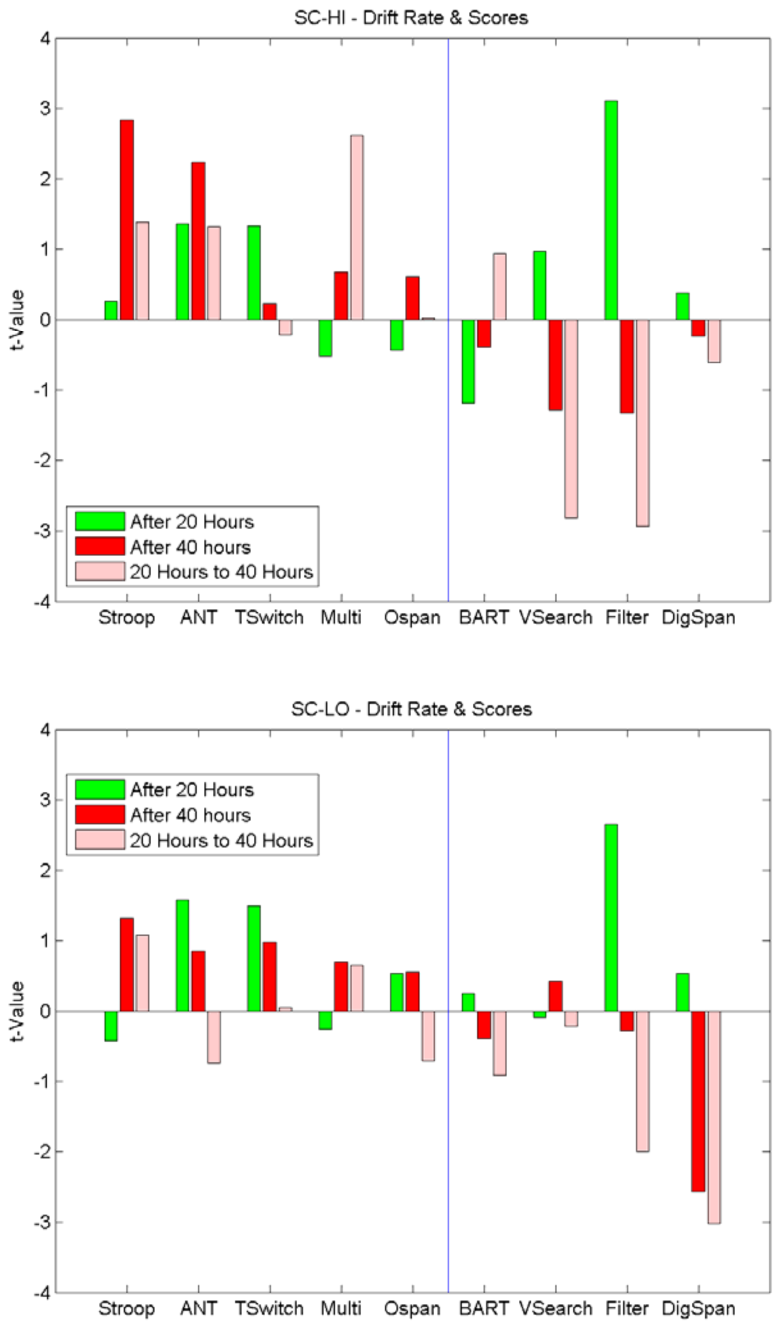

Figure 6: *t*-values for mid-test (after 20 hours), post-test (after 40 hours), and from mid-test to post-test (from 20 to 40 hours) for drift rate parameter (for ANT, Task Switching, Multi-location Memory, Visual Search and Information Filtering) and scores (for Stroop, Ospan, BART, and Digit Span)

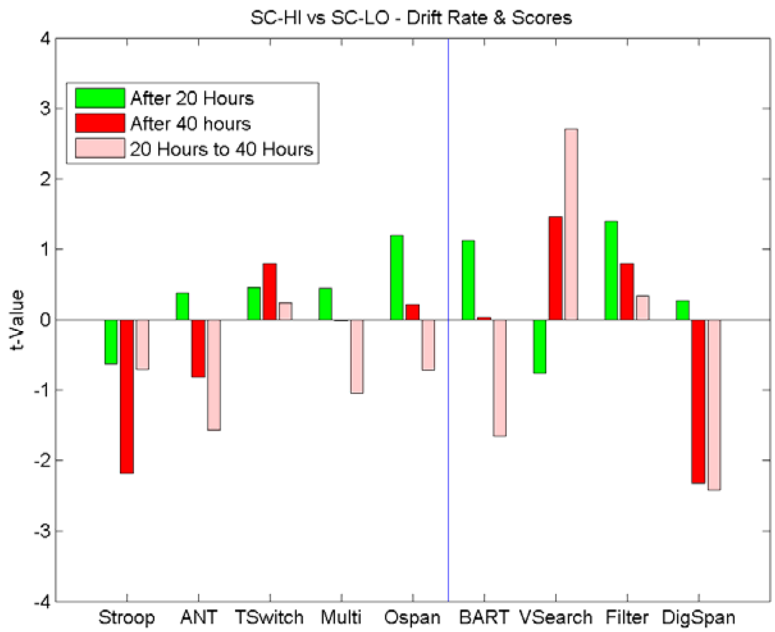

Figure 7: *t*-values for mid-test (after 20 hours), post-test (after 40 hours), and from mid-test to post-test (from 20 to 40 hours) for drift rate parameter (for ANT, Task Switching, Multi-location Memory, Visual Search and Information Filtering) and scores (for Stroop, Ospan, BART, and Digit Span), for SC-HI versus SC-LO

The *t*-values are then combined using the meta-analytic Bayes factor technique (see above). For SC-HI, the meta-analytic Bayes factor is 6.77 for Executive tasks and 0.04 for Other tasks. For SC-LO, the meta-analytic Bayes factor is 1.17 for Executive tasks, and 0.05 for Other tasks. For both SC conditions combined, the meta-analytic Bayes factor is 40.76 for Executive tasks, and 0.02 for Other tasks. Because a Bayes factor is the ratio of the probability likelihood of one model versus another (in this case an alternative model versus a null model), Bayes factor ratios below 1:3 (i.e., $0.\bar{3}$) represent evidence in favor of the null hypothesis.

	Mid-test minus Pre-test (0 to 20hrs)	Post-test minus Mid-test (20 to 40hrs)	Post-test minus Pre-Test (0 to 40hrs)
SC-LO Executive Tasks	0.32º	0.11º	1.17
SC-HI Executive Tasks	0.16º	1.73	**6.77***
SC-BOTH Executive Tasks	0.24º	0.59	**40.76****
SC-LO Other Tasks	0.50	0.03ºº	0.05ºº
SC-HI Other Tasks	0.32º	0.03ºº	0.04ºº
SC-BOTH Other Tasks	0.56	0.02ºº	0.02ºº

*BF > 3, substantial evidence for alternative hypothesis
**BF > 10, strong evidence for alternative hypothesis
ºBF < $0.\overline{3}$, substantial evidence for null hypothesis
ººBF < 0.1, strong evidence for null hypothesis

Table 8: Meta-analytic Bayes Factors for Executive Task group and Other Task group, for SC-LO, SC-HI, and SC-BOTH (combined) on mid-test versus pre-test, post-test minus pre-test, and post-test minus mid-test, versus the Sims control group

	Mid-test minus Pre-test (0 to 20hrs)	Post-test minus Mid-test (20 to 40hrs)	Post-test minus Pre-Test (0 to 40hrs)
SC-HI vs. SC-LO Executive Tasks	0.20^{o}	0.06^{oo}	0.04^{oo}
SC-HI vs. SC-LO Other Tasks	0.26^{o}	0.11^{o}	0.07^{oo}

^{o}BF < $0.\overline{3}$, substantial evidence for null hypothesis
^{oo}BF < 0.1, strong evidence for null hypothesis

Table 9: Meta-analytic Bayes Factors for Executive Task group and Other Task group, for SC-HI versus SC-LO on mid-test versus pre-test, post-test minus pre-test, and post-test minus mid-test

Meta-Analysis for Follow-Up Session

In total, 34 of 63 participants returned for a follow-up session. For this reason, in our analysis of follow-up performance, we included only the pre-test, mid-test, and post-test results only for those who completed the follow-up as well. Additionally, the SC groups are combined due to the limited number of participants.

For this subset of participants, the meta-analytic Bayes factor for Executive tasks decreased to 0.51 at follow up, from 44.43 at post-test. There was no substantial evidence for increased performance on Other tasks at post-test or follow-up.

Figure 8 illustrates the *t*-values for the SC conditions relative to Sims condition for only the participants who returned for the follow-up session. The meta-analytic Bayes factor for follow-up results in Table 10.

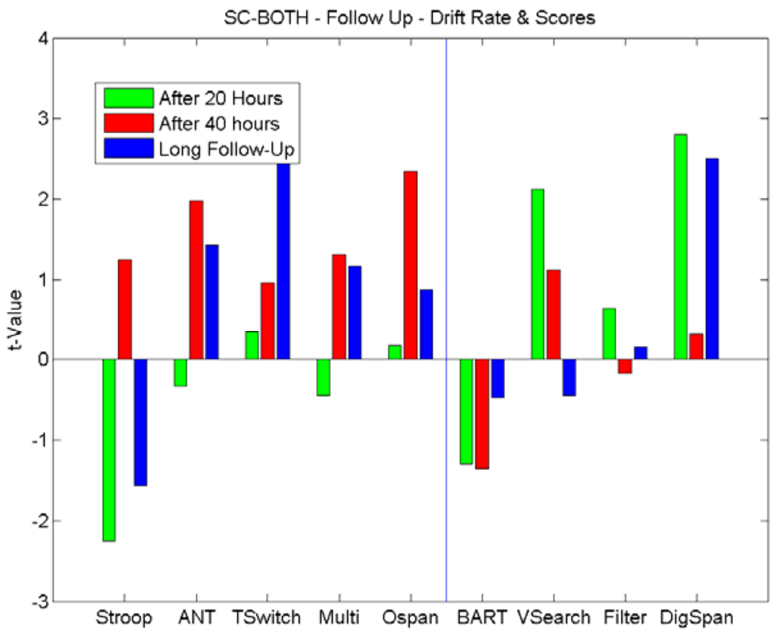

Figure 8: *t*-values for mid-test (after 20 hours), post-test (after 40 hours), and follow-up session for drift rate parameter (for ANT, Task Switching, Multi-location Memory, Visual Search and Information Filtering) and scores (for Stroop, Ospan, BART, and Digit Span)

SC-BOTH	Mid-test minus Pre-test (0 to 20hrs)	Post-test minus Pre-test (0 to 40hrs)	Follow-up Session minus Pre-test†
Executive Tasks All	0.05ºº	**44.43****	0.51
Other Tasks All	2.34	0.11º	0.29º

**BF > 10, strong evidence for alternative hypothesis
ºBF < $0.\bar{3}$, substantial evidence for null hypothesis
ººBF < 0.1, strong evidence for null hypothesis
†Table includes only those who returned for a follow-up session

Table 10: Meta-analytic Bayes Factors for Executive Task group and Other Task group, for SC conditions combined on mid-test, post-test and long term follow-up session. These results includes only those who returned for a follow-up session

Chapter 6: Discussion

General Discussion

This study sought to compare the cognitive consequences of video game training on two different forms of the same video game: a low attention-switching version of StarCraft (SC-LO) and a high attention-switching version of StarCraft (SC-HI). Results indicate that by 40 hours of gaming, those in the SC-HI condition demonstrated improved Executive task performance relative to the Sims gaming control condition. When a subset of long term follow-up participants was considered, performance for Executive as well as Other tasks were no longer substantially different relative to controls.

Comparisons were made relative to a control video gaming condition, for which participants played The Sims 2. A psychological task battery consisted of two main subsets: Executive Tasks and Other Tasks. The Executive Tasks consisted of Stroop, ANT, task switching, multi-location memory, and operating span. The Other Tasks consisted of BART, information filtering, digit span, and visual search. The task battery was administered as a pre-test before video game training began. Participants then completed 40 hours of video game training, with a mid-test task battery administration after 20 hours. A post-test task battery was administered after video game training was complete. For some participants, a long-term follow up task battery was also completed.

To identify whether performance change in speeded reaction time tasks was due to real enhancements, a diffusion modeling approach was employed. This yielded the critical parameter of drift rate, which reflects true processing enhancements. When considering all tasks together in the meta-analysis, diffusion modeling was not always appropriate or possible, and so standardized scores or critical dependent measures of interest were used.

A meta-analytical Bayes factor technique was used to determine the strength of performance changes from pre-test to mid-test, post-test, and follow up for the Executive Tasks cluster and Other Tasks cluster. At mid-test, there were no substantial performance changes for SC-LO or SC-HI relative to Sims. By post-test, there was evidence that SC-HI had improved on Executive tasks but not on Other tasks.

When the boundary separation parameter was considered for Executive tasks, there was decisive evidence that SC-HI demonstrated increased boundary separation parameters (Jeffreys, 1961 proposes Bayes factors of 3:1 signify substantial evidence, 10:1 signifies strong evidence, and 100:1 signifies decisive evidence). The boundary separation results were such that those in the SC-HI group demonstrated increased boundary separation on post-test relative to the Sims. Nevertheless, the drift rates parameter is highly separable from the boundary separation parameter, and increases in drift rate indicate a cleaner overall diffusion process with enhanced information quality in the diffusion process.

These results provide further evidence that video game training leads to psychological benefits over time. Specifically, RTS gaming can improve performance on executive functioning tasks. To reduce the impact of practice effects and the peripheral aspects of video gaming in interpreting the results, the Sims group was used a control baseline. The benefits of this comparison are two fold. First, changes in task performance that resulted simply from re-testing at mid-test and post-test are minimized by factoring out practice effects from a control group. Second, the peripheral aspects of video game training were factored out from the effects measured for the treatment conditions. In other words, the Sims group was also spending time playing a game on a laptop, just as the RTS groups were. This makes it more likely to appreciate differences detected in the RTS group as representing the true impact of RTS gaming, as opposed to the impact of increased time spent playing any sort of game or performing any sort of computer-based activity.

A previous experiment involving the training of RTS game play, Boot, Kramer, Simons, Fabiani, and Gratton (2008) found no effect of RTS gaming on a comprehensive battery which included tests of attention, memory, and executive control. It is important to note two important differences between their study and the current study. First, their study involved video game training in a laboratory setting, as opposed to home-based training. Second, and most critically, their study involved 21 hours of video game training. In accordance with their findings, the present study also did not demonstrate significant changes in Executive task battery performance after 20 hours of gaming. It was not until a full 40 hours of gaming that the RTS groups began to show performance differences on Executive tasks. There are interesting implications for this distinction between the current study and Boot, et al. (2008). First, it likely that more than 20 hours of RTS video game training are required to have an impact on attentional processes. Second, it may also be the case that increased familiarity with the task battery might be required before differences can be detected between RTS gaming and control gaming groups. Some have suggested that it may be important to familiarize participants with test procedures, and future work should consider stabilizing performance via re-testing before administering video game training (Bartels, Wegrzyn, Wiedl, Ackermann, & Ehrenreich, 2010; Hausknecht, Halpert, Di Paolo, and Gerrard, 2006; Wesnes & Pincock, 2002)

In comparison with the results from action game training, there are key differences in benefits from RTS gaming. Action gaming have been shown to reduce reaction times relative to controls for speeded reaction time tests with high visual and perceptual components. Such tasks include visual search, contrast sensitivity, mental rotation, and useful field of view (Dye, Green, & Bavelier, 2009; Feng, Spence, & Pratt, 2007; Li, Polat, Makous, & Bavelier, 2009). The present study does not show improvements resulting from RTS video gaming for low-level perceptual tasks such as visual search and information filtering. This is an important step in identifying how different forms of video games affect different cognitive and perceptual components.

In addition to the contrast with prior results in action gaming, the present study also identifies differences within RTS gaming that drive cognitive change. Two forms of the same RTS game (StarCraft) were compared: high and a low attention-switching versions. Due to a carefully controlled titration procedure, the proportion of winning was equivalent between the two groups. This signifies that gaming difficulty did not differ between the groups but rather the in-game attributes, minimizing the role of general engagement or participant interest. In the SC-LO gaming maps, there were one participant base which battled against one computer opponent base. In the SC-HI gaming maps, there were two participant bases and two opponent bases. Increases in the number of bases leads to more game characteristics to track and command. Results in the psychological test battery indicate that those in the SC-HI condition demonstrated increased performance in Executive tasks (but not Other tasks) at post-test relative to the Sims control group, after 40 hours of video game training. This novel finding indicates that game attributes drive cognitive changes.

FUTURE DIRECTIONS

Feature Analysis

Attributes of gaming behavior were recorded in combination with video game state features in real time. This allows for a unique quantitative method for linking video game behavior with the magnitude and time course of cognitive change. Thousands of game characteristics were recorded multiple times per second along with participant selection activity. Bayesian feature selection methods are used to calculate which video game features the participant is attending to and reacting to over time (Berger & Pericchi, 1996).

How Novices Become Experts

Data mining and model selection techniques distill usable parameters that represent participant strategy and performance attributes such as the scope of attended features and how feature attention changes over time. An application of this procedure is to compare the gaming behavior of novices to experts by determining the nature and time course of a novice participant's advancement in proficiency and whether these advancements coincided with cognitive changes or generalized learning.

Extracting Feature Weights for Training Regimens

An intriguing use of feature selection capabilities is the construction of a training procedure. The use of specified training procedures can be used to guide novice participants into expertise. Expert RTS gamers can be used to establish feature weights that in turn predict a novice player's information needs (Love, Jones, Tomlinson, & Howe, 2009). A key question is whether novices who received training assistance enjoy the same cognitive improvements as those who self-learned. Another interesting question is whether novices learn better by using training procedures extracted from expert players or from their own successful previous game play. The techniques outlined here, along with promising preliminary results, represent a clear pathway to understanding the relationship between the human and the environment in complex and dynamic situations.

Real-time Predictive Analysis of Behavior in Complex Tasks

With large amounts of gaming behavior data, it becomes possible to predict an individual's future actions based on real-time characteristics of the environment. Using collected data, it is possible to retroactively link environment features to player behavior at any given time. Thus, it becomes possible to use these retroactive feature weights to estimate the future probability of any given action based on the current state of the environment. The development of these procedures would represent a key advancement for the development and verification of state-action models to describe human behavior in complex and dynamic situations.

RTS Games as Method for Improving Component Processes

The present study represents a foundation for using RTS video gaming to boost executive functioning and determining the nature of these changes by modeling the relationship between video gaming factors and cognitive change. Future work should extend this research into areas of mental health. Beyond using RTS gaming to enhance executive function in healthy individuals, RTS gaming may be used to target deficits in populations with specific executive functioning dysfunction. These include attention deficit hyperactivity disorder (ADHD) and complications associated with normal aging.

Gaming Interventions for ADHD

ADHD describes an idiosyncratic constellation of inattentive, impulsive, and hyperactive behaviors (Durston, 2008). As such, the field varies greatly in the understanding and modeling of ADHD. Castellanos, Sonuga-Barke, Milham, & Tannock (2006) proposes that ADHD may be an umbrella term for separate dysfunctional processes involving *hot* versus *cool* executive functioning (EF). ADHD in cool EF may be associated with attention dysfunction while hot EF may be associated with affective dysfunction.

Interestingly, RTS games may have the unique ability to tap both of these processes, offering a more robust tool for assessment and intervention. Computerized training paradigms designed to tap only cool processes are often criticized as boring and not engaging enough for sustained use. ADHD intervention paradigms involving video game constructs and imagery have already been shown to improve intervention compliance (Prins, et al., 2011). While effective in improving working memory, it is difficult to build on these results or learn about the mechanisms involved. The authors admit that "it is not clear … which of the various elements of the game format contributed to superior training efficacy." This represents a serious hurdle as well as an important opportunity. Halperin and Healey (2011) propose that current computer-based training batteries lack a directed play and social aspect which is important in assuaging ADHD symptoms. RTS gaming has a unique ability to involve a social element, whether playing against artificially intelligent opponents or actual human opponents.

Gaming Interventions for Complications in Aging

Prior investigations have found evidence for high levels of plasticity in cognitive processes for those in late adulthood. Bherer, Kramer, Peterson, Colcombe, Erickson, & Becic (2008) demonstrated that older adults can enjoy generalized performance enhancement in dual-task scenarios from long-term training. Training occurred one hour per day for a 3-week period. Training experience transferred to various forms of dual-task paradigms that included within-modality and cross-modality task configurations. This work suggests that cognitive training in older adults is possible.

In a review of cognitive training in older adults, Hertzog, Kramer, Wilson, & Lindenberger (2009) drew a distinction between low-transfer tasks that lead to specific training improvements only, and high-transfer tasks that lead to generalized improvements in cognitive functioning. Low-transfer tasks include simple and unstructured task environments such as signal identification tasks, which result in narrow transfer effects. High-transfer tasks invoke a broad range of cognitive mechanisms, are highly engaging, and involved structured activities. In support, Mozolic, Long, Morgan, Rawley-Payne, and Laurienti (2011) have demonstrated that modality-specific attention training can lead to generalized learning in alternative modalities in healthy older adults. These authors and others point to video gaming as ideal paradigms which involve a rich blend of complex strategies and processes (Lustig, Shah, Seidler, & Reuter-Lorenze, 2009).

In a recent demonstration of video game training in older adults, Stern, Blumen, Rich, Richards, Herzberg, & Gopher (2012) report limited effectiveness in using the Space Fortress game to improve executive control in older adults. However, Basak, Boot, Voss, and Kramer, (2008) have demonstrated improvements in task switching resulting from RTS game training in older adults. This discrepancy, along with the present study's indication the importance of gaming attributes, highlight the importance of understanding the role of video game attributes and video game behavior attributes on cognitive change.

Understanding the psychological impact of video game play is important for several reasons. First, video gaming is an increasingly popular past time, and so it critical to understand the psychological impact of chronic gaming. Second, complex and immersive decision making tasks are historically difficult to research due to the intractable nature of task and behavior features. Video games are invaluable tools in the research of complex decision making due to their high levels of both immersion and complexity despite being capable of high levels of experimental control. Finally, recent advances in cognitive training research point to complex decision making environments as effective yet highly engaging tools that lead to levels of generalizability and far transfer effects not comparable with specific and repetitive single-task training paradigms. The advances put forth in the current research initiative demonstrate the ability to control aspects of the gaming environment which lead to differential cognitive enhancements. Future work should build on the aforementioned opportunities for feature analysis and behavior extraction techniques to further the understanding of human behavior in complex and immersive decision making environments.

APPENDIX: EZ-Diffusion Parameters and Assumptions

ALTERNATIVE EZ-DIFFUSION PARAMETERS

Drift rate is the primary measure of interest in speeded reaction time tasks, but the boundary separation parameter and non-decision time parameters are also available and important. Figures 9 through 11 and Tables 11 and 12 report the *t*-values and meta-analytic Bayes factors for boundary separation and non-decision time results in both the Executive tasks and Other tasks. Higher boundary separation values indicate more conservative responses. While this typically leads to longer response time, when paired with higher drift rate parameters, higher accuracy, and lower reaction times, increased boundary separation does not necessarily indicate a speed-accuracy tradeoff. Instead, it is indicative of a cleaner diffusion process with fewer errant responses that occur due to chance fluctuations.

EZ-DIFFUSION MODELING ASSUMPTION VERIFICATION

Attentional Network Test (ANT)

First, a binomial test determined the frequency of the two response key choices were best fit by a Bernoulli trial with a probability of 0.507 with 95% CIs [0.500, 0.514]. This is sufficiently fair to satisfy the assumptions of the EZ-Diffusion model. Second, D'Agostino's K-squared test confirmed that the overall RT distribution was non-normal, $X^2(2) = 7597.5$, $p < 0.0001$. This satisfies the condition of skewness for the EZ-Diffusion model. Third, the Kolmogorov-Smirnov test rejected the hypothesis that the aggregate RT distributions for correct and error responses were from the same distribution, $K = 0.39$, $p < 0.001$. This violates one assumption of the EZ-Diffusion model, but as previously discussed the EZ-Diffusion model's parameters remains ordinally intact despite this misspecification. Furthermore, when tested on a participant-level basis, the

Kolmogorov-Smirnov test was unable to reject the hypothesis that the RT distributions for correct and error responses were from different distributions, mean $K = 0.68$, $p = 0.16$.

Task Switching

First, a binomial test determined the frequency of the two response key choices were best fit by a Bernoulli trial with a probability of 0.495 with 95% CIs [0.491, 0.499]. This is sufficiently fair to satisfy the assumptions of the EZ-Diffusion model. Second, D'Agostino's K-squared test confirmed that the overall RT distribution was non-normal, $X^2(2) = 42916$, $p < 0.0001$. This satisfies the condition of skewness for the EZ-Diffusion model. Third, the Kolmogorov-Smirnov test rejected the hypothesis that the aggregate RT distributions for correct and error responses were from the same distribution, $K = 0.09$, $p < 0.001$. This violates one assumption of the EZ-Diffusion model, but as previously discussed the EZ-Diffusion model's parameters remains ordinally intact despite this misspecification. Furthermore, when tested on a participant-level basis, the Kolmogorov-Smirnov test was unable to reject the hypothesis that the RT distributions for correct and error responses were from different distributions, mean $K = 0.33$, $p = 0.38$.

Multi-Location Memory Task

First, a binomial test determined the frequency of the two response key choices were best fit by a Bernoulli trial with a probability of 0.492 with 95% CIs [0.487, 0.499]. This is sufficiently fair to satisfy the assumptions of the EZ-Diffusion model. Second, D'Agostino's K-squared test confirmed that the overall RT distribution was non-normal, $X^2(2) = 12270$, $p < 0.0001$. This satisfies the condition of skewness for the EZ-Diffusion model. Third, the Kolmogorov-Smirnov test rejected the hypothesis that the aggregate

RT distributions for correct and error responses were from the same distribution, $K = 0.07$, $p < 0.001$. This violates one assumption of the EZ-Diffusion model, but as previously discussed the EZ-Diffusion model's parameters remains ordinally intact despite this misspecification. Furthermore, when tested on a participant-level basis, the Kolmogorov-Smirnov test was unable to reject the hypothesis that the RT distributions for correct and error responses were from different distributions, mean $K = 0.35$, $p = 0.38$.

Visual Search

First, a binomial test determined the frequency of the two response key choices were best fit by a Bernoulli trial with a probability of 0.522 with 95% CIs [0.517, 0.526]. This is sufficiently fair to satisfy the assumptions of the EZ-Diffusion model. Second, D'Agostino's K-squared test confirmed that the overall RT distribution was non-normal, $X^2(2) = 24370$, $p < 0.0001$. This satisfies the condition of skewness for the EZ-Diffusion model. Third, the Kolmogorov-Smirnov test rejected the hypothesis that the aggregate RT distributions for correct and error responses were from the same distribution, $K = 0.14$, $p < 0.001$. This violates one assumption of the EZ-Diffusion model, but as previously discussed the EZ-Diffusion model's parameters remains ordinally intact despite this misspecification. Furthermore, when tested on a participant-level basis, the Kolmogorov-Smirnov test was unable to reject the hypothesis that the RT distributions for correct and error responses were from different distributions, mean $K = 0.40$, $p = 0.43$.

Information Filtering

First, a binomial test determined the frequency of the two response key choices were best fit by a Bernoulli trial with a probability of 0.507 with 95% CIs [0.499, 0.514].

This is sufficiently fair to satisfy the assumptions of the EZ-Diffusion model. Second, D'Agostino's K-squared test confirmed that the overall RT distribution was non-normal, $X^2(2) = 7560$, $p < 0.0001$. This satisfies the condition of skewness for the EZ-Diffusion model. Third, the Kolmogorov-Smirnov test rejected the hypothesis that the aggregate RT distributions for correct and error responses were from the same distribution, $K = 0.39$, $p < 0.001$. This violates one assumption of the EZ-Diffusion model, but as previously discussed the EZ-Diffusion model's parameters remains ordinally intact despite this misspecification. Furthermore, when tested on a participant-level basis, the Kolmogorov-Smirnov test was unable to reject the hypothesis that the RT distributions for correct and error responses were from different distributions, mean $K = 0.38$, $p = 0.27$.

	Mid-test minus Pre-test (0 to 20hrs)	Post-test minus Mid-test (20 to 40hrs)	Post-test minus Pre-Test (0 to 40hrs)
SC-LO Executive Tasks	1.45 0.13º	0.14º 0.12º	**10.71**** 0.12º
SC-HI Executive Tasks	1.61 0.07ºº	0.21º 0.55	**830.38**** 0.30º
SC-BOTH Executive Tasks	8.31 0.07ºº	0.24º 0.28º	**40339.60**** 0.32º
SC-LO Other Tasks	0.61 0.12º	0.05ºº 0.04ºº	0.11º 0.03ºº
SC-HI Other Tasks	0.19º 0.06ºº	0.06ºº 0.04ºº	0.12º 0.04ºº
SC-BOTH Other Tasks	0.55 0.06ºº	0.05ºº 0.03ºº	0.09ºº 0.02ºº

*BF > 3, substantial evidence for alternative hypothesis
**BF > 10, strong evidence for alternative hypothesis
ºBF < $0.\overline{3}$, substantial evidence for null hypothesis
ººBF < 0.1, strong evidence for null hypothesis

Table 11: Meta-analytic Bayes Factors for Executive Task group and Other Task group, for SC-LO, SC-HI, and SC-BOTH (combined) on mid-test versus pre-test, post-test minus pre-test, and post-test minus mid-test, versus the Sims control group, with boundary separation parameter when EZ-Diffusion modeling was possible

	Mid-test minus Pre-test (0 to 20hrs)	Post-test minus Mid-test (20 to 40hrs)	Post-test minus Pre-Test (0 to 40hrs)
SC-HI vs. SC-LO Executive Tasks	0.08°° 0.21°	0.06°° 0.03°°	0.04°° 0.07°°
SC-HI vs. SC-LO Executive Tasks	0.23° 0.19°	0.06°° 0.04°°	0.08°° 0.08°°

°BF < $0.\bar{3}$, substantial evidence for null hypothesis
°°BF < 0.1, strong evidence for null hypothesis

Table 12: Meta-analytic Bayes Factors for Executive Task group and Other Task group, for SC-HI versus SC-LO on mid-test versus pre-test, post-test minus pre-test, and post-test minus mid-test, with boundary separation parameter when EZ-Diffusion modeling was possible

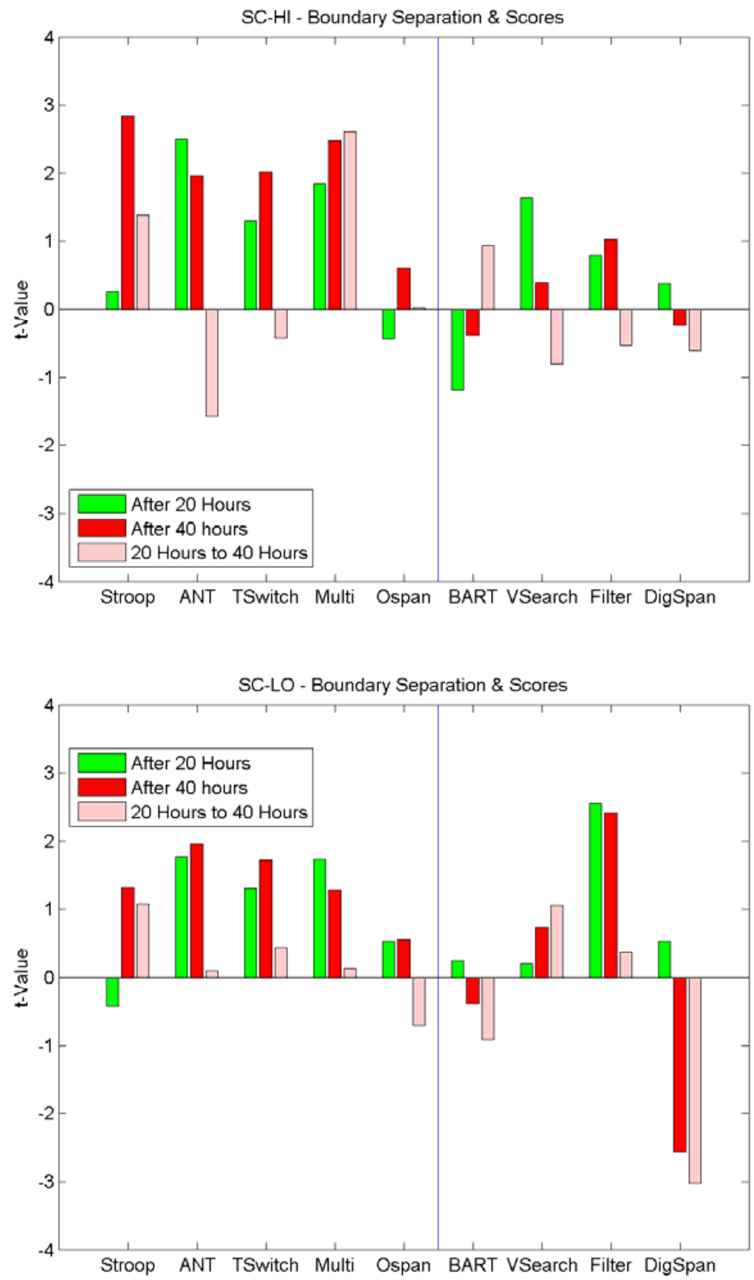

Figure 9: *t*-values for mid-test (after 20 hours) and post-test (after 40 hours) for boundary separation parameter (for ANT, Task Switching, Multi-location Memory, Visual Search and Information Filtering) and scores (for Stroop, Ospan, BART, and Digit Span). Higher boundary separation values represent less caution.

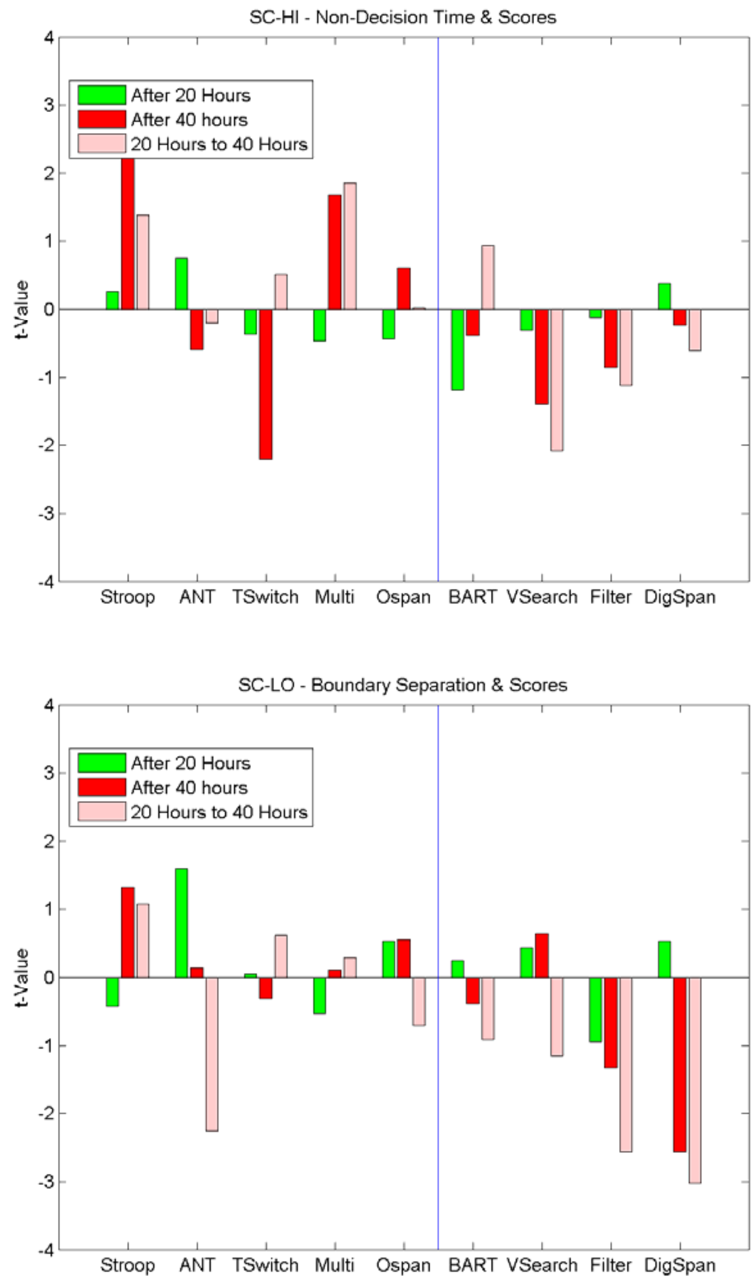

Figure 10: *t*-values for mid-test (after 20 hours) and post-test (after 40 hours) for non-decision time parameter (for ANT, Task Switching, Multi-location Memory, Visual Search and Information Filtering) and scores (for Stroop, Ospan, BART, and Digit Span)

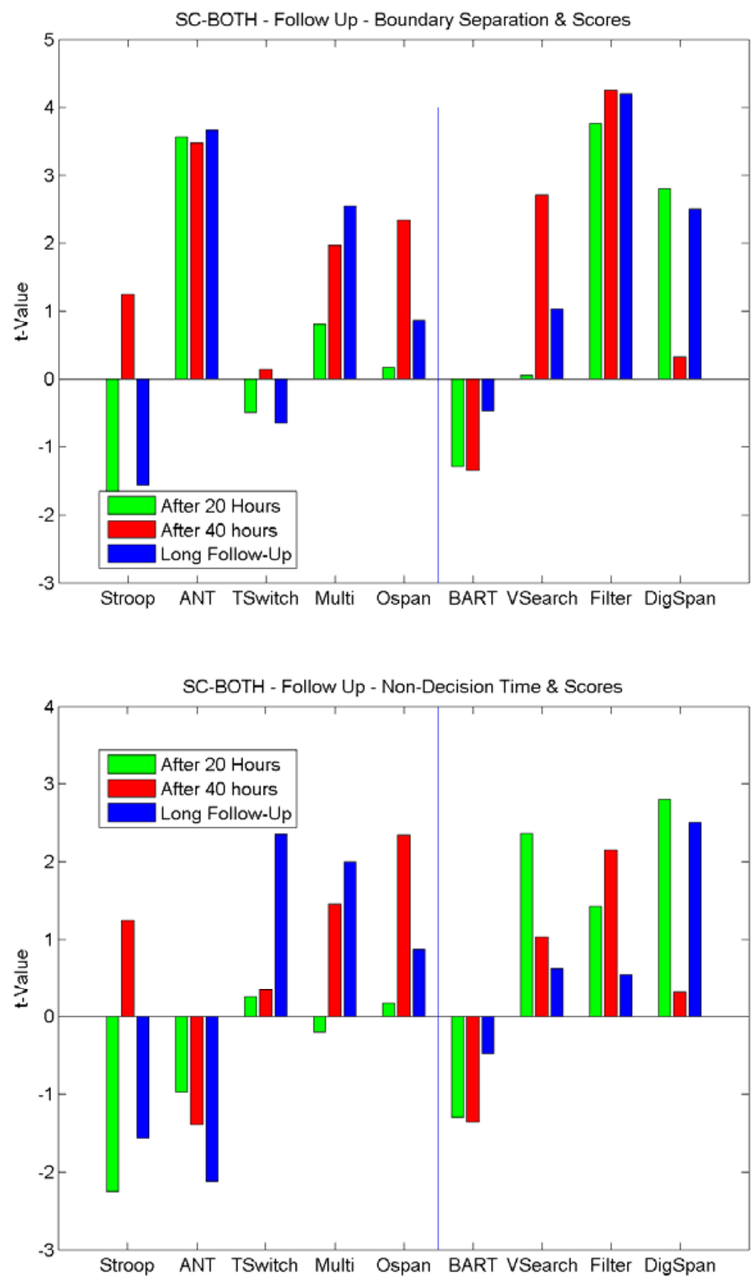

Figure 11: *t*-values for mid-test (after 20 hours), post-test (after 40 hours), and follow-up session for boundary separation and non-decision time parameter (for ANT, Task Switching, Multi-location Memory, Visual Search and Information Filtering) and scores (for Stroop, Ospan, BART, and Digit Span)

References

Aitkin, M. (1991). Posterior Bayes factors. *Journal of the Royal Statistical Society. Series B*, *53*, 111–142.

Atallah, H. E., Lopez-Paniagua, D., Rudy, J. W., & O'Reilly, R. C. (2007). Separate neural substrates for skill learning and performance in the ventral and dorsal striatum. *Nature Neuroscience, 10,* 126-131.

Bailey, K., West, R., & Anderson, C.A. (2010). A negative association between video game experience and proactive cognitive control. *Psychophysiology, 47,* 34-42.

Bartels, C., Wegrzyn, M., Wiedl, A., Ackermann, V., & Ehrenreich, H. (2010). Practice effects in healthy adults: a longitudinal study on frequent repetitive cognitive testing. *BMC Neuroscience*, *11*, 118.

Basak, C., Boot, W. R., Voss, M. W., & Kramer, A. F. (2008). Can training in a real-time strategy video game attenuate cognitive decline in older adults? *Psychology and Aging, 23, 756-777.*

Berger, J. O., & Pericchi, L. R. (1996). The intrinsic Bayes factor for model selection and prediction. *Journal of the American Statistical Association*, *91*, 109-122.

Bherer L., Kramer A. F., Peterson M. S., Colcombe S., Erickson K., & Becic E. (2008). Transfer effects in task-set cost and dual-task cost after dual-task training in older and younger adults: further evidence for cognitive plasticity in attentional control in late adulthood. *Experimental Aging Research*, *34*, 188–219.

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.

Boot, W.R., Basak, C., Erickson, K.I., Neider, M., Simons, D.J., Fabiani, M., Gratton, G., et al. (2010). Transfer of skill engendered by complex task training under conditions of variable priority. *Acta Psychologica, 135*, 349-357.

Braver, T. S., Barch, D. M., Keys, B. A., Carter, C. S., Cohen, J. D., Kaye, J. A., et al. (2001). Context processing in older adults: Evidence for a theory relating cognitive control to neurobiology in healthy aging. Journal of Experimental Psychology: General, 130, 746–763.

Botvinick, M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*, 624–652.

Burns, B. D., & Vollmeyer, R. (2002). Goal specificity effects on hypothesis testing in problem solving. *Quarterly Journal of Experimental Psychology*, *55A*, 241–261.

Bush, G., Frazier, J. A., Rausch, S. L., Seidman, L. J., Whalen, P. J., Jenike, M. A., Rosen, B. R., & Biederman, J. (1999). Anterior cingulated cortex dysfunction in attention deficit/hyperactivity disorder revealed by fMRI and the counting Stroop. *Biological Psychiatry*, *45*, 1542-1552.

Brewin, C. R., & Beaton, A. (2002). Thought suppression, intelligence, and working memory capacity. *Behavior Research and Therapy*, *40*, 923–930.

Carter, C. S., Mintun, M., Nichols, T., & Cohen, J. D. (1997). Anterior cingulate gyrus dysfunction and selective attention deficits in schizophrenia: [-sup-1-sup-5O]H-sub-20 PET study during single trial Stroop task performance. *American Journal of Psychiatry*, *154*, 1670–1675.

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*, 217–230.

Castellanos, F. X., Sonuga-Barke, E. J. S., Milham, M. P., & Tannock, R. (2006). Characterizing cognition in ADHD: beyond executive dysfunction. *TRENDS in Cognitive Sciences*, *10*, 117-123.

Charron, S. & Koechlin, E. (2010). Divided representation of concurrent goals in the human frontal lobes. *Science*, *328*, 360–363.

Chen, J.M., & Morrison, A.B. (2010). Expanding the mind's workspace: Training and transfer effects with a complex working memory span task. *Psychonomic Bulletin & Review, 17*, 193-199.

Collie, A., Darby, D. G., & Maruff, P. (2003). The effects of practice on the cognitive test performance of neurologically normal individuals assessed at brief test-retest intervals. *Journal of International Neuropsychological Society,* 9, 419-428.

Colzato, L.S., van den Wildenberg, W.P.M., Zmigrod, S., & Hommel, B. (2012). Action video gaming and cognitive control: Playing first person shooter games is associated with improvement in working memory but not action inhibition. *Psychological Research,*

D'Agostino, R. B. (1970). Transformation to normality of the null distribution of g_1. *Biometrika*, *57*, 679-681.

Dahlin, E., Nyberg, L., Backman, L., & Neely, A.S. (2008). Plasticity of executive functioning in young and older adults: Immediate training gains, transfer, and long-term maintenance. *Psychology and Aging, 23*, 720-730.

Donohue, S. E., Woldorff, M. G., & Mitroff, S. R. (2010). Video game players show more precise multisensory temporal processing abilities. *Attention, Perception, & Psychophysics, 72*, 1120-1129.

Dye, M. W. G., Green, C. S., Bavelier, D. (2009a). Increasing speed of processing with action video games. *Current Directions in Psychological Science, 18*, 321-326.

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

Durston, S. (2008). Converging methods in studying attention-deficit/hyperactivity disorder: What can we learn from neuroimaging and genetics? *Development and Psychopathology*, *20*, 1133-1143.

Erickson, K.I., Boot, W.R., Basak, C., Neider, M.B., Prakash, R.S., Voss, M.W., Graybiel, A.M., et al. (2010). Striatal volume predicts level of video game skill acquisition. *Cerebral Cortex, 20*, 2522-2530.

Falleti, M. G., Maruff, P., Collie, A., & Darby, D. G. (2006). Practice effects associated with the repeated assessment of cognitive function using the CogState battery at 10-minute, one week and one month test-retest intervals. *Journal of Clinical and Experimental Neuropsychology*, *28,* 1095-1112.

Fan, J., McCandliss, B. D., Fossella, J., Flombaum, J. I., Posner, M. I. (2005). The activation of attentional networks. *NeuroImage*, *26*, 471-479.

Fan, J., McCandliss, B. D., Sommer, T., Raz, A. & Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience*. *14*, 340–347.

Feng, J., Spence, I., & Pratt, J. (2007). Playing an action video game reduces gender differences in spatial cognition. *Psychological Science, 18*, 850-855.

Goldstone, R.L., & Sakamoto, Y. (2003). The transfer of abstract principles governing complex adaptive systems. *Cognitive Psychology, 46*, 414-466.

Good, I.J. (1992). The Bayes/Non-Bayes compromise: A brief review. *Journal of the American Statistical Association*, *87*, 597-606.

Goodman, S.N. (1999) Toward evidence-based medical statistics. 2: The Bayes factor. *Annals of International Medicine*, *130*, 1005–101.

Granek, J. A., & Gorbet, D. J., & Sergio, L.E. (2010). Extensive video-game experience alters cortical networks for complex visuomotor transformations. *Cortex*, *46*, 2010.

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*, 1465–1478.

Green, C.S. & Bavelier, D. (2007). Action video game experience alters the spatial resolution of attention. *Psychological Science*, 18, 88-94.

Green, C. S., & Bavelier, D. (2008). Exercising your brain: A review of human brain plasticity and training-induced learning. *Psychology and Aging*, *23*, 692–701.

Green, C. S., Pouget, A., & Bavelier, D. (2010). Improved probabilistic inference, as a general learning mechanism with action video games. *Current Biology*, 20, 1573-15792.

Hausknecht, J.P., Halpert, J.A., Di Paolo, N.T., & Moriarty Gerrard, M.O. (2007). Retesting in selection: A meta-analysis of practice effects for tests of cognitive ability. *Journal of Applied Psychology, 92*, 373-385.

Halperin, J. M., & Healey, D. M. (2010). The influences of environmental enrichment, cognitive enhancement, and physical exercise on brain development: Can we alter the developmental trajectory of ADHD? *Neuroscience and Biobehavioral Reviews*, *35*, 621–634.

Hertzog, C., Kramer, A. F., Wilson, R. S., & Lindenberger, U. (2009). Enrichment effects on adult cognitive development: Can the functional capacity of older adults be preserved and enhanced? *Psychological Science*, *9*, 1-64.

Howieson, Diane Black; Lezak, Muriel Deutsch; Loring, David W. (2004). "Orientation and attention". Neuropsychological assessment. Oxford [Oxfordshire]: Oxford University Press. pp. 3365–367.

Hunt, M. K., Hopko, D. R., Bare, R., Lejuez, C. W., & Robinson, E. (2005). Construct Validity of the Balloon Analog Risk Task (BART): Associations With Psychopathy and Impulsivity. *Assessment*, *12*, 416-428.

Jaeggi, S.M., Studer-Luethi, B., Buschkuehl, M., Su, Y-F., Jonides, J., & Perrig, W.J. (2010). The relationship between n-back performance and matrix reasoning – implications for training and transfer. *Intelligence, 38*, 625-635.

Kalyuga, S. (2008). Instructional designs for the development of transferable knowledge and skills: A cognitive load perspective. *Computers in Human Behavior, 25*, 332-338.

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General*, *132*, 47-70.

Karbach, J., & Kay, J. (2009). How useful is executive control training? Age differences in near and far transfer of task-switching training. *Developmental Science, 12*, 978-990.

Karni, A., & Bertini, G. (1997). Learning perceptual skills: Behavioral probes into adult cortical plasticity. *Current Opinion in Neurobiology, 7*, 530-535.

Kelley, T.A., & Yantis, S. (2009). Learning to attend: Effects of practice on information selection. *Journal of Vision, 9*, 1-18.

Kessler, R.C., Adler, L., Ames, M., Demler, O., Faraone, S., Hiripi, E., Howes, M.J., et al. (2005). The world health organization adult ADHD self-report scale (ASRS): A short screening scale for use in the general population. *Psychological Medicine, 35*, 245-256.

Jeffreys, H. (1961). *Theory of probability (3rd ed.).* Oxford: Oxford University Press, Clarendon Press.

Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L., Strong, D. R., & Brown, R. A. (2002). Evaluation of a behavioral measure of risk-taking: The Balloon Analogue Risk Task (BART). *Journal of Experimental Psychology: Applied, 8*, 75-84.

Lenhart, A., Kahne, J., Middaugh, E., Macgill, A.R., Evans, C., & Vitak, J. (2008). Teens, video games, and civics: Teens' gaming experiences are diverse and include significant social interaction and civic engagement, Pew Internet & American Life Project. http://www.pewinternet.org

Li R., Polat U., Makous W., & Bavelier D.(2009). Enhancing the contrast sensitivity function through action video game training. *Nature Neuroscience*, *12*, 549–551.

Love, B. C., Jones, M., Tomlinson, M.T., & Howe, M. (2008). Predicting Information Needs: Adaptive Display in Dynamic Environments . *Proceedings of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum Associates.

Lustig, C., Shah, P. Seidler, R., & Reuter-Lorenz, P. A. (2009). Aging, training, and the brain: A review and future directions. *Neuropsychological Review*, *19*, 504-522.

Maclin, E.L., Mathewson, K.E., Low, K.A., Boot, W.R., Kramer, A.F., Fabiani, M., & Gratton, G. (2011). Learning to multitask: Effects of video game practice on electrophysiological indices of attention and resource allocation. *Psychophysiology, 48*, 1173-1183.

Macmillan, NA.; Creelman, CD. Detection theory: A user's guide. New York: Cambridge University Press; 2004.

Masson, M.J., Bub, D.N., & Lalonde, C.E. (2011). Video-game training and naïve reasoning about object motion. *Applied Cognitive Psychology, 25*, 166-173.

McCaffrey, R. J., Ortega, A., & Haase, R. F. (1993). Effects of repeated neuropsychological assessments. *Archives of Clinical Neuropsychology, 8*, 519-524.

Mirsky, A.F., Anthony, B.F., Duncan, C.C., Ahearn, M.B., & Kellam, S.G. (1991). Analysis of the elements of attention: A neuropsychological approach. *Neuropsychological Review,* 2, 109-145.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., & Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49-100.

Monsell, S. (2003) Task switching. *Trends in Cognitive Science*, *7*, 134–140.

Mozolic, J. L., Long, A. B., Morgan, A. R., Rawley-Payne, M., & Laurienti, P. J. (2011). A cognitive training intervention improves modality-specific attention in a randomized controlled trial of healthy older adults. *Neurobiology of Aging, 32*, 655-668.

Olesen, P.J., Westerberg, H., & Klingberg, T. (2004). Increased prefrontal and parietal activity after training of working memory. *Nature Neuroscience*, *7*, 75–79.

Ophir, E., Nass, C., & Wagner, A. D. (2009). Cognitive control in media multitaskers. *Proceedings of the National Academy of Sciences*, *106*, 15583–15587.

Osman, M. (2010). Controlling uncertainty: A review of human behavior in complex dynamic environments, *Psychological Bulletin, 136*, 65-86.

Pashler, H. (2000). Task switching and multitask performance. In Monsell, S., and Driver, J. (editors). *Attention and Performance* XVIII: Control of mental processes. Cambridge, MA: MIT Press.

Posner, M.I., & Rothbart, M.K. (2007). Research on attention networks as a model for the integration of psychological science. *Annual Review of Psychology*, *58*, 1–23.

Posner, M. I., Sheese, B. E., Odludas, Y., & Tong, Y. (2006). Analyzing and shaping human attentional networks. *Neural Network*, 19, 1422–1429.

Prins, P.J.M., Dovis, S., Ponsioen, A., Brink, E., & van der Oord, S. (2011). Does computerized working memory training with game elements enhance motivation and training efficacy in children with ADHD? *Cyberpsychology, Behavior, and Social Networking, 14*, 115-122.

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for twochoice decisions. *Psychological Science*, *9*, 347-356.

Raz, A., & Buhle, J. (2006). Typologies of attentional networks. *Nature Reviews Neuroscience*, *7*, 367-379.

Rideout, V. J., Roberts, D. F., & Foehr, U. G. (2005). *Generation M: Media in the lives of 8-18 year-olds: Executive summary*. Menlo Park, CA: Henry J. Kaiser Family Foundation.

Rogers, R., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, *124*, 207-231

Rosser, J.C., Lynch, P.J., Cuddihy, L., Gentile, D.A., Klonsky, J., & Merrell, R. (2007). The impact of video games on training surgeons in the 21st century. *Archives of Surgery, 142*, 181-186.

Rouder, J. N., & Morey, R. D. (2011). A Bayes-factor meta analysis of Bem's ESP claim. *Psychonomic Bulletin & Review*, *18*, 682–689.

Rubinstein, J. S., Meyer, D. E., & Evans, J. E. (2001). Executive control of cognitive processes in task switching. *Journal of Experimental Psychology: Human Perception and Performance, 27*, 763-797.

Rueda, M., Rothbart, M., McAndliss, B., Saccomanno, L., & Posner, M. (2005). Training, maturation, and genetic influences on the development of executive attention. *Proceedings from the National Academy of Sciences*, *102*, 14931–14936.

Sanchez, C.A. (2012). Enhancing visuospatial performance through video game training to increase learning in visuospatial science domains. *Psychonomic Bulletin & Review, 19*, 58-65.

Schmidt, N. B., Richey, J. A., Buckner, J. D., Timpano, K. R. (2009).Attention training for generalized social anxiety disorder. *Journal of Abnormal Psychology*, *118*, 5–14.

Seidler, R.D. (2004). Multiple motor learning experiences enhance motor ability. *Journal of Cognitive Neuroscience, 16*, 65-73.

Sloutsky, V.M., Kaminski, J.A., & Heckler, A.F. (2005). The advantage of simple symbols for learning and transfer. *Psychonomic Bulletin & Review, 12*, 508-513.

Stern,Y., Blumen, H. M., Rich, L. W., Richards, A., Herzberg, G., & Gopher, D. (2011). Space Fortress game training and executive control in older adults: A pilot intervention. *Aging, Neuropsychology, & Cognition*, *18*, 653-677.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. Journal of *Experimental Psychology*, *18*, 643-662.

Subrahmanyam, K., & Greenfield, P.M. (1994). Effect of video game practice on spatial skills in girls and boys. *Journal of Applied Developmental Psychology 15*, 13-32.

Sungur, H., & Boduroglu, A. (2012). Action video game players form more detailed representation of objects. Acta Psychologica, 139, 327-334.

Sweller, J. (1989). Cognitive technology: Some procedures for facilitating learning and problem solving in mathematics and science. *Journal of Educational Psychology*, *81*, 457-466.

Swing, E. L. (2008). *Attention abilities, media exposure, school performance, personality, and aggression*. Unpublished master's thesis, Iowa State University, Ames, Iowa.

Tang Y. Y. & Posner, M. I. (2009) Attention training and attention state training. *Trends in Cognitive Science*, *13*, 222–227.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory & Language*, *28*, 127-154.

Unsworth, N., Heitz, R. P., Shrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods, Instruments, & Computers*, 37, 498–505.

Vandekerckhove, J., & Tuerlinckx, F. (2007). Fitting the Ratcliff diffusion model to experimental data. *Psychonomic Bulletin & Review*, *14*, 1011-1026.

van Merrienboer, J.J.G., Schuurman, J.G., de Croock, M.B.M., & Paas, F.G.W.C. (2002). Redirecting learners' attention during training: Effects on cognitive load, transfer test performance and training efficiency. *Learning and Instruction, 12*, 11-37.

Van Rullen, R., & Thorpe, S. J. (2001). The time course of visual processing: From early perception to decision-making. *Journal of Cognitive Neuroscience*, *13*, 454-461.

Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. Nature, *438*, 500 -503.

Verhaeghen, P., Cerella, J., & Basak, C. (2004). A working memory workout: How to expand the focus of serial attention from one to four items in 10 hours or less. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 1322–1337.Vogel E-K, McCollough A-W,Machizawa M-G (2005) Neural measures reveal individual differences in controlling access to working memory. *Nature*, *438*, 500–503.

Wasserman L. (2000). Bayesian model selection and model averaging. *Journal of Mathematical Psychology*, *44*, 92–107.

Wagenmakers, E.-J., van der Maas, H., & Grasman, R. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review*, *14*, 3-22.

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale—4th Edition (WAIS-IV®)*. San Antonio, TX: Harcourt Assessment.

Wesnes K, & Pincock C. (2002). Practice effects on cognitive tasks: a major problem? *The Lancet Neurology*, 1-473.

Wolfe, J. M. (2003). Moving towards solutions to some enduring controversies in visual search. *Trends in Cognitive Sciences*, *7*, 70 –76.