2020

Induction and Transferral of Flow in the Game Tetris

Kevin John O'Neill
Wright State University

INDUCTION AND TRANSFERRAL OF FLOW IN THE GAME TETRIS

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

By

KEVIN JOHN O’NEILL

B.S., Rensselaer Polytechnic Institute, 2016

2020

Wright State University

WRIGHT STATE UNIVERSITY

GRADUATE SCHOOL

AUGUST 21, 2020

I HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER MY SUPERVISION BY Kevin John O'Neill ENTITLED Induction and Transferral of Flow in the game Tetris BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF Master of Science

____
.Ion Juvina, Ph.D.
Thesis Director

____
David LaHuis, Ph.D.
Graduate Program Director

____
Debra Steele-Johnson, Ph.D.
Chair, Department of Psychology

Committee on Final Examination

____
Pamela Tsang, Ph.D.

____
Joseph Houpt, Ph.D

____
Barry Milligan, Ph.D.
Interim Dean of the Graduate School

ABSTRACT

O'Neill, Kevin John. M.S. Department of Psychology, Wright State University, 2020. Induction and Transferral of Flow in the Game Tetris.

We looked at the facilitation and transfer of a flow state in a cognitive context. Subjects played a manipulated version of the game Tetris, and we gathered data on their gameplay performance on pre- and post-tasks, as well as a set of questionnaires which measure flow and perceived task effort. The altered version of Tetris includes an artificial intelligence agent that continually assesses the participant's skill and adapts the challenge level of the game to match the participant's skill. An adaptive condition characterized by challenge-skill balance was hypothesized to induce flow, reduce perception of effort, and improve performance. We found differences in reported flow state between conditions, with the easy condition inducing greater flow than adaptive condition, which induced greater flow than the hard condition. We did not find significant differences for performance measures.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I'd like to acknowledge my advisor, Ion Juvina, and my lab-mates, Michael Collins, Alex Hough, Peter Crowe, and Jarean Carson, for their advice, suggestions, and willingness to help pilot test the experiment software.

Introduction

Flow State is a state of high concentration and focus that people often enter into when fully absorbed in a task, such as reading, writing, or programming. People in flow generally tend to lose a degree of awareness of their surroundings and of the passage of time, and tend to do tasks longer without experiencing conscious fatigue. Flow is also intrinsically motivating, as people who experience flow when doing a task usually want to continue doing that task.

Flow as a concept first arose through study of tasks which are intrinsically motivating, and the types of personalities that are most likely to enter into a flow state (Czikszentmihalyi, 1990). Csikszentmihalyi termed this tendency towards flow the autotelic personality, giving a distinction between flow as a personality trait and the flow state. Trait flow is a characteristic of an individual person, much like any other personality trait, whereas state flow is a more short term experience, affected by the current task and environment, as well as one's own dispositional trait flow (Nakamura & Csikszentmihalyi, 2009).

Necessary characteristics of a task which would induce a flow state include a balance between perceived challenges and one's skills to deal with those challenges, as well as clear, proximal goals and immediate feedback about performance on the task (Nakamura & Csikszentmihalyi, 2009).

The subjective state of flow exhibits characteristics of intense concentration, a merging of action and awareness, a loss of reflective self-consciousness, a sense of being in control, a sense that time has passed faster than normal, and an experience of the activity as being intrinsically rewarding (Nakamura & Csikszentmihalyi, 2009).

Flow state has traditionally been measured via interview, where subjects are asked to qualitatively describe their flow experiences (Jackson, 1995), or via the Experience Sampling

Method, where subjects are prompted at random times during the day via a pager or similar device to record their current subjective experience in a journal (Csikszentmihalyi & Larson, 2014).

A third method of flow measurement is through questionnaire, which offers a more quantitative description of flow state. Among the most well used are the Flow State Scale (FSS) and Dispositional Flow Scale (DFS) (Jackson & Marsh, 1996), and their updated versions, the FSS2 and DFS2 (Jackson & Eklund, 2002), as well as short versions of the same scales (Jackson, Martin, & Eklund, 2008). The FSS and related versions measure the degree to which one experienced flow state in an activity, while the DFS measures how one's traits and personality predispose them to enter into flow states, something akin to Csikszentmihalyi's ideas of an autotelic personality (Nakamura & Csikszentmihalyi, 2009).

Flow state has been described in the past as a powerful and beneficial experience. People in flow can do tasks for hours without losing concentration or experiencing the feeling of fatigue. If this experience could be induced and transferred to existing tasks which require high concentration for extended periods of time, the benefits would be immediate.

Much of the previous research in the area of flow has focused on flow from a passive and distant perspective. The work of Csikszentmihalyi and his collaborators largely focused on identifying and describing flow, and understanding the attributes of tasks which are more likely to induce flow, as well as the personalities of people who find themselves in a flow state more often. This is a necessary and important step in the transition of a concept understood from a lay, "common-sense" perspective to one described rigorously and empirically. However, the work is not complete, and these foundations need to be built upon.

Relatively few threads of research have explored the concept of flow from a more controlled, cognitive, and attentional perspective. In conducting this research, I hoped to add to this more internal and basic understanding of the psychological phenomena that make up flow.

One of the questions concerning flow which has yet to be fully explored is that of how flow state may arise during a given task. Czikszentmihalyi identified nine dimensions of flow which contribute to the flow experience, and which can be manipulated through the design of a task in order to better facilitate the flow experience of someone performing the task. Not all tasks can be manipulated in such a way; some tasks are restricted in the changes which can be made to them, either due to the high cost or difficulty of changing core components of a task, or because the intended outcome precludes any further changes in that direction. In such scenarios, it would be useful if changes could be made to unrelated tasks which might be performed in close temporal proximity to, or even concurrent with, the desired task, which might allow some of the benefits of a flow state to be transferred from an unrelated task to the desired one.

This concept of transfer of beneficial effects is similar to, and indeed takes direct inspiration from, the concept of transfer of learning. The idea of learning itself is fairly intuitive; generally, the more one does a certain task, the better one gets at that task. Transfer of learning occurs when knowledge and skills obtained in learning one task are applied to a different task. In literature this is usually split into near transfer and far transfer, where near transfer refers to transfer of learning between similar tasks, while far transfer describes transfer of learning between dissimilar tasks. The current conclusions of the field are that, when transfer does appear, far transfer occurs much more rarely than near transfer (Perkins & Saloman, 1992).

For example, driving a car and driving a truck are similar tasks; many of the same skills are required for both, and so skills gained in learning one task will likely be useful in learning the other task, facilitating near transfer. However, driving a car and piloting an airplane are mostly dissimilar tasks. Despite this, there might be some skills that can be acquired from learning to drive a car that could be of use in piloting an aircraft. These skills will likely not be numerous or easily applied, but if experience driving a car were to improve performance in piloting an airplane, that would be categorized as far transfer.

The comparison of learning transfer to flow transfer is not perfect; flow is better described as a state of mind or a mood than as a skill or type of knowledge. However, it doesn't seem unreasonable that an analogous effect might exist, one which may also show a distinction between near and far transfer. Previous research has shown that moods persisting from previous situations can affect future behaviour (Hills, Hill, Mamone, & Dickerson, 2001), so it seems possible that flow may also persist, in a manner analogous to learning transfer.

This idea of transfer of flow constitutes the core research questions of this work: "how can flow be effectively induced, and is it possible for manipulations of a flow-inducing task to affect performance on a task which has not been altered to induce flow?"

In order to test these questions, we needed to identify a task which could be manipulated to be more or less flow inducing, as well as a "baseline" task which could be used to test the effect of any transfer of flow. We chose the video game Tetris as the task which could be manipulated to affect flow state, and the unaltered version of the same game as the baseline task.

Tetris is one of the most popular video games of all time, and has an existing body of literature in several disciplines which have explored its aspects. The game involves rotating and translating a falling tetromino, to place it either on the bottom of a 2d game board, or on top of previously placed zoids on that game board. In existing Tetris research, these tetrominoes (also called "Tetris pieces") are termed "zoids" (see Fig. 1), and the pile of previously placed zoids at the bottom of the game board is called the “accumulation.” The player’s goal is to arrange the zoids such that they form one or more completely filled rows in the accumulation, at which point the row or rows will disappear, the pieces above the row or rows will descend to fill the gap, and the player’s score will increase. This sequence of zoid placements (each called “episodes” in the literature) continues indefinitely until the accumulation reaches the top of the game board. At that point, there is no room to place a new piece, and the player will get a “game over” message and be prompted to start a new game. The speed at which the zoids fall increases over the course of a game, as this speed is proportional to the number of rows cleared thus far. This fact all but assures that the player will lose, since the zoids will eventually fall faster than any human’s possible reaction time, and end up in places not intended by the player.

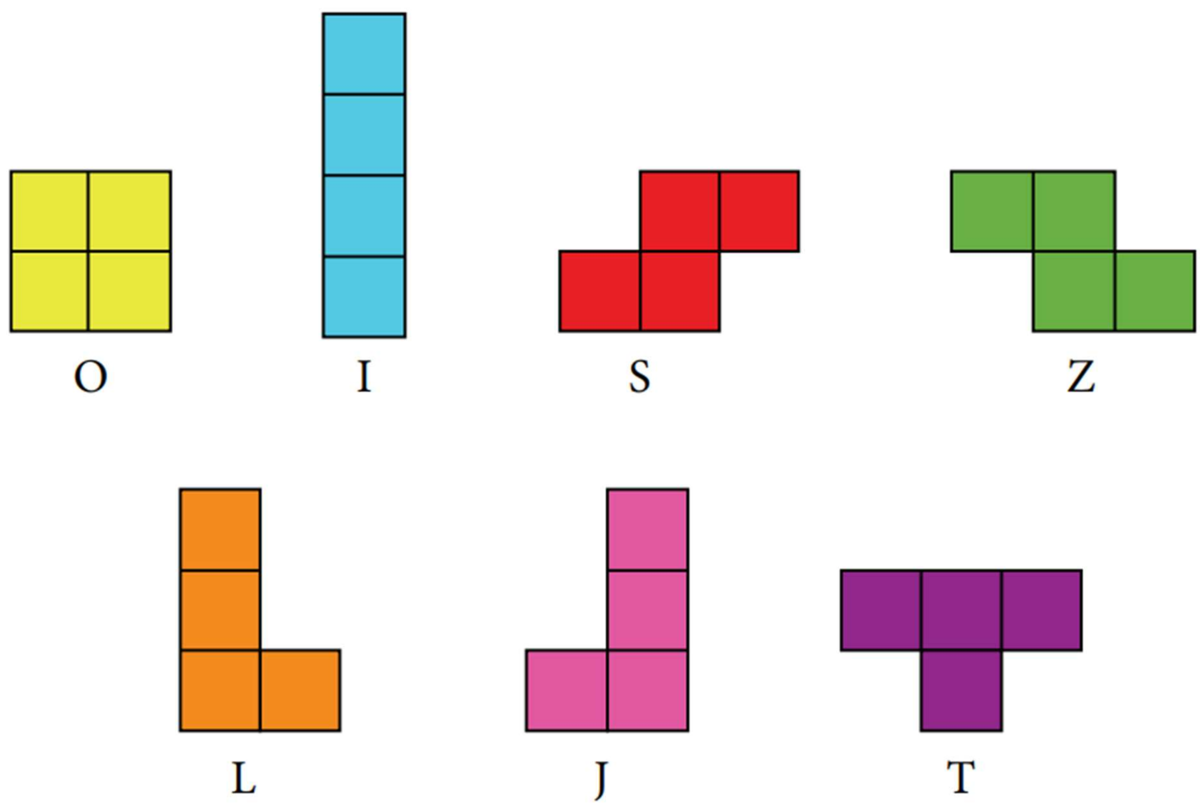

Figure 1. The seven Tetris Tetrominoes, also known as zoids, along with their standard naming convention. Adapted from “Generalisation over details: the unsuitability of supervised backpropagation networks for tetris” by I. J. Lewis and S. L. Beswick, 20

A benefit of using Tetris in a research setting is that there exists a version of it created for behavioral research, Meta-T (Lindstedt & Gray, 2015), which offers high flexibility in manipulating the task, as well as a robust logging system.

In this experiment, Tetris also served as the baseline task, which was observed to determine the effects of the manipulation on performance. This was useful both from an experimental design standpoint
(one fewer thing to debug!) as well as the fact that the subject was changing modes of thought between similar tasks (both Tetris) as opposed to dissimilar ones. Presumably flow is a fairly fragile state, and forcibly pulling someone out of a task and into a different one would likely be more damaging to the state of flow than pulling someone from one task and placing them into a similar task.

Previous studies have looked at inducing flow through game-like conditions similar to the ones used in this study. Engeser and Rheinberg (2008) studied flow and its components under several tasks, including the game of pac-man. Their pac-man experiment manipulated the game difficulty in order to test its effect on flow, and they found a degree of support for the assumption that challenge-skill balance is necessary for flow.

Many studies have utilized Tetris itself in the study of flow. Chanel, Rebetez, Bétrancourt, and Pun (2008) used Tetris as a task to induce flow, and manipulated it by modulating the game difficulty through the falling speed of the pieces. Their method was to first measure a player's skill, by finding the difficulty level at which the player reported feeling engaged. They would then set this baseline skill level as the difficulty of the medium condition, from which easier and harder conditions were generated. They measured flow through both questionnaire responses and physiological indicators. In both of these measures, they found support for the idea that different levels of difficulty in the game Tetris lead to different levels of flow in participants, with higher difficulty (and thus more quickly falling pieces) leading to anxiety, and lower difficulty (more slowly falling pieces) leading to boredom.

Keller and Bless (2008) investigated the challenge-skill balance component of flow, also using Tetris. They utilized an adaptive framework which re-evaluated the player's skill every 30 zoids, and adjusted the fall speed accordingly, to place them in a boredom, adaptive, or overload condition. They found that subjects reported higher fit of skills and task demands in the adaptive condition vs. the boredom and overload conditions. Subjects in the adaptive condition also performed better at the Tetris task than subjects in the boredom and overload conditions.

The Tetris methodology of Keller and Bless (2008) was utilized in two other studies. Keller and Blomann (2008) utilized this Tetris paradigm to find correlation between flow state and internal locus of control. Keller, Bless, Blomann, and Kleinböhl (2011) used this adaptive Tetris paradigm to study flow in conjunction with physiological measures. They found links with flow and biological markers of stress, concluding that flow experiences are subjectively positive, but physiologically stressful, reflecting higher mental load.

Plotnikov et al. (2012) utilized Tetris with conditions of varying difficulty to study the relationship between flow and EEG readings, and were able to distinguish flow from a state of boredom. Harmat et al. (2015) investigated links between flow state and neurological systems using a Tetris game with varying difficulty levels, finding limited support for a correlation between flow state and activity of the parasympathetic nervous system.

A common theme across many of these studies was the use of questionnaires to measure a subject's experience of flow. In line with this trend, in this experiment we used the

Dispositional Flow Scale (DFS) and Flow State Scale (FSS) to measure their initial state of flow, and their flow as affected by the experimental Tetris task, respectively.

Another theme was the use of an adaptive paradigm to manipulate flow state, as is seen in Keller and Bless (2008) and studies that draw from it. This is a natural fit for flow research, as presenting a good match for a subject's available attentional resources is not a trivial task. Things like fatigue and momentary distractions may alter the amount of attention the subject has at their disposal to devote to the task, and a static difficulty level cannot adjust to compensate for changes of this sort. An adaptive framework allows challenge to more closely match skill throughout the task, which should make the task as a whole more conducive to flow.

This adaptive strategy has many similarities to adaptive training, which involves the use of dynamic tasks in instruction and training to better improve performance by basing the functioning of the task on the subject's performance, skill, or other individual differences (Park & Lee, 2004; Vanderwaetere, Desmet, & Clarebout, 2011; Spain, Priest, & Murphy, 2012; Landsberg et al., 2012). Adaptive training has been of great interest due to the possibility of performance improvement not being dependent on the use of human instructors.

In the adaptive training paradigm, experimental designs which measure one's skill beforehand and set the difficulty of the task to match this initial measurement, such as that of Chanel et al. (2008), would be analogous to macro-adaptation, wherein the adaptation is based on an initial assessment given prior to instruction. Likewise, experimental designs which continually or periodically take a measure of one's Tetris performance, such as that experiments following the design of Keller and Bless (2008), as well as the experiment described in this document, would be analogous to micro-adaptation, in which instruction is tailored based on the student's current performance on the training task. Previous work in the domain of adaptive training has emphasized a holistic approach towards the various approaches towards adaptive training, but notes that micro-adaptation allows for quicker feedback and is most receptive to student needs (Mödritscher et al., 2004; Landsberg et al., 2012). These aspects suggest that it may be a more fruitful approach to inducing flow state.

Much of the more attentional research on flow has found it to be closely intertwined with perceived effort. Subjects who reported feeling more in flow generally also reported that tasks required less cognitive effort (Harris, Vine, & Wilson, 2017). This introduces the question: do subjects who report being in flow actually use less cognitive and attentional resources on the task? Or do they only perceive it to be easier, despite requiring the same amount of attention as it otherwise would. The latter seems to be more intuitive, since in this scenario, it's the same task whether or not they are in flow. It stands to reason the same amount of cognitive effort is required to perform the same exact task.

To understand how reported flow experience is related to perceived task effort, we introduce a simple effort rating to the experiment, which asks the subjects to give a rating from 1-10 on the difficulty of the task they just performed.

Methods

The experiment was approximately three hours in duration, which was spread out over two sessions to lessen the effects of fatigue due to long periods of high attention. In the first session, each subject played 50 minutes of a baseline (unmanipulated) Tetris task, called the Pre task. In the second session, which took place between 1 and 10 days after the first session, the subject played one hour of the experimental (manipulated) Tetris task, followed by another 50 minutes of baseline Tetris, called the Post task.

Demographic information was gathered at the beginning of the study, and the subject's flow was queried right before the Pre task (via the DFS) and right after the Experimental task (via the FSS). Additionally, subjects were asked to report their perception of their expended cognitive effort on a 10-point scale after each Tetris task (i.e. Pre, Experimental, and Post).

A visual summary of the experiment protocol can be seen in Figure 2. In this diagram, ovals represent Tetris tasks, rectangles represent flow scales, and octagons represent task effort rating items.

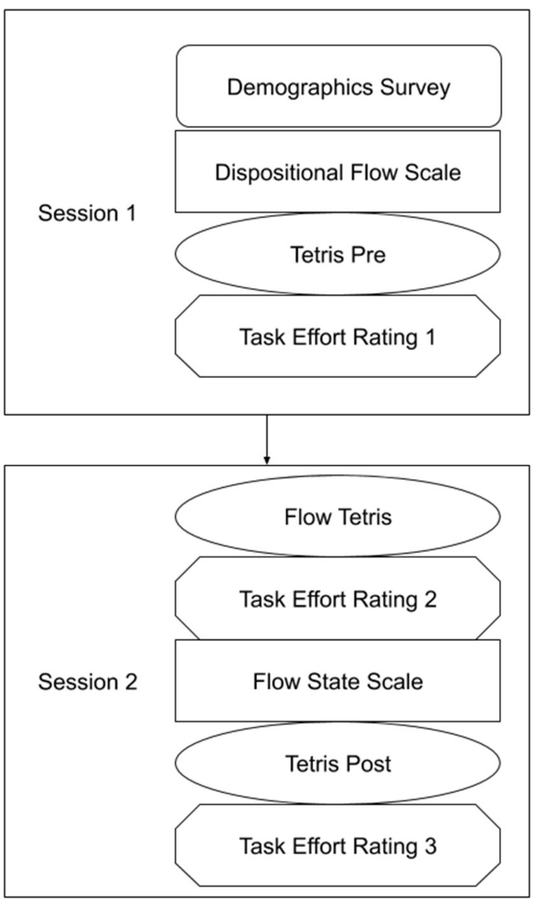

Figure 2. Diagram of experiment protocol

Change in flow state was measured by taking the difference between scores on the FSS and the DFS, per subject. This was to control for individual differences in dispositional flow between subjects. Some subjects may tend more towards entering into flow states than others, and thus would naturally score higher on the FSS. For such subjects this would not be indicative of a change in flow state as a result of the experiment, so subtracting their DFS score controls for this possibility.

Additionally, the DFS and FSS are nearly identical in structure and wording. Each item in the DFS corresponds to a near-duplicate item in the FSS. The difference between the two is that the DFS phrases item in a habitual tense ("I lose my normal awareness of time") and asks the subject to rate the frequency they experience the item, while the FSS phrases questions in more concrete past tense, referring to the task the subject just did ("I lost my normal awareness of time"), and asks subjects how much they agree that the item described their experience.

The initial flow measurement is given at the beginning of the experiment. It seemed likely that giving subjects the FSS at the beginning, before they had performed any tasks, would confuse them more than anything, as the FSS would ask them about flow experience in a task they had just performed. The DFS is a more natural way to query their initial flow, and since it is nearly identical in its measurement to the FSS, we felt it was an appropriate way to control for individual differences in dispositional flow.

Players were sorted randomly into one of 3 conditions, which correspond to one of 3 versions of the FlowTetris task. These versions differ from the baseline tasks in that they had been manipulated to have different levels of difficulty. The "easy" task should be easier for the subjects to do, the "hard" task should be harder for the subjects to do, and the "adaptive" task is so called because it combines elements of easy and hard to varying extents, depending on how well or poorly the player is currently doing. In other words, the adaptive version of Tetris is reacting in a dynamic manner, instead of in a static one. The level of difficulty in the adaptive version is not preprogrammed in the FlowTetris task. It is determined on-the-fly, according to the concept of micro-adaptation mentioned above.

To create these tasks of varying difficulty, I altered the algorithm which determines which piece is given to the player to place onto the Tetris board for each episode. In the baseline tasks, this was done randomly, as it is done in many implementations of Tetris.

In the FlowTetris task, the zoid which was given to the player is determined by a Tetris AI, which has analyzed every possible placement position and orientation of each zoid for the current board state, and then ordered the zoids in terms of easiest to hardest. Players in the easy FlowTetris condition were always given the zoid which was easiest to find a good placement for, while players in the hard condition were always given the zoid which was hardest to find a good placement for. Players in the adaptive condition were given a zoid which varied in its ease of placement, corresponding to how well the player was currently doing.

In the FlowTetris adaptive condition, the state of the Tetris board is used as a proxy for the participant's performance. Since the aim of the game is to complete rows, it is fairly intuitive (and generally agreed on within the Tetris community) to consider a player who has managed to keep their accumulation low on the board as performing better than a player whose accumulation has risen near the top. The input of the algorithm was the maximum height of the accumulation, (i.e. the height of the highest block in the current accumulation). This was mapped linearly to the sorted zoids, and the player was given a zoid which corresponded to their current performance. If the player was doing well, and thus the accumulation was fairly low, the system will give them pieces which are more difficult to place, to give the player a greater challenge. If the player was doing poorly, the accumulation would be very high, and the system would give them pieces which were easy to place, to lower the challenge to an overwhelmed player. Thus, which zoid they were given was a function of the maximum height of the accumulation where the maximum height was assumed to be a reasonably accurate proxy for performance.

The pieces were sorted into a best-to-worst list based on a Tetris solver AI model, specifically a weighted feature-sum model using the Dellacherie set of features and weights (Fahey 2003). Use of weighted feature-sums is standard practice in the Tetris AI community. The Dellacherie Tetris solver is a good balance of computational complexity and excellent performance. Computational complexity and efficiency was important because this algorithm needed to run in real time, in the time between when a player placed a zoid and was given the next zoid, for all 7 zoids, up to 40 times per zoid (10 rows * 4 orientations), to evaluate all possible placements. The Dellacherie AI controller uses only six features, and could complete 660,000 lines on average before failing, which is quite impressive for a simple linear evaluation function (Algorta and Şimşek, 2019).

The Meta-T software can compute a number of these features based on the state of the game board. These can range from simple features such as max height, which is the height of the highest point in the accumulation, to more complex features such as measures of jaggedness of the accumulation. Each feature is more or less desirable in a Tetris game; for example, the player would want to keep the maximum and average heights of the accumulation low, as well as to keep to a minimum the number of pits and wells (empty regions in the accumulation which are covered and uncovered by a piece on top, respectively). A placement of a piece will generate numeric values for these features, which can be multiplied by their respective weights and then summed together to produce a single value for that placement. A higher value for this weighted feature sum indicates a better placement, while a lower value indicated a worse placement.

In the adaptive Tetris condition of this experiment, each piece was evaluated at every possible position and orientation in which it could be placed for the current board. The zoids were then sorted by the maximum score of all the possible placements for that piece, which is to say the score of the most ideal placement for a given zoid.

The highest scoring zoid is the one which should be easiest to place, since it has the best evaluated placement, and thus it was what was given to the player in the Easy condition. The lowest scoring zoid should be hardest to place, since its best evaluated placement was the worst of all the zoids. This was the zoid which was given to the player in the hard condition. In the adaptive condition, the zoid that the player was given scaled linearly based on the maximum height of the accumulation.

Player performance was measured as a function of their game score. Players earn points by clearing rows, with more points earned for clearing multiple rows at once.

Because of the variable nature of the zoids given, players can sometimes get into situations where there are no good choices. An inconvenient zoid (or sequence of zoids) will have no good possible placements; players must choose between placements that are "least bad." Situations such as these can lead to a lower score for a game due to the variability of the zoid choice mechanism, rather than a player's skill or mental state. To control for this, performance in the analysis was measured by a "criterion score," which was defined as the average of the four highest-scoring games a player completed within the time limit. This is standard practice when using the *Meta-T* software (Sibert, Gray, & Lindstedt, 2017).

To measure change in player performance from baseline, we utilized a post-minus-pre design, where subject performance change due to the manipulated condition was measured as the difference in criterion score between the post condition (performed after the manipulated condition) and the pre condition (performed before the manipulated condition). This allowed us to control for individual differences by ensuring player performance after the manipulated condition is measured in comparison to an initial measurement.

We could have also utilized a two-way ANOVA design, such that difficulty was a between-subjects factor, as each subject is placed into an easy, adaptive, or hard condition, and time was a within-subjects factor, as all subjects did a pre, experimental, and post tetris task. However, this would introduce some additional complications, since the experimental task is not the same for all participants. Subjects in the easy condition experience an easy task during the experimental condition, subjects in the adaptive condition experience the adaptive task, and subjects in the hard condition experience the hard version of the task. This aspect of the design may have led to issues using an ANOVA design, since it could be argued that the two factors are not completely independent, with the difficulty factor dictating what task is played during one portion of the time factor. For the sake of ease of analysis and interpretation, we chose instead a simpler pre-test post-test methodology, but this is not necessarily the only way to analyze this data.

As stated above, the experiment was constructed to measure whether flow was effectively induced or discouraged, as well as to test the amount of transfer of beneficial effects between the experimental task and a subsequent task, in this case the post Tetris task.

It seemed likely that the adaptive manipulation would induce flow, as it is similar to successful manipulations used in Tetris-based flow experiments in the past (Chanel, Rebetez, Bétrancourt, & Pun, 2008; Keller & Bless, 2008; Keller & Blomann, 2008; Keller, Bless, Blomann, & Kleinböhl, 2011; Harmat et al., 2015).

It also does not seem unreasonable that flow may have a beneficial effect on performance in subsequent tasks which the subject might do immediately after experiencing a flow state, since a similar effect can be seen with analogous psychological phenomena such as near transfer (Perkins & Saloman, 1992) and persistence of mood (Hills, Hill, Mamone, & Dickerson, 2001).

Thus, this experiment was designed to test the following two hypotheses:

Hypothesis 1: Subjects who play in the "adaptive" experimental condition will report higher state flow than those in the "easy" or "hard" conditions.

Hypothesis 2: Subjects who play in the "adaptive" experimental condition will perform better in the post task, relative to their performance in the pre task, than subjects who play in the "easy" or "hard" conditions.

In addition to these two specific hypotheses, this study can also be used to test a more general hypothesis that integrates the relationships between the "adaptive" manipulation, effort perception, flow, and performance in a single statistical model. Thus, adaptivity was hypothesized to have a positive direct effect on performance and a positive indirect effect on performance via flow. In addition, the positive effect of adaptivity on flow was expected to be mediated by effort perception: the adaptive condition would reduce effort perception which in turn would increase flow.

Results

I ran 137 subjects total, dropping 23 of them due to incomplete data. Of the remaining 113 subjects, their gender (female=69, male=43, non-binary/other=1) and age (M=20.02, SD=3.59) were not out of the norm for what would be expected of a undergraduate college population.

For the majority of the experiment, each of the 113 subjects were randomly assigned to one of the three conditions. However, for the initial week or so of testing, a misconfigured config file led to the first batch of subjects being assigned exclusively to the adaptive condition. After this error was noted and fixed, random assignment worked as intended, with a moderate number of subjects in conditions (adaptive=48, easy=32, hard=33). Due to time and organizational constraints, it was not practical to gather data from further subjects to correct this imbalance, so for this analysis it has been allowed to persist.

A summary of the subject performance on each Tetris task can be seen in Figure 3. As is to be expected, performance is very similar across all conditions for the pre task, since all subjects would have had the same treatment up through that task. The easy condition produced higher scores than baseline, and the hard condition produced lower scores than baseline, which were the intended effects of those manipulations.

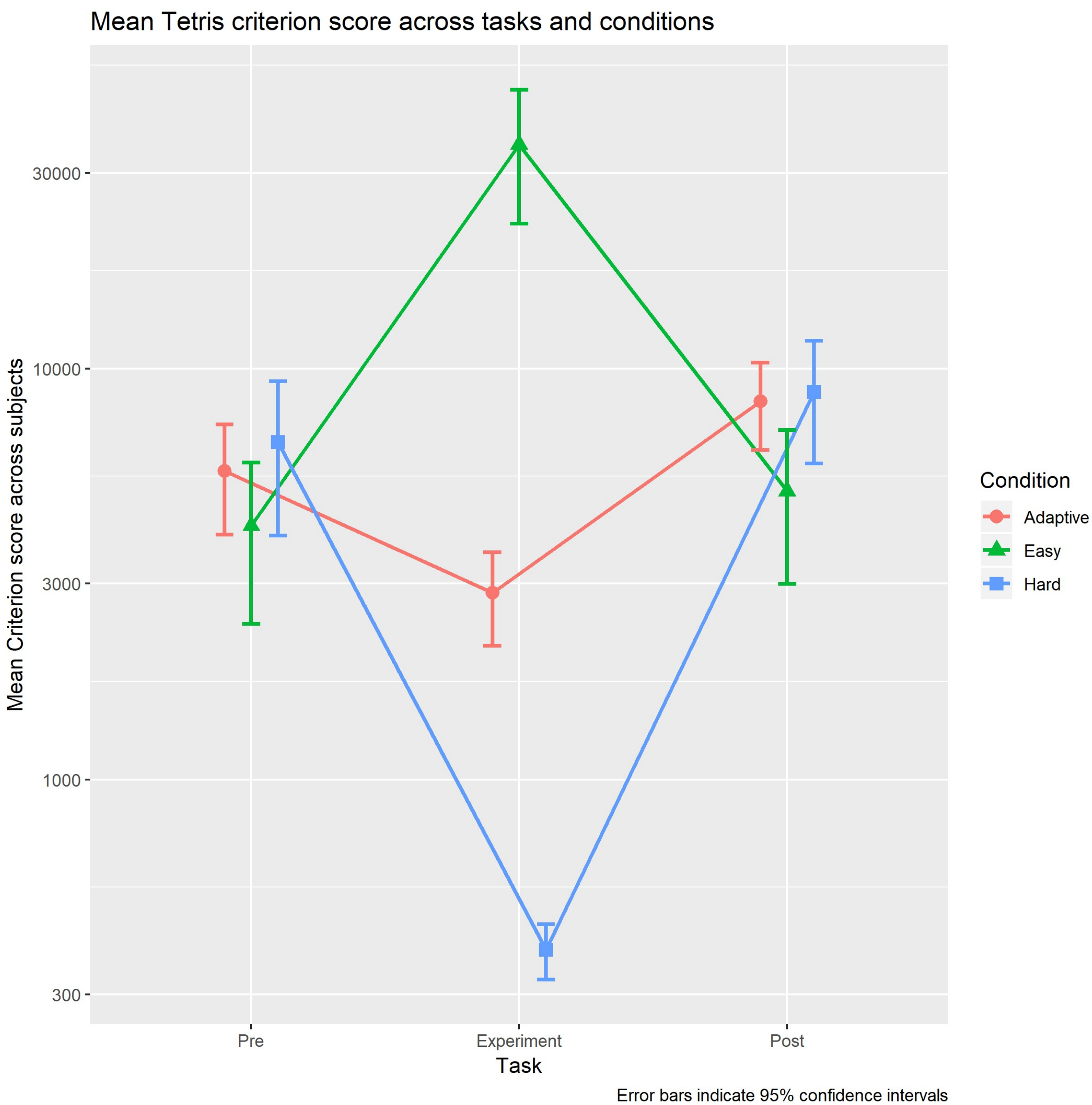

Figure 3. *Time course analysis across both between- and within-subjects conditions, summarized across criterion score. Note the log scale.*

To address hypothesis 1, I ran a one-way ANOVA to determine the effect of tetris difficulty manipulation on the difference in reported flow between the flow questionnaires which occurred before and after the experimentally manipulated task (the DFS and FSS respectively). These questionnaires utilized 5-point Likert-type scales, which were converted into integer values where "Strongly Disagree" (FSS) and "Never" (DFS) correspond to a value of 1, and

“Strongly Agree” (FSS) and “Always” (DFS) correspond to a value of 5. The result showed a significant effect of tetris condition on reported flow state, $F(2,110)=19.37$, $p<0.001$; Eta^2 = 0.26. A Tukey’s HSD followup test found significant differences between all pairwise comparisons of experimental groups, at $p<.05$ for all. A summary of the scores for each condition can be found in Table 1, and they can be seen visually in Figure 4.

Condition	Mean	SD
Adaptive	.0295	.556
Easy	.321	.447
Hard	-.481	.559

Table 1. Mean and standard deviation of differences between flow scores for each condition. Positive mean indicates greater flow reported in post-test than in pre-test.

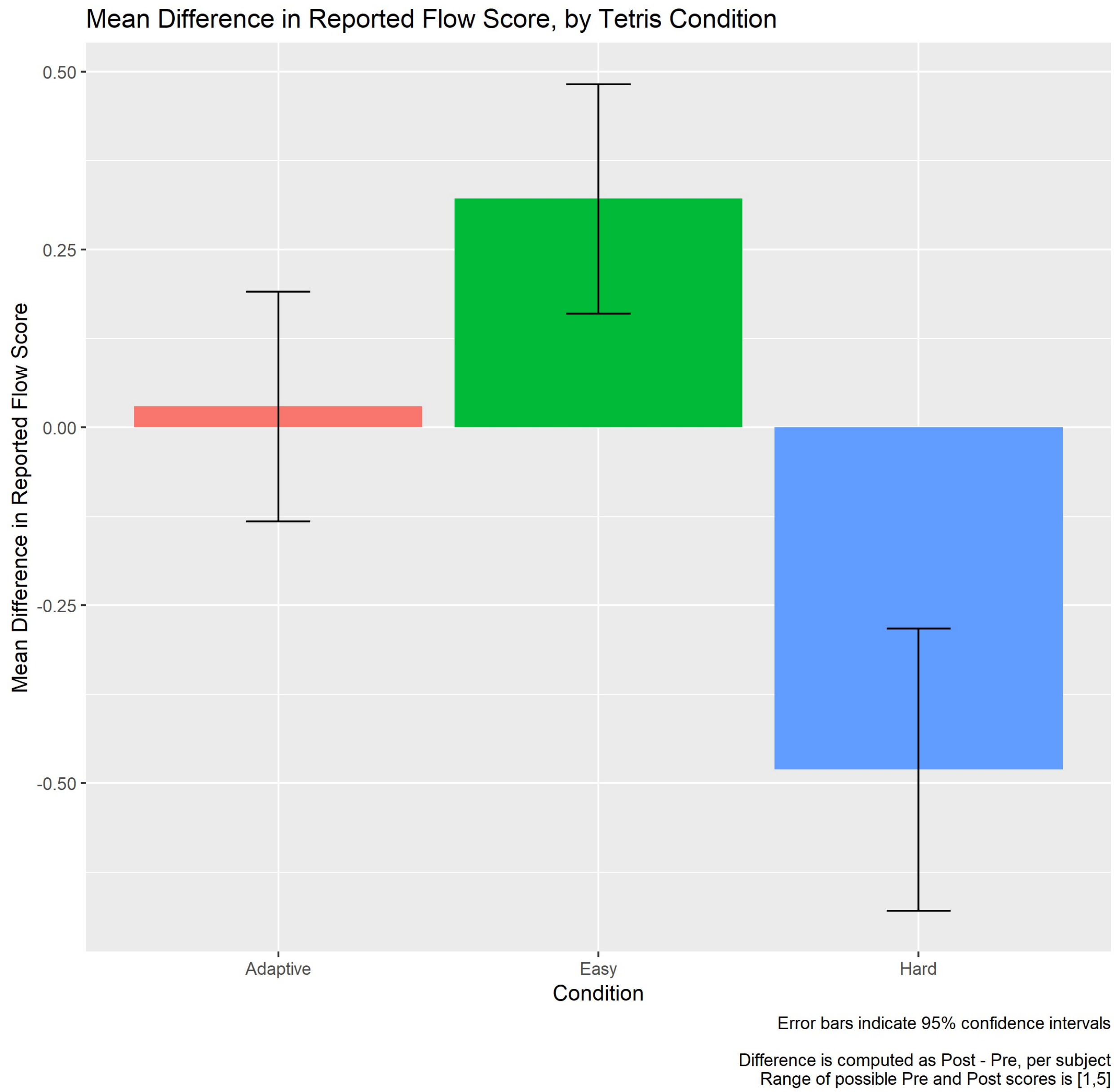

Figure 4. Mean difference in reported flow score across conditions

To address hypothesis 2, I ran a one-way ANOVA to determine the effect of Tetris difficulty manipulation on the difference in criterion scores for the baseline and contrast Tetris tasks, occurring before and after the experimentally manipulated tetris task. The results showed no significant effect of Tetris difficulty condition on performance difference, $F(2,110)=2.031$, $p=0.135$, $Eta^2 = 0.035$; see Figure 5.

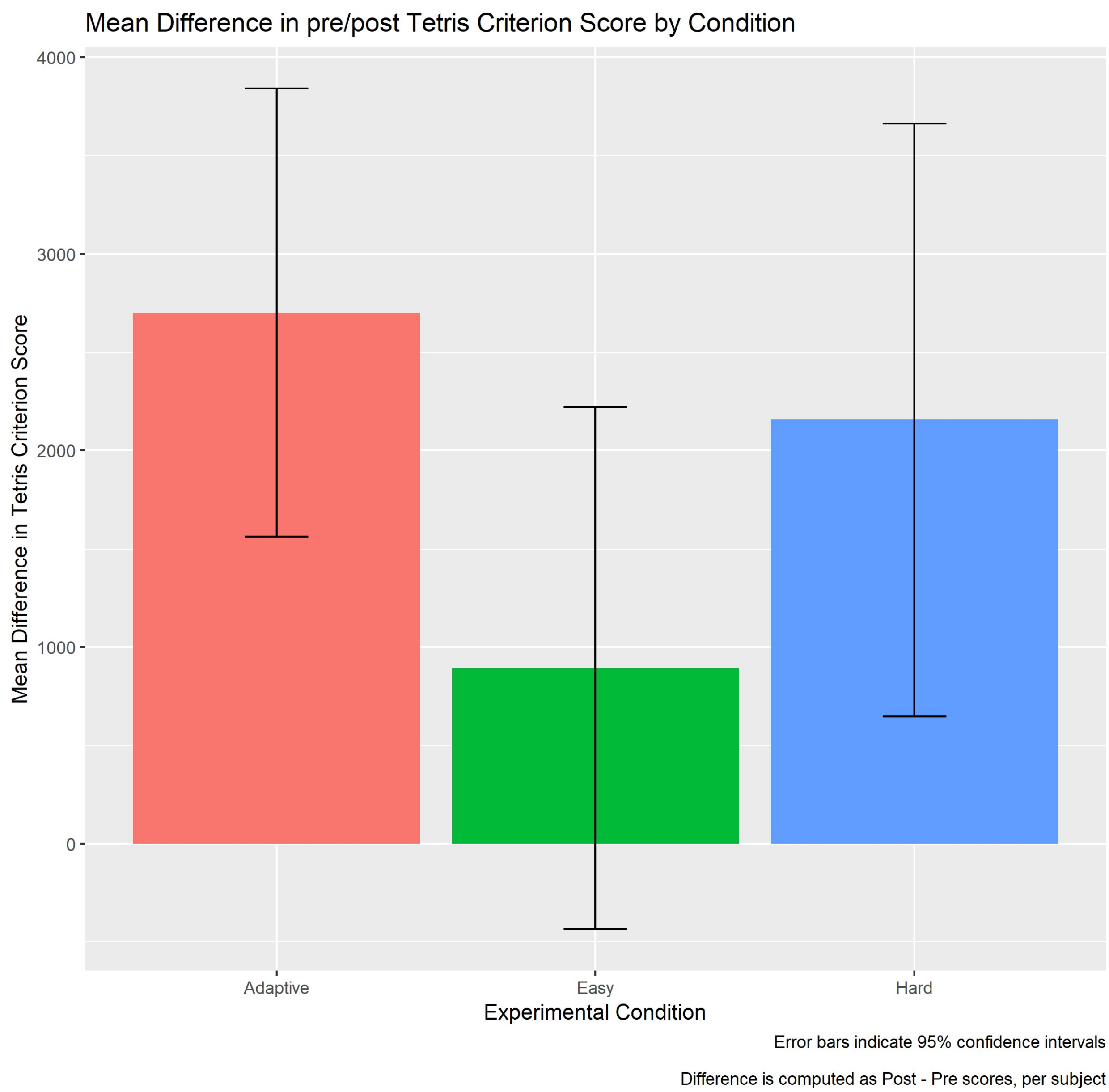

Figure 5. Mean difference in pre/post Tetris Criterion Score by Condition

There were no significant differences in reported task effort immediately after the manipulated task between any of the conditions, F(2,110)=.055, P=.138. This can be seen in Figure 6, as the middle block of bars, which represent the task effort measurement of the experimental task itself. As expected, there are no significant differences either for the pre- and post-tasks, since subjects in all conditions are doing the exact same task.

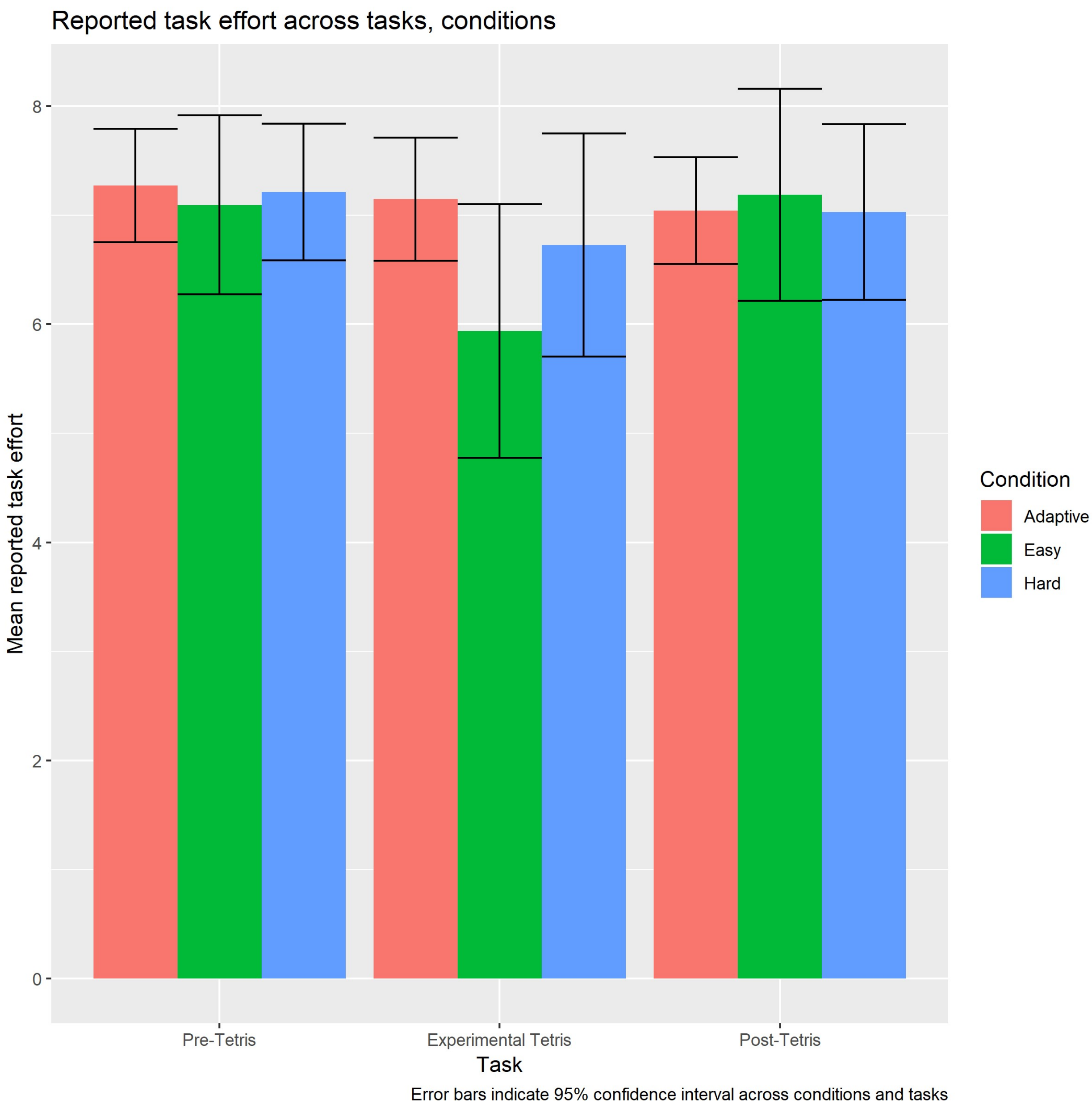

Figure 6. Reported task effort across tasks and conditions.

The question remains as to what extent these results were affected by the fact that a number of subjects at the beginning of data collection were placed exclusively into the adaptive condition. Based on the logs of who was run before the error was discovered, I filtered out 9 subjects who were placed in the adaptive condition, leaving the condition totals at adaptive=39, easy=32, and hard=33.

Hypothesis 1, looking at the effect on reported flow score from different Tetris conditions, did not show a marked difference. The filtered results showed $F(2,101)=18.09$, $p<0.001$; $Eta^2=0.26$; as opposed to the unfiltered results (as reported above) being $F(2,110)=19.37$, $p<0.001$; $Eta^2 = 0.26$. While the unfiltered results showed significant differences between all pairwise groups, the filtered results only showed significant differences between easy and hard, and adaptive and hard. The difference between easy and adaptive became non significant at $p=.076$.

Hypothesis 2, looking at the effect on performance change from different Tetris conditions, similarly did not show a marked difference. The filtered results showed $F(2,101)=1.72$, $p=.184$, $Eta^2=.032$, as compared to the unfiltered results (again, as reported above) being $F(2,110)=2.031$, $p=0.135$, $Eta^2 = 0.035$.

Discussion

As hypothesized, there were significant differences across conditions for changes in reported flow score. However, while I hypothesized that the adaptive condition would lead to the highest reported flow, in fact the easy condition led to the highest reported flow. The hard condition successfully frustrated subjects, as subjects in that condition reported lower flow afterwards. However, the adaptive condition largely did not affect flow score for subjects, with subjects on average reporting very similar flow scores after the experiment to their scores beforehand.

Theoretically, an easy task should lead to low reported flow, as it would be a mismatch between challenge and skill. The skill available to handle the task far outstrips the challenge of the task itself, leading to a state of boredom instead of flow. This is what is seen in much of the previous literature; for example, as stated above, Chanel, et al. (2008) found that subjects reported boredom rather than flow when exposed to an easy condition.

This high-reporting of flow in the easy condition may be due to a challenge floor present in the Tetris task itself. Playing Tetris at low skill levels can be challenging even at the lowest levels, because there is always a time pressure present. Players have no option to pause the descent of the falling zoid. There is roughly five seconds of falling time between the top and bottom of the tetris board on the easiest level, and if the subject cannot find their ideal final placement for that piece in a significantly shorter time span, they are unlikely to ever enter into a state of boredom. They will remain engaged and at some balance of challenge and skill (and therefore flow) by the natural progression of difficulty of the game, both due to increasing zoid falling speeds (as the level increases), as well as faster required reaction time (as through play the height of the accumulation increases, leading to there being less space between where the piece is generated at the top of the board, and where it can land). The demographic information (shown in figure 9) indicates that a majority of the subjects identify themselves as novices at

Tetris, lending some credence to this possibility.

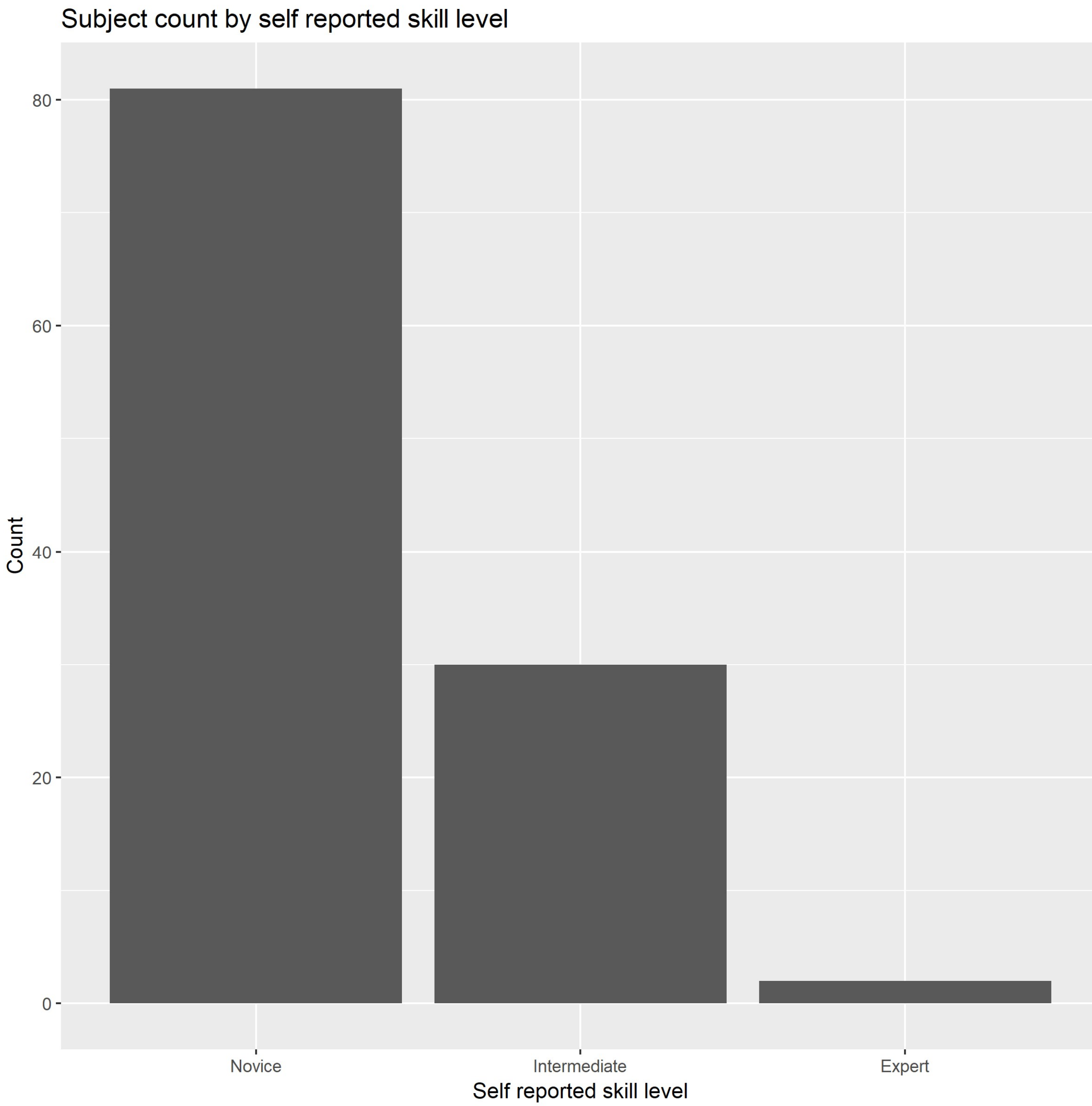

Figure 7. Subject count by self-reported skill level. Subjects were asked to estimate their skill level by placing themselves in exactly one category out of [Novice, Intermediate, Expert]

Most previous studies, including Chanel, et al. (2008), manipulated difficulty by altering the speed at which the Tetris pieces fall at, thereby reducing the effect of this aspect of Tetris on the challenge felt by the player.

The same challenge floor effect could also explain why the adaptive condition was not associated with lower perceived effort than the hard condition. There is a minimum amount of effort needed to determine where to place a given piece. While the easy condition was designed to minimize this amount of effort, by always giving the player a piece to which the ideal placement is likely obvious, the adaptive and hard conditions do not.

The adaptive condition will only give the player pieces which are easy to place when the player is close to losing. In such a scenario, the accumulation is near to the top of the game board, and because of this, there is little time for the piece to fall from the top before it hits the accumulation. This means that although the pieces given are the same as would be given in the easy condition, they still require high attentional load, since the player has a much shorter length of time to decide where to place the piece and then execute the sequence of keypresses which will place the piece there. As the height of the accumulation decreases, which gives the player longer to decide where to place the pieces, the adaptive system gives pieces which are harder to place, offsetting any lessening of required attentional load gained from getting more time to react. This may explain why the adaptive and hard conditions showed similar levels of reported attentional load.

Future research should more fully explore manipulations intended to alter the task challenge and the corresponding effort perception. Tetris is a complex task, with many dimensions that may be altered to increase or decrease the task difficulty. The challenge of the game, and the amount of attention it demands, can even vary from moment to moment, as players may create problems for themselves, or solve those same problems through their own behavior and their decisions of where to place each piece. This property of the game makes it interesting and compelling, contributing both to its commercial success as well as the interest researchers (both previous and present) have taken in the game for exploring the concept of flow.

However, this complex nature also makes it difficult to find a one-size-fits-all solution for manipulating its task difficulty in experimental settings. It's possible that altering the speed at which pieces fall may indeed be a more efficacious way to alter task challenge, or that altering other aspects of the game may be a better method to induce flow to a greater or lesser extent. There may even be other games or tasks which might more clearly show differences in this area. Further research into Tetris and other games as tasks to explore flow would be beneficial for the field.

Conclusion

While the manipulations did effectively induce flow to a greater or lesser extent across conditions, it produced the highest flow in the easy condition, rather than the adaptive condition, as expected. As described in the discussion section, this may be due to a floor effect in the difficulty of the Tetris task.

No significant differences were found in the post-task Tetris performance, meaning that no lingering effects of flow state were measured purely via performance on the Tetris task given after the experimentally manipulated task.

Overall, the results of this experiment do not show evidence of flow affecting future performance. However, future research may make use of this established methodology and may better establish the relationship of flow with performance and task effort.

As mentioned earlier in the document, the core question of this work was, “How can flow be effectively induced, and is it possible for manipulations of a flow-inducing task to affect performance on a task which has not been altered to induce flow?”

We can address the first part of this question via the previous research into this area, and through the experiment presented here: at the very least, it seems likely that it is possible to effectively induce flow. Czikszentmihalyi identified nine dimensions of flow, three of which (clear proximal goals, immediate feedback, and challenge-skill balance) are aspects of a task which cause those doing a task to be more likely to enter into flow. These aspects have been utilized throughout the history of flow research to create tasks which can induce flow to a greater or lesser extent in subjects, and a number of these experiments have been described in the Introduction section of this document. And as can be seen in the results section, the manipulations to the task results in subjects reporting significant differences in flow across conditions. However, in one aspect, the flow reported from subjects was contrary to our expectations, in that the easy condition led to significantly higher flow than the adaptive condition. These differences were still significant, and some explanation for why this behavior was observed is given in the discussion, but future work is required in order to have confidence about these manipulations effectively inducing flow state.

The second part of the question, “is it possible for manipulations of a flow-inducing task to affect performance on a task which has not been altered to induce flow?”, is more difficult to answer. If this transfer of flow was possible, it could be integrated into real-world tasks to increase performance and reduce fatigue, and this possibility was what motivated this work. Everything, no matter how unlikely, is possible; we can’t conclusively say that it is impossible for flow transfer to exist. But given our experimental findings, flow transfer seems improbable.

There were no significant effects on performance from the manipulation conditions, and so the gathered evidence seems in favor of the nonexistence of flow transfer.

Still, there are applications that can be drawn from this work that would be beneficial to the field of human factors. The methodology for inducing flow state in participants seems to be sound. It supports Czikszentmihalyi's assertion that a balance between challenge and skill is necessary for flow, and the idea that manipulating this balance can have effects on how subjects experience flow. Future studies on flow state can make use of such methodology to manipulate flow state in subjects.

References

Algorta, S., & Şimşek, Ö. (2019). The Game of Tetris in Machine Learning. *arXiv preprint arXiv:1905.01652.*

Aubé, C., Brunelle, E., & Rousseau, V. (2014). Flow experience and team performance: The role of team goal commitment and information exchange. Motivation and Emotion, 38(1), 120-130.

Chanel, G., Rebetez, C., Bétrancourt, M., & Pun, T. (2008, October). Boredom, engagement and anxiety as indicators for adaptation to difficulty in games. In *Proceedings of the 12th international conference on Entertainment and media in the ubiquitous era* (pp. 13-17). ACM.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive psychology, 4*(1), 55-81.

Cooper, L. A. (1975). Mental rotation of random two-dimensional shapes. *Cognitive psychology, 7*(1), 20-43.

Csikszentmihalyi, M., & Csikszentmihalyi, I. S. (Eds.). (1992). *Optimal experience: Psychological studies of flow in consciousness*. Cambridge university press.

Csikszentmihalyi, M., & Larson, R. (2014). Validity and reliability of the experience-sampling method. In *Flow and the foundations of positive psychology* (pp. 35-54). Springer Netherlands.

Csikszentmihalyi, M. (1975). Play and intrinsic rewards. *Journal of Humanistic Psychology, 15*(3), 41-63. http://dx.doi.org/10.1177/002216787501500306

Csikszentmihalyi, M. (1990). *Flow: The Psychology of Optimal Experience*. New York: Harper & Row.

Csikszentmihalyi, M. (1997). *Finding flow: The psychology of engagement with everyday life.* Basic Books.

Eisenberger, R., Jones, J. R., Stinglhamber, F., Shanock, L., & Randall, A. T. (2005). Flow experiences at work: For high need achievers alone?. Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior, 26(7), 755-775.

Engeser, S., & Rheinberg, F. (2008). Flow, performance and moderators of challenge-skill balance. *Motivation and Emotion*, 32(3), 158-172.

Ericsson, K. A. (Ed.). (2009). *Development of professional expertise: Toward measurement of expert performance and design of optimal learning environments*. Cambridge University Press.

Fahey, C. P. (2003). Tetris AI, Computer plays Tetris. http://colinfahey.com/tetris/tetris_en. Html.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.

Fullagar, C. J., Knight, P. A., & Sovern, H. S. (2013). Challenge/skill balance, flow, and performance anxiety. Applied Psychology, 62(2), 236-259.

Gopher, D., & Braune, R. (1984). On the psychophysics of workload: Why bother with subjective measures?. *Human Factors, 26*(5), 519-532.

Gray, W. D., & Lindstedt, J. K. (2017). Plateaus, dips, and leaps: Where to look for inventions and discoveries during skilled performance. Cognitive science, 41(7), 1838-1870.

Harris, D. J., Vine, S. J., & Wilson, M. R. (2017). Is flow really effortless? The complex role of effortful attention. Sport, Exercise, and Performance Psychology, 6(1), 103.

Harmat, L., de Manzano, Ö., Theorell, T., Högman, L., Fischer, H., & Ullén, F. (2015). Physiological correlates of the flow experience during computer game playing. *International Journal of Psychophysiology*, 97(1), 1-7.

Heyne, K., Pavlas, D., & Salas, E. (2011, September). An investigation on the effects of flow state on team process and outcomes. In Proceedings of the human factors and ergonomics society annual meeting (Vol. 55, No. 1, pp. 475-479). Sage CA: Los Angeles, CA: SAGE Publications.

Hills, A. M., Hill, S., Mamone, N., & Dickerson, M. (2001). Induced mood and persistence at gaming. Addiction, 96(11), 1629-1638.

Hjortskov, N., Rissén, D., Blangsted, A. K., Fallentin, N., Lundberg, U., & Søgaard, K. (2004). The effect of mental stress on heart rate variability and blood pressure during computer work. *European journal of applied physiology, 92*(1-2), 84-89.

Hockey, R. (2013). *The psychology of fatigue: Work, effort and control*. Cambridge University Press.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal, 6*(1), 1-55.

Jackson, S. A. (1995). Factors influencing the occurrence of flow state in elite athletes. *Journal of applied sport psychology*, *7*(2), 138-166.

Jackson, S. A. (2000). Joy, fun, and flow state in sport. *Emotions in sport*, 135-155.

Jackson, S. A., & Marsh, H. W. (1996). Development and validation of a scale to measure optimal experience: The Flow State Scale. *Journal of sport and exercise psychology, 18*(1), 17-35.

Jackson, S. A., & Eklund, R. C. (2002). Assessing flow in physical activity: The flow state scale–2 and dispositional flow scale–2. *Journal of Sport and Exercise Psychology, 24*(2), 133-150.

Jackson, S. A., Martin, A. J., & Eklund, R. C. (2008). Long and short measures of flow: The construct validity of the FSS-2, DFS-2, and new brief counterparts. *Journal of Sport and Exercise Psychology, 30*(5), 561-587.

Jackson, S. A., Thomas, P. R., Marsh, H. W., & Smethurst, C. J. (2001). Relationships between flow, self-concept, psychological skills, and performance. Journal of applied sport psychology, 13(2), 129-153.

Jorna, P. G. (1992). Spectral analysis of heart rate and psychological state: A review of its validity as a workload index. *Biological psychology, 34*(2-3), 237-257.

Juvina, I., Nador, J., Larue, O., Green, R., Harel, A., & Minnery, B. (2018). Measuring individual differences in cognitive effort avoidance. *Proceedings of the 40th Annual Meeting of the Cognitive Science Society*, Madison, WI.

Keller, J., & Bless, H. (2008). Flow and regulatory compatibility: An experimental approach to the flow model of intrinsic motivation. *Personality and social psychology bulletin, 34*(2), 196-209.

Keller, J., & Blomann, F. (2008). Locus of control and the flow experience: An experimental analysis. *European Journal of Personality: Published for the European Association of Personality Psychology, 22*(7), 589-607.

Keller, J., Bless, H., Blomann, F., & Kleinböhl, D. (2011). Physiological aspects of flow experiences: Skills-demand-compatibility effects on heart rate variability and salivary cortisol. *Journal of Experimental Social Psychology*, 47(4), 849-852.

Kool, W., McGuire, J. T., Rosen, Z. B., & Botvinick, M. M. (2010). Decision making and the avoidance of cognitive demand. *Journal of Experimental Psychology: General, 139*(4), 665.

Landsberg, C. R., Astwood Jr, R. S., Van Buskirk, W. L., Townsend, L. N., Steinhauser, N. B., & Mercado, A. D. (2012). Review of adaptive training system techniques. *Military Psychology, 24*(2), 96-113.

Lazarovitz, S. M. (2004). Team and individual flow in female ice hockey players: The relationships between flow, group cohesion, and athletic performance (Doctoral dissertation, ProQuest Information & Learning).

Lewis, I. J., & Beswick, S. L. (2015). Generalisation over details: the unsuitability of supervised backpropagation networks for tetris. Advances in Artificial Neural Systems, 2015, 4.

Lindstedt, J. K., & Gray, W. D. (2015). Meta-T: Tetris® as an experimental paradigm for cognitive skills research. *Behavior research methods, 47*(4), 945-965.

Mödritscher, F., Garcia-Barrios, V. M., & Gütl, C. (2004). The Past, the Present and the Future of adaptive E-Learning. *Proceedings of ICL* 2004.

Nakamura, J., & Csikszentmihalyi, M. (2009). Flow theory and research. *Handbook of positive psychology*, 195-206.

Nakamura, J., & Csikszentmihalyi, M. (2014). The concept of flow. In *Flow and the foundations of positive psychology* (pp. 239-263). Springer Netherlands.

Park, O. C., & Lee, J. (2004). Adaptive instructional systems.

Perkins, D. N., & Salomon, G. (1992). Transfer of learning. International encyclopedia of education, 2, 6452-6457.

Pilegard, C., & Mayer, R. E. (2018). Game over for Tetris as a platform for cognitive skill training. *Contemporary Educational Psychology, 54*, 29-41.

Plotnikov, A., Stakheika, N., De Gloria, A., Schatten, C., Bellotti, F., Berta, R., ... & Ansovini, F. (2012, July). Exploiting real-time EEG analysis for assessing flow in games. In *Advanced Learning Technologies (ICALT), 2012 IEEE 12th International Conference on* (pp. 688-689). IEEE.

Sibert, C. (2015). *Improving novice Tetris players with feedback from AI model based tutors* (Master's Thesis, Rensselaer Polytechnic Institute).

Sibert, C., Gray, W. D., & Lindstedt, J. K. (2015). Tetris-: Exploring Human Performance via Cross Entropy Reinforcement Learning Models. In *CogSci*.

Sibert, C., Gray, W. D., & Lindstedt, J. K. (2017). Interrogating Feature Learning Models to Discover Insights Into the Development of Human Expertise in a Real-Time, Dynamic Decision-Making Task. *Topics in cognitive science, 9*(2), 374-394.

Spain, R. D., Priest, H. A., & Murphy, J. S. (2012). Current trends in adaptive training with military applications: An introduction. *Military Psychology*, 24(2), 87-95.

Taatgen, N. A., Juvina, I., Schipper, M., Borst, J. P., & Martens, S. (2009). Too much control can hurt: A threaded cognition model of the attentional blink. *Cognitive psychology, 59*(1), 1-29.

Thiery, C., & Scherrer, B. (2009). Building controllers for Tetris. *Icga Journal, 32*(1), 3-11.

Thompson, J. J., McColeman, C. M., Stepanova, E. R., & Blair, M. R. (2017). Using Video Game Telemetry Data to Research Motor Chunking, Action Latencies, and Complex Cognitive-Motor Skill Learning. *Topics in cognitive science, 9*(2), 467-484.

Tozman, T., Magdas, E. S., MacDougall, H. G., & Vollmeyer, R. (2015). Understanding the psychophysiology of flow: A driving simulator experiment to investigate the relationship between flow and heart rate variability. *Computers in Human Behavior, 52*, 408-418.

Vandewaetere, M., Desmet, P., & Clarebout, G. (2011). The contribution of learner characteristics in the development of computer-based adaptive learning environments. *Computers in Human Behavior, 27*(1), 118-130.

Westbrook, A., & Braver, T. S. (2015). Cognitive effort: A neuroeconomic approach. *Cognitive, Affective, & Behavioral Neuroscience, 15*(2), 395-415.

Webster, J., Trevino, L. K., & Ryan, L. (1993). The dimensionality and correlates of flow in human-computer interactions. *Computers in human behavior, 9*(4), 411-426.

Vollmeyer, R., & Rheinberg, F. (2006). Motivational effects on self-regulated learning with different tasks. *Educational Psychology Review, 18*(3), 239-253.

Appendix 1: Flow Scales

The following flow questionnaires are adapted from the Dispositional Flow Scale 2 (DFS-2) and Flow State Scale 2 (FSS-2), (Jackson & Eklund, 2002), with minor changes in prompting to ensure that the subject evaluates their flow experience as a habitual mental state (as in the DFS-2) or with respect to the task which immediately preceded the administration of the scale (as in the FSS-2)

DFS-2

Responses are prompted by the following:

When you do things in your everyday life, how often would you say the following statements match your mental state?

Responses to all items follow a Likert-type scale format, requiring subjects to choose among the following answers:

- **Never**
- **Rarely**
- **Sometimes**
- **Frequently**
- **Always**

Subjects are presented with the following

1. When I do a challenging task, I believe my skills will allow me to meet the challenge.
2. When I do a task, I make the correct movements without thinking about trying to do so.
3. When doing tasks, I know clearly what I want to do.
4. When I'm doing a task, it is really clear to me how my performance is going.
5. When I do a task, my attention is focused entirely on what I am doing.
6. When I do a task, I have a sense of control over what I am doing.
7. When I do a task, I am not concerned about what others may be thinking of me.
8. When I do a task, time seems to alter (either slows down or speeds up).
9. When I do a task, I really enjoy the experience.
10. When I do a challenging task, my abilities match the high challenge of the situation.
11. When I do a task, things just seem to be happening automatically.
12. When I do a task, I have a strong sense of what I want to do.
13. When I do a task, I am aware of how well I am performing.
14. When I do a task, it is no effort to keep my mind on what is happening.
15. When I do a task, I feel like I can control what I am doing.
16. When I do a task, I am not concerned with how others may be evaluating me.

17. When I do a task, the way time passes seems to be different from normal.
18. When I do a task, I love the feeling of the performance and want to capture it again.
19. When I do a challenging task, I feel I am competent enough to meet the high demands of the situation
20. When I do a task, I perform automatically, without thinking too much.
21. When I do a task, I know what I want to achieve.
22. When I do a task, I have a good idea while I am performing about how well I am doing.
23. When I do a task, I have total concentration.
24. When I do a task, I have a feeling of total control.
25. When I do a task, I am not concerned with how I am presenting myself.
26. When I do a task, it feels like time goes by quickly.
27. When I do a task, the experience leaves me feeling great.
28. When I do a task, the challenge and my skills are at an equally high level.
29. When I do a task, I do things spontaneously and automatically without having to think.
30. When I do a task, my goals are clearly defined.
31. When I do a task, I can tell by the way I am performing how well I am doing.
32. When I do a task, I am completely focused on the task at hand.
33. When I do a task, I feel in total control of my body.
34. When I do a task, I am not worried about what others may be thinking of me.
35. When I do a task, I lose my normal awareness of time.
36. When I do a task, the experience is extremely rewarding.

FSS-2

Responses are prompted by the following:

How much would you agree that each statement below matched your mental state during the task you just did?

Responses to all items follow a Likert-type scale format, requiring subjects to choose among the following answers:

- **Strongly Disagree**
- **Disagree**
- **Neither agree nor disagree**
- **Agree**
- **Strongly Agree**

Subjects are presented with the following

1. I was challenged, but I believed my skills would allow me to meet the challenge.

2. I made the correct movements without thinking about trying to do so.
3. I knew clearly what I wanted to do.
4. It was really clear to me how my performance was going.
5. My attention was focused entirely on what I was doing.
6. I had a sense of control over what I was doing.
7. I was not concerned about what others may have been thinking of me.
8. Time seemed to alter (either slowed down or sped up)
9. I really enjoyed the experience.
10. My abilities matched the high challenge of the situation.
11. Things just seemed to be happening automatically.
12. I had a strong sense of what I wanted to do.
13. I was aware of how well I was performing.
14. It was no effort to keep my mind on what was happening.
15. I felt like I could control what I was doing.
16. I was not concerned with how others may have been evaluating me.
17. The way time passed seemed to be different from normal.
18. I loved the feeling of the performance and want to capture it again.
19. I felt I was competent enough to meet the high demands of the situation.
20. I performed automatically, without thinking too much.
21. I knew what I wanted to achieve.
22. I had a good idea while I was performing about how well I was doing.
23. I had total concentration.
24. I had a feeling of total control.
25. I was not concerned with how I was presenting myself.
26. It felt like time went by quickly.
27. The experience left me feeling great.
28. The challenge and my skills were at an equally high level.
29. I did things spontaneously and automatically without having to think.
30. My goals were clearly defined.
31. I could tell by the way I was performing how well I was doing.
32. I was completely focused on the task at hand.
33. I felt in total control of my body.
34. I was not worried about what others may have been thinking of me.
35. I lost my normal awareness of time.
36. I found the experience extremely rewarding.

Appendix 2: Pilot Experiments

Introduction

The form of the experiment seen in the above document occurred as the result of several years of design iteration. Two full experiments were performed on previous versions of the protocol, and a pilot study was performed on the current version of the protocol. The following sections describe the design and results of these previous iterations.

Experiment 1

Our study into the flow state has focused on how it can be induced, and whether the flow state can be transferred from one task to another. Flow state seems to offer a number of positive benefits, namely that flow-inducing activities are intrinsically motivating and are subjectively perceived as being less cognitively demanding (Nakamura & Csikszentmihalyi, 2009). Being able to transfer these benefits to tasks which don't themselves induce flow could lead to improvements such as reduced operator fatigue and increased concentration. Many tasks which have high-stakes and require constant attention are not necessarily flow-inducing, and so the ability to transfer flow or its positive effects to such a task could improve performance and in some circumstances might save human lives.

Previous research has shown that flow state can be induced, and *Tetris* seems to be one of the more prevalent paradigms used to do so (Chanel & Rebetez 2008; Harmat et al. 2015; Keller & Bless, 2008; Keller,Bless, Blomann & Kleinböhl, 2011; Plotnikov et al., 2012). Accordingly, in our experiment, we used a version of *Tetris* to manipulate the subjects' flow states.

Tetris seems to have many of the qualities necessary to induce a flow state. Primarily, it provides a balance between challenge and skill by starting out at a slow pace, and increasing in pace as the player clears lines. A more skilled player will clear lines more quickly than a less skilled player, so a more skilled player will find themselves at a higher piece fall rate, and thus a higher difficulty, more quickly than a novice player. In this way, *Tetris* ensures that the player is not in a state of boredom, where challenge is lower than skill. A game of *Tetris* will also quickly end once the player's skill is not sufficient to keep up with the challenge, as the pieces will pile up, the player will lose the game, and will be able to start a new game, reverting them to the initial difficulty level. This system keeps the player from being overwhelmed for a prolonged period by the challenge outpacing their skill. These aspects naturally cause the player to be at a balance of challenge and skill more often than not, which is why *Tetris* is an ideal starting point for flow research.

The version of Tetris used in this experiment has two conditions: a non-adaptive condition, which is unaltered from normal *Tetris* gameplay, and an adaptive condition, which has had two alterations made to better induce flow state. The first alteration is that subjects cannot get a game over which forces them to start a new game. When the accumulation (the pile of previously placed *Tetris* pieces while begins at the bottom of the game board) nears the top of the game board, the system will remove rows so that the accumulation is lower and the subjects will never lose the game. This system removes rows in order of the number of filled spaces in that row; rows that are mostly filled will be removed sooner than rows which are mostly empty. The second alteration is that subjects are able to control the speed the pieces fall at by pressing on a foot pedal, in a manner analogous to a car's accelerator. Thus, these alterations make the Tetris game adaptive, in the sense that the lower skill participants will be "helped" by the system and not be allowed to fail, while the higher skill participants could increase the speed to increase the challenge as they wish. These manipulations we apply to *Tetris* are intended to affect these characteristics of *Tetris* which are flow inducing. In the adaptive condition, they are intended to facilitate flow by facilitating challenge skill balance, while in the non-adaptive condition, they are intended to inhibit it in the same manner.

To test whether flow or any related effects are transferred from *Tetris* to another attentional control task, subjects do the Attentional Blink Task (ABT) (Taatgen, Juvina, Schippe, Borst & Martens, 2009) both before and after playing Tetris. If there was transfer of flow from Tetris to this task, then there should be a difference between performance on the first ABT and the second ABT. Additionally, we measured flow state before and after each task using the FSS2 and DFS2 (Jackson & Eklund, 2002).

We gathered data from 58 subjects, and discarded 6 subjects because of incomplete data or subjects not following instructions. Of the 52 analyzed, 18 were male and 34 were female, with age (M=20.40, SD=2.97). The likert-type flow report scales were normalized between 0 and 1, where 1 represents a subject fully in flow, and 0 represents a subject not at all in flow. Results were inconclusive; no significant difference was found between the flow/non-flow conditions in terms of difference in reported flow state, $t(50)=1.35$, $p=.18$, $d=0.374$ (see Figure 10).

Figure 8. Mean flow score difference by Tetris condition. Calculated as post-pre

Likewise, there was no significant difference in performance on the ABT for the flow/non-flow conditions, t(50)=-1.3361, p=.19, d=.371 (see Figure 11).

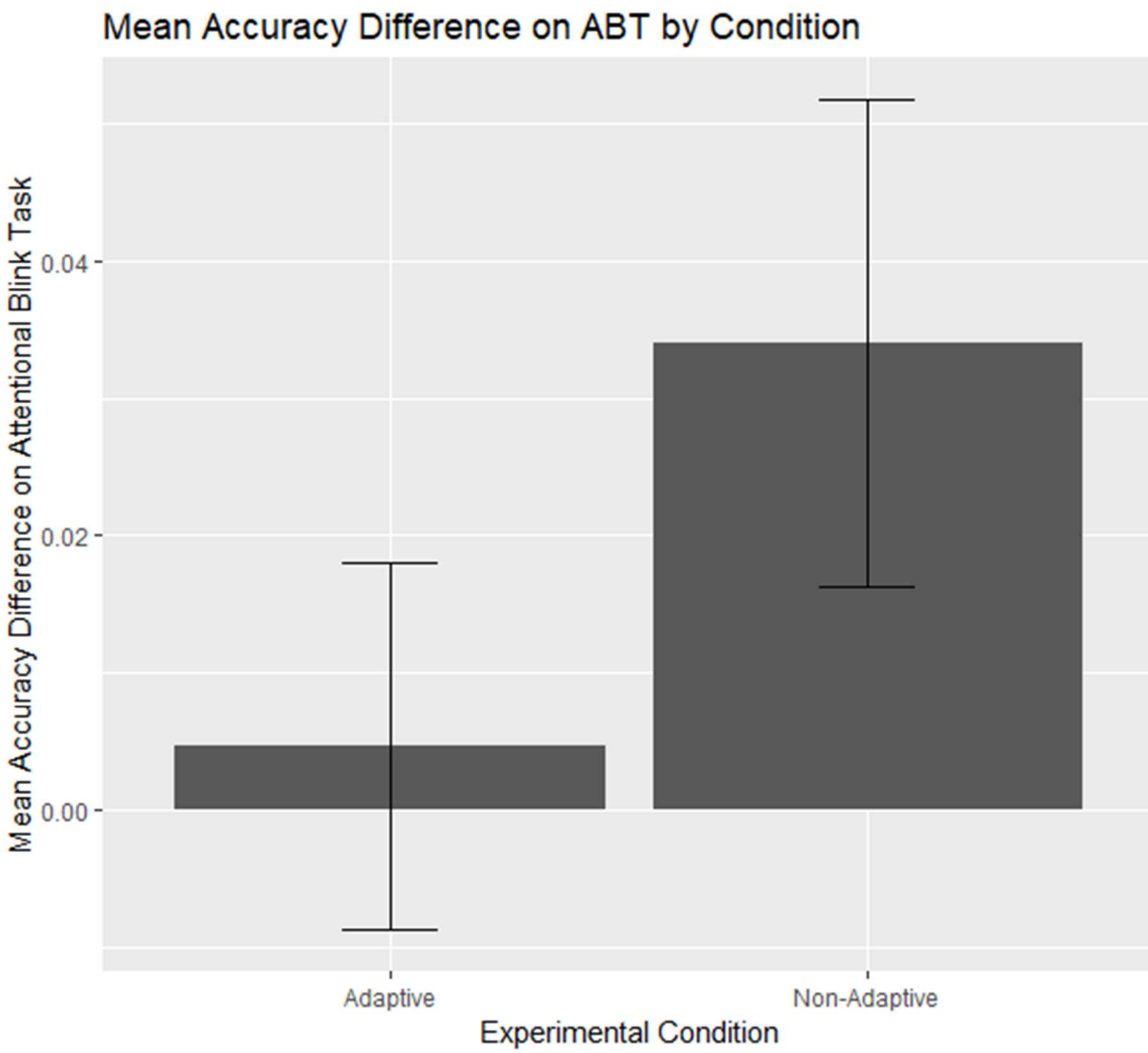

Figure 9. *Mean accuracy difference on Attentional Blink Task, by Tetris condition. Calculated as post-pre*

It seems probable that the conditions were not different enough to induce a significant effect. Only 4 subjects out of 27 in the adaptive condition actually pressed the foot-pedal for more than 500ms, and it seems likely that 20 minutes was not enough time for many of the subjects to reach a point in their game where the row removal system was necessary. Presumably, if these systems were never utilized, then the differences between the adaptive and non-adaptive conditions would be negligible, and a lack of any significance between them would be expected.

Experiment 2

The second experiment was designed to fix the above-mentioned problems with the first experiment. The mechanism to remove rows from the game board when the accumulation nears the top, which was present in the first experiment, has not been implemented in the second experiment. This was for several reasons. First, as mentioned above, it was noted that many players in the first experiment never reached the point where the accumulation was close enough to the top of the board that this mechanism was activated, and thus the accumulation could not have been a factor in any inducement of flow state they might experience, because it never occurred in their games.

Second, some players nearer the more skilled end of expertise in Tetris can reach the higher levels of Tetris within the time period. What level of Tetris is being played directly maps the speed that the pieces fall at, and when the pieces fall too quickly, the player does not have enough time to move them from the center position they are dropped from, leading to a tall narrow tower of pieces forming in the center of the game board. When this tower reaches the top of the game board, the row removal mechanism will remove the most filled rows first, which means that the accumulation the player has been working on will be mostly or entirely removed by the time the game speed is reduced to a level the player can manage once again. This would leave those players with a single narrow tower of pieces in the center of the board, requiring players to fill it in on the left and the right to get back to a normal manageable game board. This seemed to be non-conducive to flow, as the player is quickly thrust into a separate situation from the one they have been working on.

Thirdly, the piece-choice mechanism was likely a more productive avenue to pursue than the row-removal mechanism with regards to inducing flow state, since it more directly addresses the challenge-skill balance aspect central to flow. Because of these reasons, the row-removal mechanism has not been re-implemented in this iteration of the experiment.

An additional change was the method of altering difficulty for the subject. Instead of giving subjects the option of using the speed of the falling pieces to affect the task difficulty, the second experiment uses an algorithm that gives the subject Tetris pieces based on what would be better or worse for a given game state. We believe that this method of manipulating game difficulty allows us to manipulate challenge-skill balance much more precisely than was possible in the previous experiment, and is present whether or not the subject consciously decides to make use of it.

A corollary of this is that the preview box, which normally shows the next piece to be given, has been removed in the experimental tasks, since in those tasks the game hasn't determined which zoid will be given next until the player has placed the zoid in the current round.

To allow for easier implementation of this alteration of the piece-choice mechanism, the project was transitioned over to the *Meta-T* framework, a *Tetris* analog designed and optimized for behavioral science research (Lindstedt & Gray, 2015). This implementation has been designed to allow a great deal of control over game parameters, and gives richer log data than our previous implementation. This framework has also been used in a number of other experiments into expertise and cognitive-motor skills within the *Tetris* paradigm (Pilegard, & Mayer, 2018; Sibert 2015; Sibert, Gray, & Lindstedt, 2015; Sibert, Gray, & Lindstedt, 2017; Thompson, McColeman, Stepanova, & Blair, 2017), and our experiment will be building upon this work.

This modification involved sorting between all of the possible Tetris pieces, called *zoids*, to evaluate their usefulness for a particular board state. There are 7 unique zoids (see Figure 3), and for any given board configuration, it is possible to order them from "best" to "worst" based on a number of features of the game board. These features include things such as maximum height of the accumulation (the pile of zoids at the bottom of the game board), number of wells (holes in the accumulation which are open at the top), and number of pits (holes in the accumulation which are not open at the top). (For a more complete list, see Thiery and Scherrer, 2009). Higher or lower values among these features are considered better or worse for performance; for example, a good AI will minimize the maximum height, since the higher the accumulation is, the closer the AI is to losing the game. Likewise, numbers of pits and wells are to also be minimized, since pits and wells make it more difficult to clear rows.

Use of feature sets is a standard practice in AI approaches to Tetris (Fahey, 2003; Thiery & Scherrer, 2009; Sibert, Gray, & Lindstedt, 2015), and *Meta-T* has the ability to output and log these features, as well as an AI system which can determine a feature-based score for every possible placement for a given zoid and board configuration.

The modifications to the existing *Meta-T* framework consists of a subsystem which uses this AI to generate the potential feature score for each of the possible orientations and placements of each zoid, returning the maximum feature score for a possible placement, indicating the "goodness" of the most ideal placement of that zoid. In this manner, each zoid can be rated as being more or less ideal for the player to be given, and the system can choose to give better or worse zoids in order to manipulate game difficulty.

There are three difficulty conditions being used in this experiment, a "best" condition, a "worst" condition, and an adaptive condition. In the "best" condition, each zoid being given would be, when placed in the most ideal spot, the most ideal for the given accumulation, and presumably would lead to a player doing better at the game because of this. In the "worst" condition, each zoid being given, when placed in the most ideal spot for that zoid to be placed on that particular game board, is the least ideal for the given game board, and presumably would lead to the player doing worse at the game because of this.

In the adaptive condition, the piece being given scales from best to worst based on the maximum height of the accumulation (that is, the height of the highest block in the accumulation). When the accumulation is empty, or its maximum height is very near the bottom, the worst pieces are given, because the player is probably doing well and should be given more challenge. When the maximum height of the accumulation is near the top, the best pieces are given, because the player is doing poorly and could use some assistance. When the accumulation is between these maximum and minimum values, the "goodness" of the zoid given is linearly proportional to the height of the accumulation, binned into one of seven categories, one for each of the seven possible zoids.

Analysis of Conditions

While "adaptive" is descriptive of the actual functioning of the game in the condition, "easy" and "hard" are relative terms, and it is not necessarily assured that the manipulations which are assumed to correspond to decreased and increased difficulty actually do. To verify if this is actually the case, I ran several ANOVA's on different metrics of performance within the manipulation conditions. The three metrics of performance used were minimum path difference, initial latency, and feature score difference.

Minimum path difference

Minimum path difference is the number of unnecessary translations and rotations performed in placing a piece in a particular placement. It is computed by taking the actual number of translations and rotations performed by the player, and subtracting the minimum necessary translations and rotations to move the piece into the proper location and orientation for its final placement. This is indicative of performance because a player who is performing better will likely have a better idea of where they want to place a piece, and will generally move the piece to that point in a more straightforward manner. On the other hand, a player who is performing poorly will be less sure about where they want to place a piece, and may move and rotate the piece in ways not necessary to reach the eventual placement, either in aborted attempts to place the piece in a different position than the final position, or as a manner of spatial manipulation to aid in their determination of where the piece might fit. Cooper (1975) showed that the rotation of 2d shapes from an initial orientation is proportional to the time required to make a determination about whether its form matches another form. Given this, it seems reasonable that subjects, when given the ability to affect the rotation and translation of the object, would manipulate it in the world space as opposed to in the mental space, both to lessen mental processing load and decrease mental processing time, an important factor in a time-sensitive game such as Tetris. We would expect a player playing an easier version of Tetris would have a lower minimum path difference than a player playing a harder version, as they would make fewer unnecessary moves.

I conducted a one-way between subjects ANOVA and determined there was a significant effect of the manipulation of the Tetris task (i.e. easy, hard, adaptive) on performance on the minimum path difference metric, $F(2,9576)=39.17$, $p<.05$, $\eta^2=.0081$. A follow-up Tukey's HSD test determined that all conditions were significantly different from each other, with $p<.05$, and that the scores appear in the order expected, with “Easy” having the lowest minimum path difference, “Hard” having the highest minimum path difference, and Adaptive having a minimum path difference found between the two. This can be seen in Figure 12.

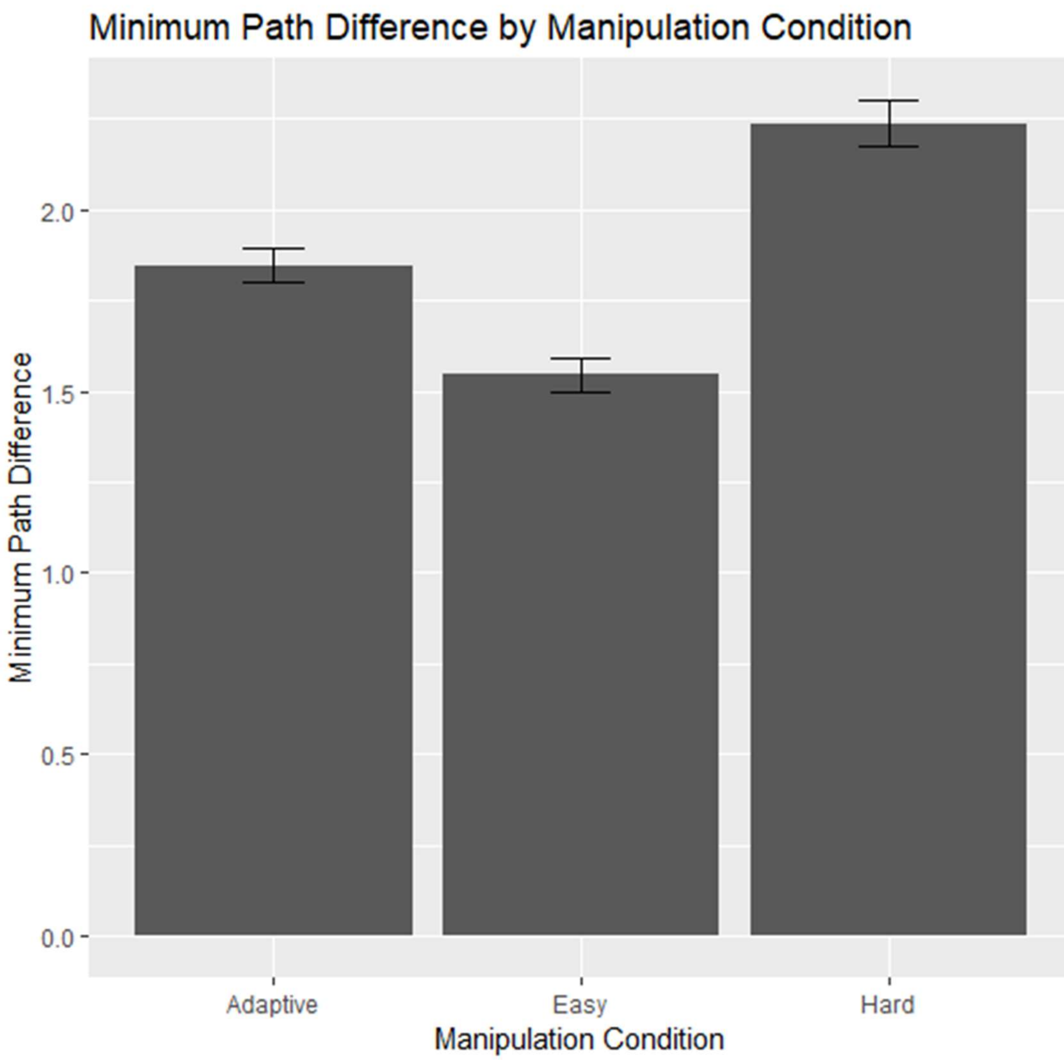

Figure 10. Difference of Minimum Path Measure by Tetris condition. Minimum Path Difference is defined as the number of unnecessary translations and rotations performed in placing a piece in a particular placement.

Initial latency

Initial latency is the second metric of performance examined, and is defined as the time (in ms) between initial presentation of a Tetris piece and when the subject first presses a button to begin manipulating the piece. A subject who is under higher mental workload would likely be unable to respond to the presentation of a new zoid as quickly as a subject who is under lower mental workload. This measure can be thought of as analogous to a reaction time measure. We would expect subjects playing an easier version of Tetris to have a lower initial latency than subjects who are playing a harder version of Tetris.

I conducted a one-way between subjects ANOVA and found a significant effect of experimental manipulation (easy, hard, or adaptive) of the Tetris task on the initial latency metric, $F(2,9576)=631.5$, $p<.05$, $\eta^2=0.12$. A follow-up Tukey's HSD found that all conditions were significantly different from each other. This can be seen in Figure 13. The "Easy" condition had the lowest initial latency, and the "Hard condition had a higher initial latency than the easy condition. However the adaptive condition had an initial latency significantly higher than both the "easy" and "hard" conditions. Its possible this due to the changing conditions of the Adaptive task, which means players may find it more difficult to generate strategies to place the pieces, since a strategy which is valid when the accumulation is low, and therefore pieces are likely more difficult to place, may not be valid when the accumulation is high, and the pieces would probably be easier to place.

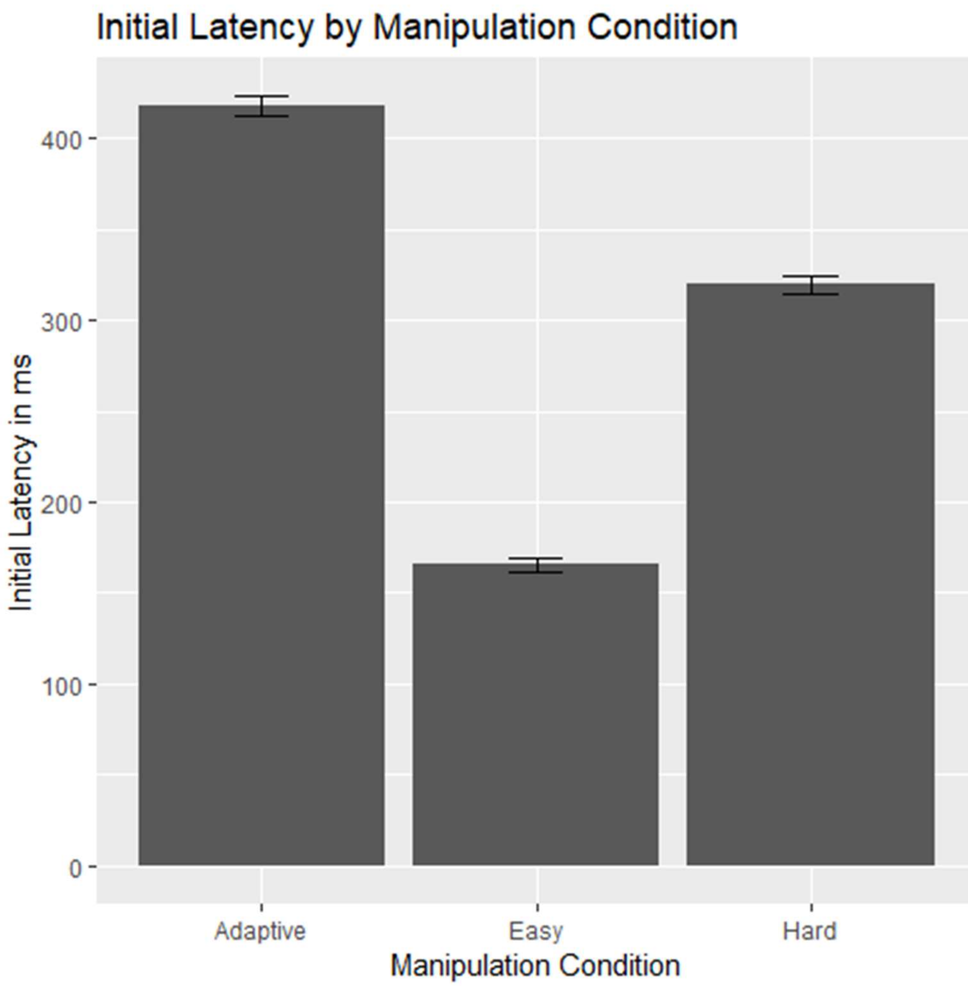

Figure 11. Initial Latency measure by Tetris condition. Initial latency is defined as the time (in ms) between initial presentation of a Tetris piece and when the subject first presses a button to begin manipulating the piece

Feature score difference

Feature score difference is the difference in weighted feature scores for the post-placement accumulations of the player's Tetris piece placement, and the placement of a zoid recommended by Meta-T's AI capabilities, in this case using the Dellacherie controller (Fahey 2003). This AI generates a weighted feature sum (as described above) for each possible placement, and then picks the placement which has the greatest score. If we take a difference score between the AI's placement score and the player's placement score, we get an indication of how well the player is playing compared to the difficulty of the current accumulation, which would be indicated by the magnitude of the AI move's feature score.

Lower numbers on this feature score difference measure indicate the player chose a "better" placement as evaluated by the AI. A score difference of zero would mean that the player either chose the same spot, or a spot with an equivalent feature score, meaning a spot which is just as good (as evaluated by the Dellacherie controller) as the one chosen by the AI. A negative score difference indicates that the player found a better placement of the piece than the AI did, which is possible because the AI only considers positions which can be reached solely by dropping a piece in a particular orientation from the top of the game board. Positions which require the player to slide a piece underneath a portion of the accumulation are not considered by this algorithm, in part because of the larger computational overhead required to evaluate all possible paths to a given placement on top of simply evaluating all orientations from the top of the board. However, even though it is possible for the player to score higher than the AI (and does occur in a minority of cases), the feature score difference measure is still indicative of the player's placement ability relative to the placement difficulty of a particular accumulation, and is still informative.

We would expect that a player in an easier Tetris condition would get lower scores on this metric than a player in a harder Tetris condition, as players with more available cognitive resources will be able to evaluate more possible placements, and will be more likely to choose a better placement, as opposed to one with fewer available cognitive resources, who would be more likely to satisfice when confronted with the constraints of the task.

I performed a one-way between subjects ANOVA and found a significant effect of the experimental manipulation on the Tetris task (either easy, hard, or adaptive conditions) on the feature score difference metric, $F(2,9576)=7.89$, $p<.05$, $\eta^2=.0016$. A Tukey's HSD follow-up test found significant differences between easy and adaptive, and easy and hard, not between hard and adaptive. It also found that easy had higher average feature score differences than hard, as expected, but that adaptive had lower feature score differences than easy and hard (although, as mentioned above, not significantly lower than the hard condition). This can be seen in Figure 14.

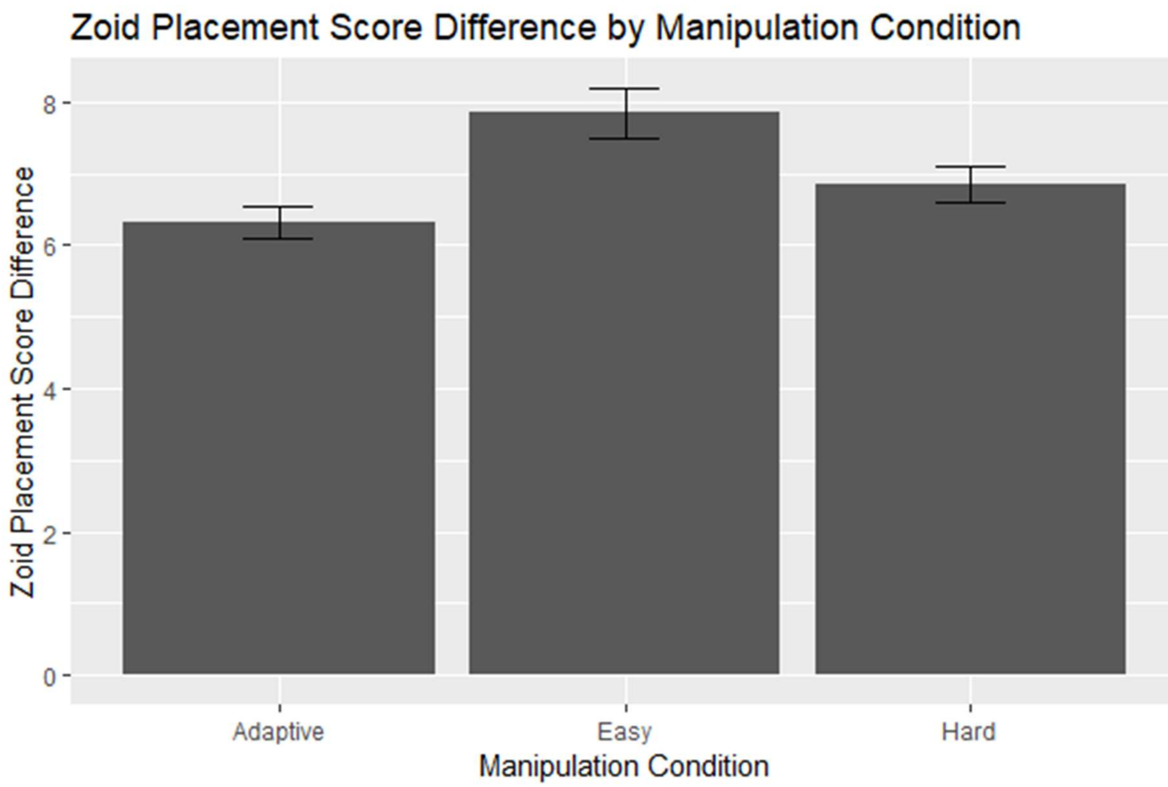

Figure 12. Zoid Placement Score Difference by Tetris Condition.

Zoid Placement Score difference is defined as the difference in weighted feature scores for the post-placement accumulations of the player's Tetris piece placement, and the placement of a zoid recommended by Meta-T's AI capabilities using the Dellacherie controller

It is important to note that each of the above is only a part of the whole story. Each measure is an indicator of overall performance, but does not describe the entirety of performance. The main measure we use as a more descriptive measure of performance is the game score, which is logical to use because it is also what the players themselves will use to determine their performance. As can be seen in the results concerning hypothesis 2 below, overall score in the three conditions, relative to individual performance on the baseline tetris task, behaves as expected, with players scoring relatively highest on the easy condition, followed by the adaptive condition, followed by the hard condition.

Experimental Design

This interaction between the subject and the piece-choice system can be conceived of as an instance of human-machine teaming, with the system being an AI agent. Both the player and the agent have an influence over the performance of the game; the player has control over where the pieces are placed on the game board, and the agent has control over what pieces will be given to the player based on its evaluation of the game board. It stands to reason that there are actions which the agent may make which improve or worsen the player's performance in Tetris, and that those actions which keep the player in flow (which are present in the adaptive condition) will lead to better performance than those which do not (in the easy or hard condition). This interaction can then be characterized as an instance of team flow, where the actions of the AI agent can have an effect on the flow of the player, and that cooperative interaction between the AI and the player will lead to better performance than an antagonistic interaction.

The experiment lasts three hours over two sessions. In the first session, subjects take the DFS, to get a measure of their trait flow, and then do a baseline *Meta-T* task for 50 minutes. This task gives subjects a random zoid each time, and is closer to the *Tetris* games some players might be familiar with than the experimental tasks are. The purpose of this task is to familiarize the subjects with the game and leave enough time for any practice effects to occur, as well as to gather data for the post-experimental task to be compared against.

When gaining expertise in a task, people will go through periods of learning, which will increase performance to a task. Eventually they will reach a point where their performance plateaus, and further performance improvements will not occur unless they alter their strategy, allowing them to transcend the ceiling imposed by their previous strategy (Ericsson, 2009; Gray & Lindstedt, 2017). Ideally, the first *Meta-T* task will give the players long enough time to reach a plateau, which would reduce the amount of learning effects across the experimental task and into the final *Meta-T* post task, hopefully diminishing its influence as a confound.

In the second session, subjects will do the experimental *Meta-T* task for one hour, in which they have been randomly assigned to either the "best" condition, the "worst" condition, or the adaptive condition. Following this task, they will take the FSS, to gain a measure of their state flow immediately following the experimental task. After this experimental task, they then play a second unaltered *Meta-T* game for 50 minutes, to determine if there is any transfer of flow from the experimental task, differing based on the different experimental conditions.

Performance on each Tetris task is evaluated using a "criterion score." This is a measure intended to remove some of the variation in Tetris score caused by the semi-random nature of gameplay. Sometimes players will get a disadvantageous sequence of zoids, which could cause them to have poor performance in a single game, which wouldn't be representative of their overall ability. To counter this, we take the average of the scores of a player's four highest-scoring games, and use it as a measure of performance. This method of score evaluation has been used in *Meta-T* research previously (Sibert, Gray, & Lindstedt, 2017).

We have generated two hypotheses with respect to outcomes of this experiment, relating to induction of flow state, and transfer to external tasks. We believe the manipulation which adaptively balances challenge and skill will lead to higher flow state than manipulations which produce a mismatch between challenge and skill. In this area, we believe we will find results similar to those found by Keller et al. (2011) and Tozman, Magdas, MacDougall, and Vollmeyer (2015), who found that a challenge skill balance led to a higher flow assessment.

We believe that, if a flow state or some correlate of such is transferred from the experimental task to the transfer task, there should be an increase in performance in the transfer task. Many studies suggest a relationship between flow state and improved performance (Eisenberger, Jones, Stinglhamber, Shanock,& Randall, 2005; Fullagar, Knight, & Sovern, 2013; Jackson, Thomas, Marsh, & Smethurst, 2001), with some (Jackson et al. 2001) positing a causal relationship between flow and improved performance. Therefore we believe that if a flow state has transferred from an experimental task to a transfer task, there should be increased performance on the transfer task relative to baseline.

Hypothesis 1: Subjects who play in the “adaptive” experimental condition will report higher state flow than those in the “easy” or “hard” conditions.

Hypothesis 2: Subjects who play in the “adaptive” experimental condition will perform better in the post task, relative to their performance in the pre task, than subjects who play in the “easy” or “hard” conditions.

Pilot data Thirty-five subjects total have been run, of which 13 were excluded due to incomplete data. The remaining 23 subjects were predominantly female (13 female, 10 male), of varying ages (M=21.56 years, SD=5.08).

To test hypothesis 1, an F test was performed to determine if there was a significant effect of the experimental condition on reported flow $F(2,20)=2.18$, $p=0.14$ (see Figure 15). No significant difference was found, likely due to the small sample size.

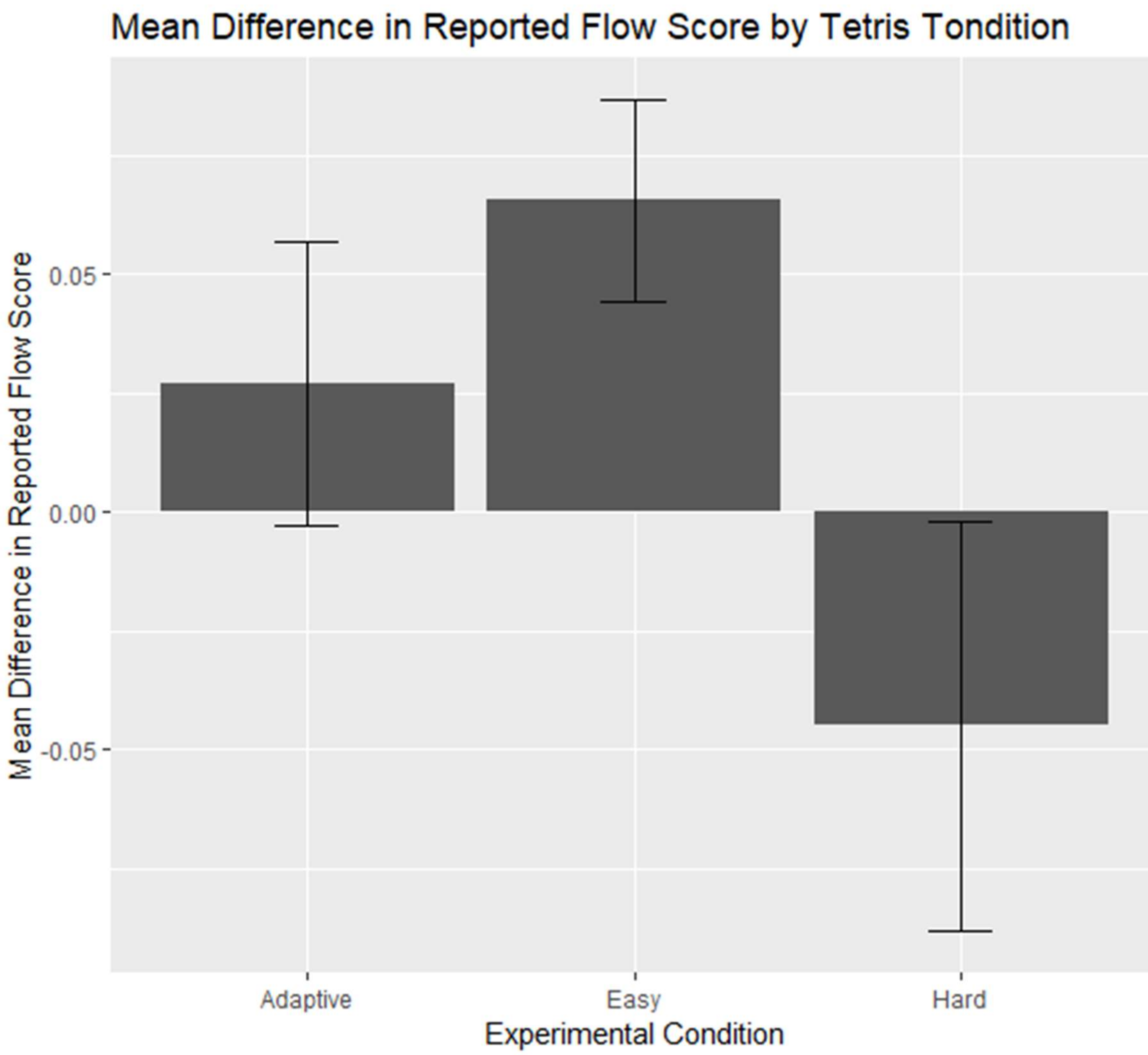

Figure 13. Mean difference in reported flow score by Tetris Condition. Calculated as post-pre

To test hypothesis 2, an F test was performed to determine if there was a significant effect of the experimental condition on difference in performance between the pre and post Tetris tasks $F(2,20)=0.32$, $p=0.73$ (see Figure 16). No significant difference was found.

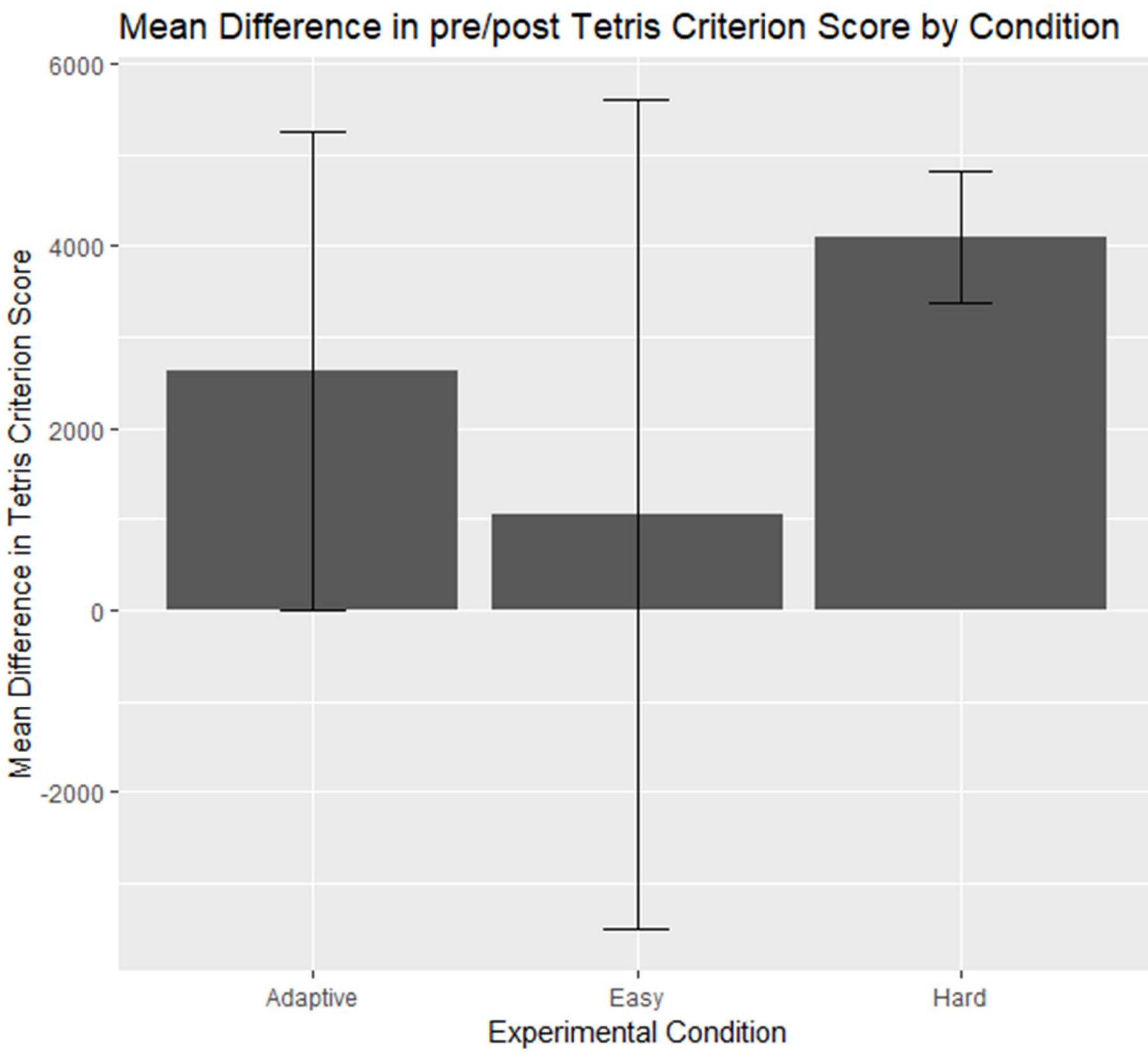

Figure 14. Mean difference in Tetris Criterion Score by Tetris condition, calculated as post-pre.

One of the concerns involved with an attentional control task such as Tetris is that of the subject experiencing learning effects, and improving their performance over time for reasons unrelated to flow. This is part of the motivation for having the subject play the baseline task for 50 minutes before experiencing the experimental condition, as our reasoning was that if there was a strong learning effect present in performance on the task, it would eventually plateau as players became acclimated to the task (Ericsson, 2009; Gray & Lindstedt, 2017). To determine if 50 minutes was a sufficient amount of time, we examined the learning curves of performance (both individual, and by experimental group) on the initial baseline task and the experimentally manipulated task. In these conditions, both in the final score curves as well as curves measuring minimum path difference, initial latency, and feature score difference, no clear pattern was found which would indicate a strong distinguishable learning effect. We reason that since there was no visible learning effect found within these conditions, it is likely not a strong effect overall, and its role as a confounding variable will likely be minimal.