Adaptive Interface Promotes a Composite of Performance and Flow in Tetris

Ion Juvina
Wright State University
3640 Colonel Glenn Hwy, Dayton, OH 45435
ion.juvina@wright.edu

Kevin O'Neill
Wright State University
3640 Colonel Glenn Hwy, Dayton, OH 45435
koneill1994@gmail.com

ABSTRACT
An interactive version of the game Tetris based on the Meta-T game is presented. In this version, a human plays the standard Tetrix game while an AI algorithm attempts to determine the player's skill and dynamically alter the game difficulty to match it. The AI determines which Tetris piece (called zoid) to give to the human by analyzing for each zoid every possible placement position on the current board state, and then ordering the zoids from easiest to hardest. An empirical study pitching this adaptive condition against easy and hard conditions show that the adaptive condition had a positive effect on a composite criterion made of 60% performance and 40% flow. Arguably, this is a realistic criterion for many human performance domains. The adaptive condition, powered by the AI algorithm, does well on this composite criterion because it avoids the pitfalls of the easy and hard conditions: the easy condition hurts performance while the hard condition hurts flow.

Author Keywords
Adaptive interface; flow; challenge-skill balance.

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms
Human Factors; Design; Measurement.

DOI: 10.37789/rochi.2022.1.1.20

INTRODUCTION AND BACKGROUND
Adaptive forms of instruction, in which the instructional material changes dynamically based on learner performance, abilities, needs, or other individual differences [19], can be more effective than corresponding non-adaptive instruction [1]. While adaptive technologies may impact learning and performance in multiple ways, one possible explanatory construct is the experience of flow [4]. Some adaptive task environments induce an experience of flow [6,10] and, in theory, flow is associated with increased task performance [4]. However the evidence for the link between flow and task performance is mixed [6,15].

Flow is a state of effective mind-environment coupling, experienced when people are fully immersed in a task, such as reading, writing, or computer programming. One of the antecedents of flow is the challenge-skill balance, a situation in which a person's skill matches the demands of the task at hand [4]. Teams achieving high degrees of coordination and performance have been reported to collectively experience flow as well [20]. Csikszentmihalyi [4] defined the autotelic personality as a disposition to enter the state of flow, distinguishing between flow as a personality trait and the flow state. Trait flow is a characteristic of an individual, whereas state flow is a more transient experience, affected by the current task and environment, as well as one's own trait flow [13]. Several studies have utilized the video game Tetris to study the experience of flow. For example, Chanel, Rebetez, Bétrancourt, and Pun [3] used Tetris as a task environment to induce flow and manipulated it by modulating the game difficulty through the falling speed of the pieces. They found that different levels of difficulty in the game Tetris lead to different levels of experience in participants, identified as anxiety, engagement, and boredom.

Flow has also been related to perceived and real cognitive effort. Subjects in a challenge-skill balance condition, who reported feeling in flow, generally also reported that their tasks required a lower amount of cognitive effort than subjects in a difficult condition, while the objective cognitive effort was maximal in the flow condition [8].

The goals of the work presented here were to explore a novel technique to induce flow and to study the relationships between the experience of flow, perception of cognitive effort, and task performance. We summarize the results of an empirical study that were reported in a master's thesis [14] and add the results of a new exploratory analysis that were not included in the master's thesis.

METHOD
We first describe the experimental task and the technique we used to operationalize the balance between the demands of the task environment and the fluctuations in performance as users completed the task. Then we describe the methodology of the study.

Task and Software
Tetris is one of the most popular video games of all time. The game involves rotating and translating a falling tetromino (a.k.a., zoid or Tetris piece), to place it either on the bottom of a two-dimensional game board, or on top of previously placed zoids on that game board. 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 completed row (or rows) will disappear, the pieces above 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 assures that the player will eventually lose, since the zoids will end up falling faster than any human's possible reaction time.

We developed an interactive version of the game Tetris by extending the Meta-T game [12,18], which offered high flexibility in manipulating the task, as well as a robust logging system. In this version, a human plays the standard Tetrix game while an artificial intelligence (AI) algorithm continuously assesses the player's skill and dynamically alter the challenge level of the game to match it. The AI determines on-the-fly which zoid to give to the human to place onto the Tetris board for each episode by analyzing every possible placement position and orientation of each zoid for the current board state, and then ordering the zoids from easiest to hardest. This ranking of zoids formed the basis of our experimental manipulation, which will be described in the Experimental Design section. In the baseline task, this was done randomly, as it is done in many implementations of Tetris.

The AI algorithm that sorted Tetris pieces into a best-to-worst list was a weighted feature-sum model using the Dellacherie set of features and weights [7]. 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 [2].

The Meta-T software can compute the features needed for zoid ranking 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 (i.e., 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.

Participants

A sample of 137 volunteers was recruited from the population of undergraduate students in Psychology at a medium-size Midwestern university through Sona Systems (https://www.sona-systems.com/) in exchange for course credits. The participants who had incomplete data (24) were excluded. 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 range of an undergraduate college population.

Experiment Design

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

Participants in the easy condition were always given the zoid that was easiest to find a good placement for, while players in the hard condition were always given the zoid that was hardest to find a good placement for. Participants in the adaptive condition were given a zoid that varied in its ease of placement, corresponding to the player's continuously assessed skill. In other words, the adaptive version of Tetris reacted in a dynamic manner.

To the extend that the AI was able to accurately diagnose and then coordinate with the player's skill, the expectation was that this adaptive manipulation would induce flow in the human player. The state of the Tetris board was used as a proxy for the participant's performance. Since the aim of the game was to complete rows, we considered a player who kept their accumulation low on the board as performing better than a player whose accumulation was 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 that corresponded to their current performance. If the player was doing well, and thus the accumulation was fairly low, the system would give them pieces that were 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 that 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. By using this technique, we implemented one of the finer modalities of adaptation, referred to as micro-adaptation in Landsberg et al. [11] and the "step loop" in the adaptivity grid [1].

Measures

We measured performance in the Tetris game, self reported dispositional and situational flow, and perceived task effort. These measures were obtained repeatedly to detect changes from a baseline (see Fig. 1).

Change in flow was measured by taking the difference between scores on the Flow State Scale (FSS) and the Dispositional Flow Scale (DFS) [9] for each participant to control for individual differences in dispositional flow. The DFS and FSS are nearly identical in structure and wording. 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. DFS was administered at the beginning of the experiment before the participants had performed any tasks.

The game score was used as a basis for computing player performance. 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. 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 in studies using the *Meta-T* software [18].

Change in player performance 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).

To understand how reported flow experience was related to perceived task effort, we introduced a simple effort rating to the experiment, which asked the subjects to give a rating from 1 to 10 on the difficulty of the task they had performed.

Procedure

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 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 flow was measured right before the pre task (via the DFS) and right after the experimental task (via the FSS). Additionally, participants were asked to report their perception of 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 1. In this diagram, ovals represent Tetris tasks, rectangles represent flow scales, and octagons represent task effort rating items.

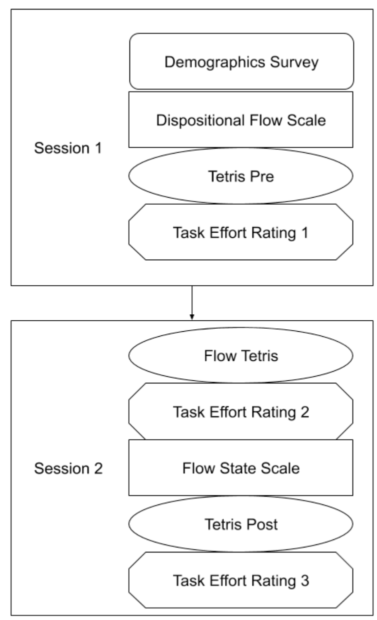

Figure 1. Diagram of the experimental protocol

Hypotheses

We expected that the adaptive manipulation would induce flow, as in other Tetris-based flow experiments from the literature [3,10], and flow would have a beneficial effect on performance in subsequent tasks which the subject might do immediately after experiencing a flow state, since a similar effect was reported with regard to other psychological phenomena such as near transfer [16] and persistence of mood [5].

We have also defined a more general hypothesis that integrated the relationships between the adaptive manipulation, effort perception, flow, and performance in a single statistical model (see Fig. 5). 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 AND DISCUSSION

Figure 2 shows that performance was very similar across all conditions for the pre task, as expected, since the participants were randomly assigned to the three conditions. Next, the easy condition produced higher scores than the baseline, and the hard condition produced lower scores than the baseline, which were the intended effects of those manipulations. The adaptive condition also produced lower scores than the baseline, but not as low as the hard condition.

A one-way analysis of variance (ANOVA) with the difference in performance between the pre and post tasks as a dependent variable and condition as a factor yielded non-significant differences between the three conditions, $F(2,110) = 2.03$, $p = 0.14$, $Eta^2 = 0.04$, even though the adaptive and hard conditions had non-zero performance differentials (see Fig. 3).

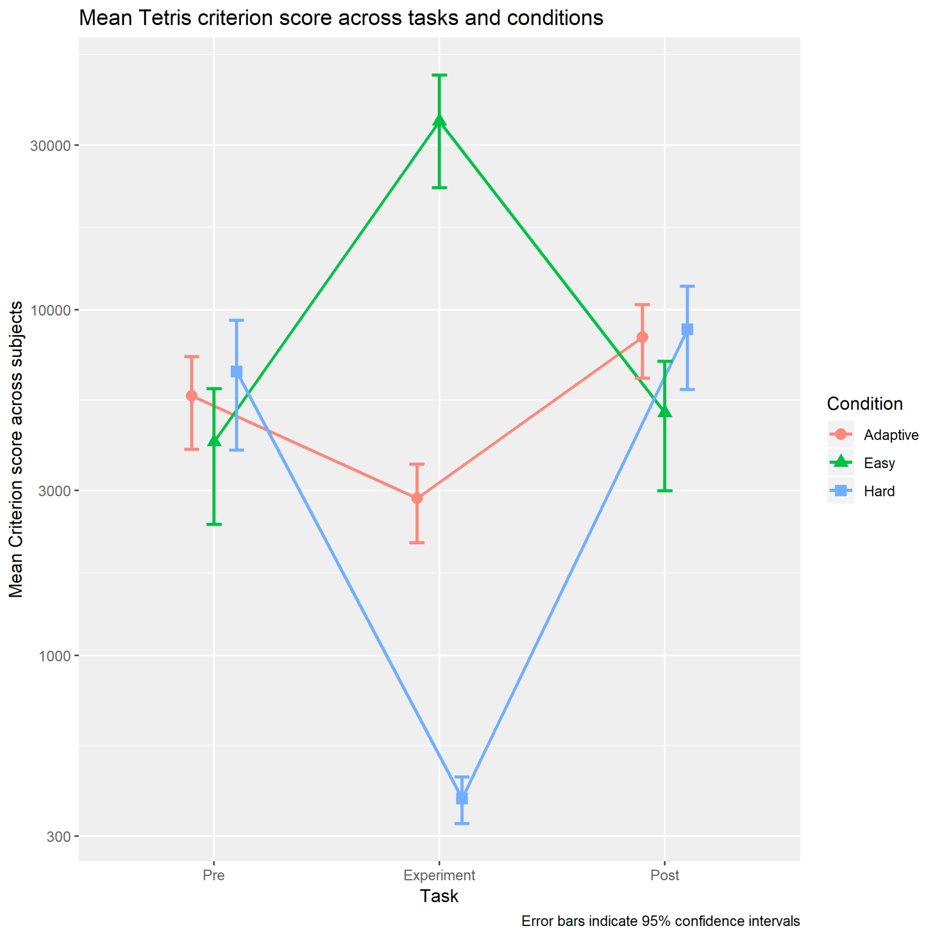

Figure 2: Time course of performance across condition and measurement repetition. Criterion score is plotted on a logarithmic scale because speed in Tetris increase as a function of time in a game.

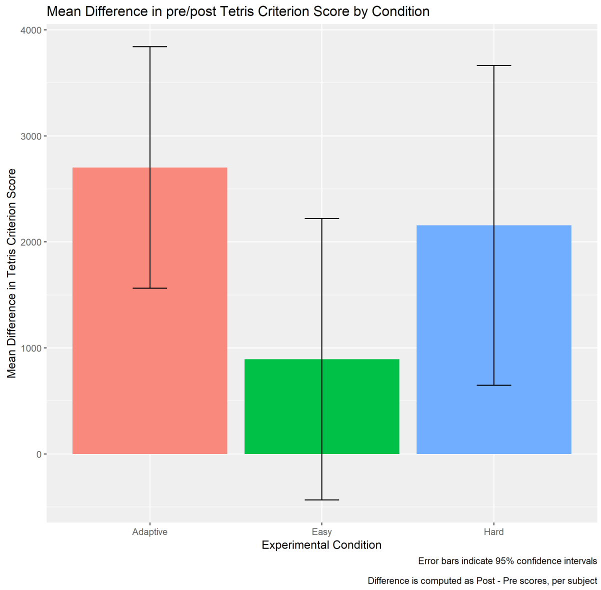

Figure 3. Performance differential by condition

We found a main effect of condition on the flow differential (i.e., DFS minus FSS), $F(2,110) = 19.37$, $p < 0.001$, $Eta\^2 = 0.26$. A Tukey's HSD follow-up test found significant differences between all pairwise comparisons of experimental groups, at $p < 0.05$. However, contrary to our expectations, it was not the adaptive condition but the easy condition that significantly induced a flow state above and beyond the participants' dispositional flow, while the hard condition actually hindered the participants disposition to experience a state of flow (see Fig. 4).

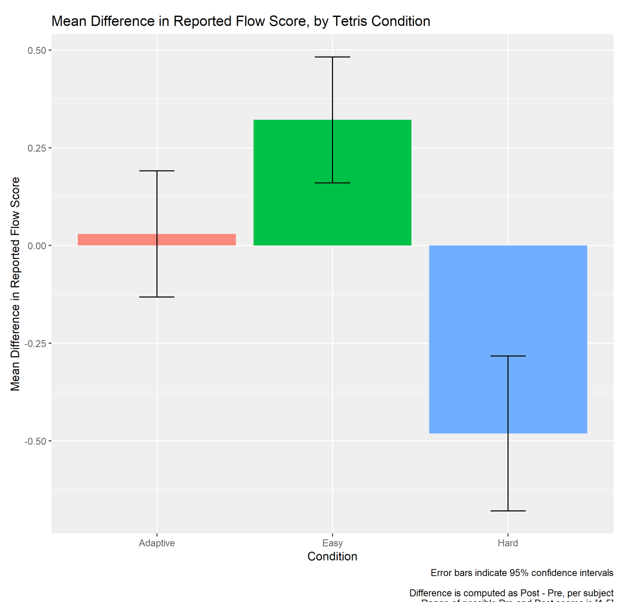

Figure 4. Flow differential by condition

Theoretically, an easy task should lead to low reported flow, as the challenge of the task would be lower than the skill of the player, and would result in boredom. 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 players 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 info indicated that a majority of the subjects identified themselves as novices at Tetris, lending some credence to this possibility.

Participants in the easy condition reported slightly lower levels of perceived task effort; however, the difference did not reach statistical significance, $F(2,110) = 0.6$, $p = 0.14$. The perceived task effort in the adaptive condition was as high as in the hard condition.

To understand the relationships between the adaptive treatment, flow perception, task performance, and effort perception in a more holistic way, we took a structural equations modeling (SEM) approach. We fitted a hypothesized structural model (see Fig. 5) against the empirical data.

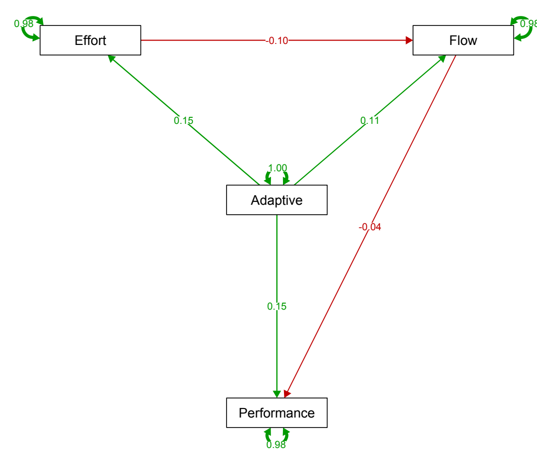

Figure 5. A structural equation model showing the relationships between the *adaptive* treatment, *flow* and *effort* measures as mediators, and *performance* as dependent variable. Numbers are standardized path coefficients.

Even though none of the direct, indirect, or total effects were significant, some of the effects were trending in the expected direction. The direct effects of adaptivity on flow and performance, respectively, were positive, as expected. In addition, the effect of perceived effort on flow was negative, as expected. In contrast, the effect of adaptivity on effort perception trended in the unexpected direction: the adaptive treatment did not reduce but increased effort perception. Lastly, the expected positive effect of flow on performance was virtually zero.

These results are tentative and in need of replication with a larger sample; if proven correct, they challenge some of the assumptions from the literature and suggest a different approach. The finding that performance and flow are uncorrelated suggests that they could be seen as two different and potentially complementary outcomes. The adaptive condition may have a positive impact on a composite outcome formed by combining performance and flow scores. The composite outcome can be a weighted combination of performance and flow.

It would be helpful to understand the effect of adaptivity on such a composite criterion. A weighted sum of normalized flow and performance scores was computed. The raw values were normalized by scaling their ranges between 0 and 1. The normalized values were then added together, with their proportion varying as a free parameter. Then we determined the value of this parameter that yielded the greatest difference between the adaptive condition and the other two conditions.

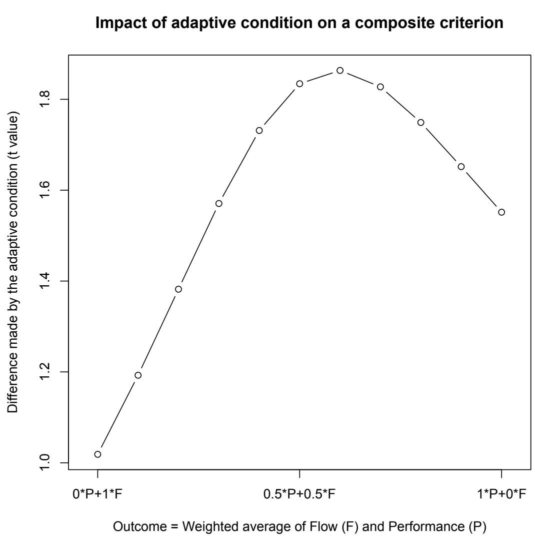

Figure 6. The impact of the adaptive treatment on a composite criterion that includes various proportions of flow and performance.

As can be seen in Figure 6, when different weights for performance and flow are considered, the adaptive condition appears to have a larger positive effect on a composite criterion that is loaded more with performance than with flow. Specifically, the highest positive impact is obtained when the criterion is composed of about 59% performance and 41% flow, $t(98.5) = 1.86$, $p = 0.065$, $d = 0.35$ (small-medium effect size).

CONCLUSION

While these results are preliminary, they suggest that the adaptive treatment has a positive effect on a composite criterion made of 60% performance and 40% flow. Composite criteria have practical value in many human performance domains [17]. Arguably, this is a realistic criterion, as designers of instructional applications aim to increase both performance and engagement. The adaptive version of Tetris presented here, powered by the AI algorithm, does well on this composite criterion because it avoids the pitfalls of the easy and hard conditions: the easy condition hurts learning and performance while the hard condition hurts flow.

ACKNOWLEDGMENTS

IJ’s contribution to this work was supported in part by The Office of Naval Research grant number N00014-16-l-2047 P00004 to Brandon Minnery, Ion Juvina, and Assaf Harel.

REFERENCES

1. Aleven, V., McLaughlin, E. A., Glenn, R. A., & Koedinger, K. R. (2017). Instruction based on adaptive learning technologies. In R. E. Mayer & P. Alexander (Eds.), *Handbook of research on learning and instruction* (2nd ed., pp. 522-560). New York: Routledge.
2. Algorta, S., & Şimşek, Ö. (2019). The Game of Tetris in Machine Learning. *arXiv preprint arXiv:1905.01652.*
3. 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.
4. Czikszentmihalyi, M. (1990). Flow: The psychology of optimal experience. New York: Harper & Row.
5. Ekman, P. (1984). Expression and the nature of emotion. In K. Scherer & P. Ekman (eds), *Approaches to emotion*. Hillsdale, NJ: Erlbaum, pp. 319-344.
6. Engeser, S. & Rheinberg, F. (2008). Flow, performance and moderators of challenge-skill balance. *Motivation and Emotion, 32*(3), 158–172.
7. Fahey, C. P. (2003). Tetris AI, Computer plays Tetris. http://colinfahey.com/tetris/tetris_en.html
8. 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), 103–114. https://doi.org/10.1037/spy0000083

9. 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.

10. 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.

11. 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.

12. 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.

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

14. O'Neill, K. J. (2020). *Induction and Transferral of Flow in the Game Tetris* [Master's thesis, Wright State University]. OhioLINK Electronic Theses and Dissertations Center.

15. Palomäki, J., Tammi, T., Lehtonen, N., Seittenranta, N., Laakasuo, M., Abuhamdeh, S., Lappi, O., & Cowley, B. U. (2021). The link between flow and performance is moderated by task experience. *Computers in Human Behavior, 124,* Article 106891.

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

17. Schmidt, F. L., & Kaplan, L. B. (1971). Composite vs multiple criteria: A review and resolution of the controversy. *Personnel Psychology, 24*(3), 419–434.

18. 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.

19. 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.

20. van den Hout, J.J.J., Davis, O.C., & Weggeman, M.C.D.P. (2018). The Conceptualization of Team Flow, The Journal of Psychology, 152:6, 388-423.