Computational Brain & Behavior (2018) 1:215–227
https://doi.org/10.1007/s42113-018-0014-4

The Tortoise and the Hare: Understanding the Influence of Sequence Length and Variability on Decision-Making in Skilled Performance

Catherine Sibert[1] · Wayne D. Gray[1]

Published online: 13 November 2018

Abstract

Tetris is a complex task which taps into several human skills: among them perceptual learning, planning, motor skills, and sequential decision-making. Following a divide-and-conquer strategy, we adopt a machine modeling approach to isolate the contribution of sequential decision-making from the other three skills. In two studies, we test three sets of 1,771,561 feature-based machine players (MPs) (11^6, 11 weights for each of 6 features) of Tetris for both long-running (tortoise) and short-running (hare) MPs. Tortoise models run until they die. Hare models are stopped after 506 episodes. For both studies, we select the longest running tortoise model and compare its score and behavior with that of the best scoring hare model. The best tortoise models adopt an endurance strategy which emphasizes single-line over multi-line clears. The best hare models adopt an escalation strategy which stresses multi-line clears. In contrast, our human players tend to adopt the escalation strategy early in their game but switch to the endurance strategy as speed demands increase. Unexpectedly, across three model runs, each with a different random seed, we obtain three different sets of "best fitting" models, that is, the MPs overfit the data even though that data is generated by an essentially infinite random sequence. However, in each model run, the best tortoise adopted the endurance strategy and the best hare adopted escalation.

Keywords Tetris · Sequential decision-making · Expert performance · Expertise · Overfitting · Cue and category validity · Category learning · Strategies · Methods · Machine modeling

Introduction

Slow and steady wins the race. This principle, oft quoted to impatient children, is the moral of Aesop's fable of *The Tortoise and the Hare*. In the story, the tortoise and the hare have a race. When the race begins, the hare quickly pulls ahead, while the tortoise makes slow progress. As the race progresses, the hare grows tired and takes a rest before finishing the race. As the hare rests, the tortoise crosses the finish line. The lesson for children is clear: constant hard work will eventually lead to success, while overconfidence can lead to costly mistakes. But, had the race course been shorter, the hare would have reached the finish line before getting tired, the clear winner. It is only due to the external environment, the length of the race, that the tortoise comes out victorious. These interactions between behavior and environment provide insight into human performance in Tetris and highlight issues that may arise when modeling a task without carefully considering the constraints imposed by its task environment (Anderson 1991a, 1991b; Simon 1991, 1992).

Tetris is a dynamic, puzzle game created in the 1980s. Today, it is available on nearly every computerized platform, making it one of the most played games in the world (Stuart 2010). Tetris has a rich history of use in psychological studies (Mayer 2014) most often exploring the side effects of its use as a treatment condition for topics as diverse as sex differences in spatial skills (Terlecki et al. 2008), relieving PTSD flashbacks (Holmes et al. 2009), or improving the abilities of engineering students (Martin-Gutierrez et al. 2009). It also presents a challenge to the machine modeling community (Fahey 2015; Szita and Lorincz 2006), who view it as a way of testing methods of model search.

In contrast, we consider Tetris as a research paradigm (Lindstedt and Gray 2015) for cognitive science theories, that

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s42113-018-0014-4) contains supplementary material, which is available to authorized users.

✉ Catherine Sibert
siberc@rpi.edu

[1] Cognitive Science Department, Rensselaer Polytechnic Institute, Troy, USA

is, as an example of *Games as Experimental Paradigms* (Game-XP) (Gray 2017). In common with the *extremely simple paradigms* (ESPs) used in most behavioral science laboratories, including our own, Game-XP may focus on basic processes of cognition, perception, action, dynamic decision-making, method discovery, and/or skill acquisition. Game-XP occupies the niche between typical psychology laboratory paradigms and real-world tasks; in common with Chess, most computer games cannot be mastered in an hour laboratory session and skilled performance may continue to increase over hours, days, years, or decades of practice. However, also in common with Chess, various levels of skilled performance exist in the college-age population making it possible to sample behavior across a broad-range of expertise (i.e., cross-sectional analysis) (see Gray 2017; Stafford and Dewar 2014, for a discussion of these issues).

Tetris is a complex task which taps into at least four human skills, namely, perceptual learning, planning, motor skills, and sequential decision-making. For human players (HPs), multiple skills must be deployed together in a dynamic task environment. In this report, we follow a divide-and-conquer strategy by adopting a machine modeling approach that simplifies the Tetris task, seeking to isolate the importance of the sequential decision-making skill. These simplifications enable us to focus on game length per se as an important factor in determining the optimal strategy for game play. That is, we will show that Machine Players (MPs) who are allowed to play until they die follow an endurance strategy whereas those MPs who play human length games adopt an escalation strategy. In contrast, we will also show that HPs tend to favor the escalation strategy during most of the game, but deploy the endurance strategy when the game speeds up so much that player expertise in other skills (e.g., perceptual learning, planning, and/or motor skills) do not suffice to keep pace with the game.

Background

During a game of Tetris, players navigate a sequence of seven shapes, each constructed of four connected squares that fall from the top of the screen into a pile at the bottom. Whenever a row is filled across the entire width of the screen, the blocks in that row or line vanish, lowering the height of the pile (A game in progress is shown in Fig. 1a, left side.). A game of Tetris is lost when the pile reaches the top of the screen, which can take as much or as little time as the player's skill allows. During human play, there is always one active piece on the board, and the next piece in the sequence is visible in the Preview Box (on the upper right portion of the screen). The number of points (Score), number of lines cleared, and current level of Tetris play are shown on the right side of the Tetris board.

Points are earned with each line that is cleared. The point value increases with (a) the number of lines cleared at once (as shown on the right of Fig. 1a) and (b) the game level. At level zero, clearing a single line is worth 40 points, while clearing four lines at once, a maneuver known as a "Tetris", is worth 1200 points, 7.5 times more points than clearing each of four lines, one line at a time. With every ten lines cleared, the game increases in level, and as the level increases, so does the point value of each type of line clear.

Tetris as a Game for Human Players

At level zero, it would take a zoid 16 s to fall from top to bottom. The speed at which zoids fall increases for each of the first 9 levels. Above level 9, speed increases are more sporadic and top out at level 29 where zoids fall from top to bottom in 1/3 of a second. The element of time pressure is a huge difference between the Tetris experience of HPs and MPs.

To be successful in Tetris, a player must score as many points as possible by clearing lines until their mistakes pile up too high or until the pieces fall too quickly to be moved into position. Whether HP or MP, no matter how skilled the player is, eventually they will lose.

Most humans play Tetris by shifting between two strategies for clearing lines: the endurance strategy or the escalation strategy.

(A) For the endurance strategy, players attempt to clear one line at a time, usually the bottom line. For this strategy there is little strategic interest in accumulating zoid blocks above the bottom row. Rather, the focus is on "plugging the gaps" by filling and clearing the bottom row. This approach minimizes the role of planning and minimizes risk by keeping the board height as low as possible.

(B) The escalation strategy is generally used, especially in the early game, by all-but-novice HPs. As clearing multiple lines results in an escalation in points awarded, this strategy attempts to clear two lines, three lines, or four lines at once. Players who opt for this strategy often try to shape the board cumulation by attempting to fill in all cells except, for example, those in the right-most column (see the right side of Fig. 1a). This human strategy is risky as it calls for a higher average board height than does the endurance strategy.[1] The risk is that the player may lose control of the accumulating pile of zoids and

[1] Attempting to keep, say, the right-most column clear while completely filling up the first four rows of the other nine columns is a common, human tactic. This configuration allows the right-most column to be plugged by an I-beam zoid, clearing four lines at once, and is called "a Tetris." However, building this configuration requires getting zoids that would leave pits (e.g., see Fig. 2) if placed in the first four rows out of the way, by placing them on the left-side of the top of the growing pile, while using other pieces to build the "four high wall," then waiting for the I-beam to appear, and finally placing the I-beam into the (at least) 4 deep slot, after that wall is built.

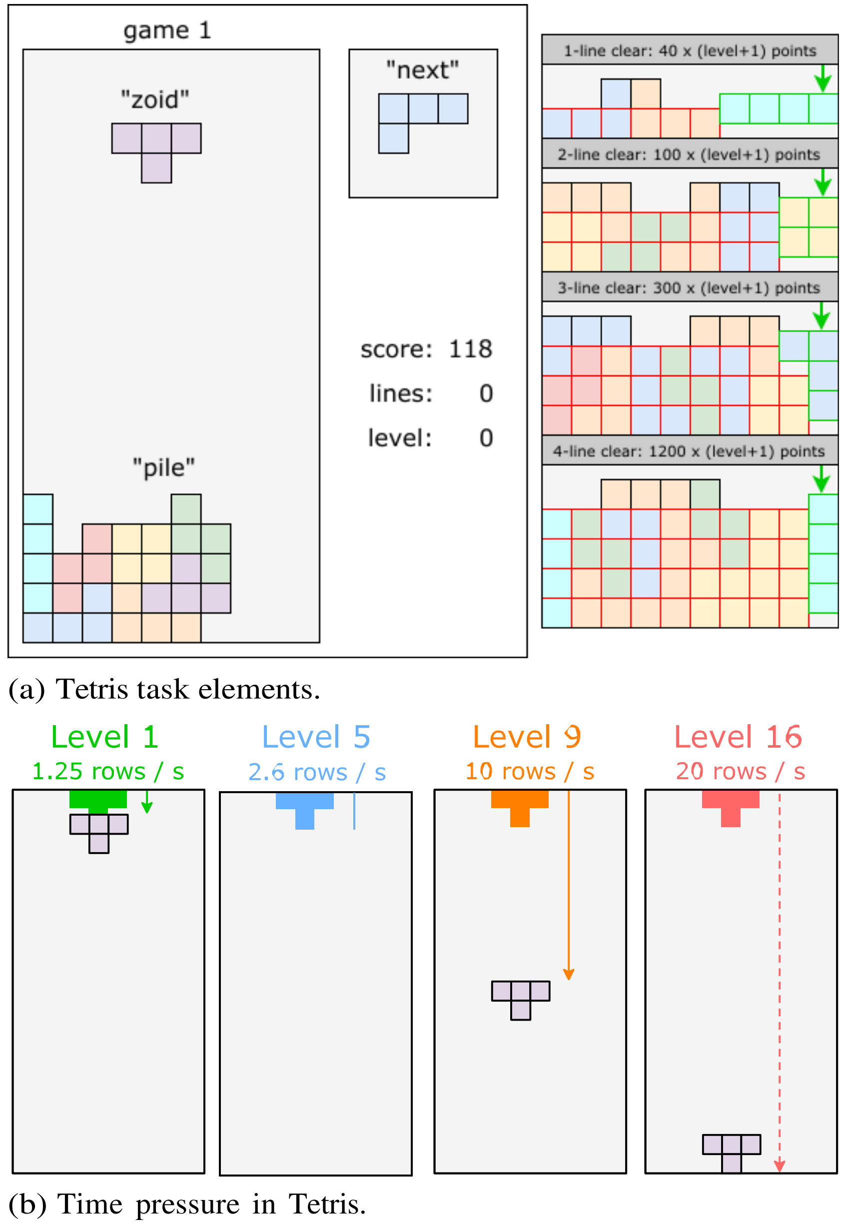

Fig. 1 Elements of the game of Tetris. **a** (left) shows the "zoid"—the active game piece that falls from the top of the screen and can be manipulated by the player; the "pile"—the accumulated zoid segments at the bottom of the game space; and the "next" zoid—in its preview box. The current score, lines cleared this game, and current difficulty level are displayed, as well as the current game number in the session. **a** (right) shows examples of filling and clearing 1, 2, 3, and 4 lines. Point values scale for both the type of line clear and the difficulty level. **b** shows the "fall rate" increasing as a function of the current game level. (Thanks to John Lindstedt for permission to use these figures)

then lose the game when the height of the pile exceeds the height of the Tetris board.

As we discuss later, HPs generally do not adhere completely to either one of these strategies, but tend to start their game with the escalation strategy and switch to the endurance strategy as the game speeds up. However, throughout most of their game, expert HPs resemble hares (Lindstedt and Gray 2018; Sibert, Gray, and Lindstedt 2017), opting to set up the riskier but higher scoring moves over the lower scoring but safer one-line clears. An example strong player, playing an hour of gameplay in our lab, cleared 27% of the total lines with single-line clears, 20% with two-line clears, and just 4% with three-line clears. Nearly half of the total lines cleared, 49%, were cleared with four-line clears. This u-shaped function shows the premium that many HPs place on clearing four lines at once.

Tetris as a Game for Machine Players

From a machine modeling perspective, Tetris provides a rich testbed for exploration. It is reasonably straightforward to create a model that plays Tetris well. These models are usually feature-based models that identify some structural elements of the board state and assign each one a weight corresponding to how desirable or undesirable a high value of that weight is. For example, a game of Tetris ends when the pile of pieces reaches the top of the screen, therefore a model that emphasizes a high value for board height would probably not be successful.

The challenge, then, comes not from creating the model itself, but finding the optimal set of weights that will produce the best performing model. It is this challenge that has interested the machine modeling community, and there exist many papers (e.g., Baccherini and Merlini 2008; Gabillon et al. 2013; Szita and Lorincz 2006; Thiery and Scherrer 2009a, b) detailing methods of feature-space search that produce high-performance Tetris models. However, what machine modelers consider high performance is not necessarily the same as what HPs consider high performance, and in their search for the best method of feature-space search, critical changes were made to the game environment of Tetris that make it nearly a separate task than what HPs face.

In a human game, the steadily increasing time pressure is a major part of strategy and performance. Playing 500–600 zoids would be a long game for a human as the increasing time pressure makes it harder, in the time available, to identify good zoid placements and move zoids to those locations. In contrast, for MPs that can objectively evaluate every possible placement of a piece and instantly move the piece to that location with no fear of a movement error, time pressure is irrelevant, and machine models of Tetris simply eliminate time pressure as an environmental factor. With time pressure eliminated, there is no real gain to scoring points early by clearing 3 or 4 lines at once. Instead, the best Tetris MPs play very long games (> 100,000 zoids) by adopting the slow and steady endurance strategy of clearing one or two lines at a time. This results in models that reach levels and scores far beyond the ability of any HP.

Divide-and-Conquer

For reasons discussed above, the methods and strategies used by HPs are more complex and more difficult to study than those used by MPs. Our initial attempt to describe the range of issues and solutions influencing human Tetris play is discussed in Lindstedt and Gray (2018). Adopting a divide-and-conquer approach, in this paper, we focus on what MPs do best, namely, make decisions on where to place a given zoid, one zoid at a time.

Most [but not all, see Şimşek et al. (2016)] MPs of Tetris are created to explore the utility of machine modeling techniques to converge on an optimal set of weights for a preselected set of features (e.g., Szita and Lorincz 2006; Thiery and Scherrer 2009b). Such models calculate the best board placement for each successive zoid, one zoid at a time, based on a small set of (usually 6) features. For Cross-Entropy Reinforcement Learning (CERL) models, what is learned across many generations of game play and feature weight tunings is an optimal weighting of each feature in terms of each other feature in the set—where optimal is defined by the highest score (see Sibert et al. 2017, for a more detailed discussion). For each zoid, the MP calculates the placement score for each possible location. The position location with the highest score is the move chosen (ties are decided by a random choice).

In Fig. 2, we show two of the 34 possible locations where the T-zoid can be placed by rotating and/or moving it laterally (referred to in Tetris as a "transposition"). For each of the six Dellacherie features, Table 1 illustrates the different feature values that arise from these two possible moves.

The landing height of the active piece, defined as the number of rows from the bottom of the board that the active piece comes to rest on, is shown as a red line in Fig. 2, column 1, landing height. For move A, the landing height is higher than that of move B, indicating an overall higher pile. As the model weighs landing height as a negative value (see Table 1 row 1), the higher the height, the more negative the placement.

The second feature, eroded cells, is defined as the number of cells that will be removed from the board once the piece placement is executed, and is illustrated as an orange bar, in move B, over the cells that will disappear. For move A, no lines are cleared, and so the value of eroded cells is zero. A single line is cleared in move B, bringing the feature value to ten.

For each row and column, a "transition" occurs when a filled cell is adjacent to an empty cell. Column 3 of Fig. 2 shows row transitions and column 4 shows column transitions (also see the corresponding rows in Table 1). In the figure, these transitions are illustrated by yellow triangles for rows and green triangles for columns.

The fifth Dellacherie feature is pits, shown as blue squares. Pits are any empty cell(s) that are completely surrounded by full cells. As the row in which the pit occurs cannot be removed until the pit is filled in, clearing a pit requires removing pieces that surround it, thereby opening the cell, so that the pit may be filled. For humans, filling a pit can be a complex maneuver that requires planning and action across many zoids. Hence, pits can lead to high boards that eventually overflow and end the game. In our Table 1 example, pits are assigned the highest negative weight.

The sixth and final feature is cumulative wells, illustrated as purple u-shapes. A well is an empty cell or series of empty cells that are surrounded on three sides by filled cells and open at the top. Wells are weighed by their number and depth. Moves A and B both have two wells, each with a depth of one.

For each candidate move, the feature values are multiplied by the weight of the corresponding feature, and then summed to form an overall score for each possible move. The move scores for move A and move B are calculated in Table 1, using the feature weights chosen by Dellacherie (Fahey 2015). As move B's weight is higher than move A's, it would be preferred. However, it is important to remember that, unlike the two alternatives discussed in our

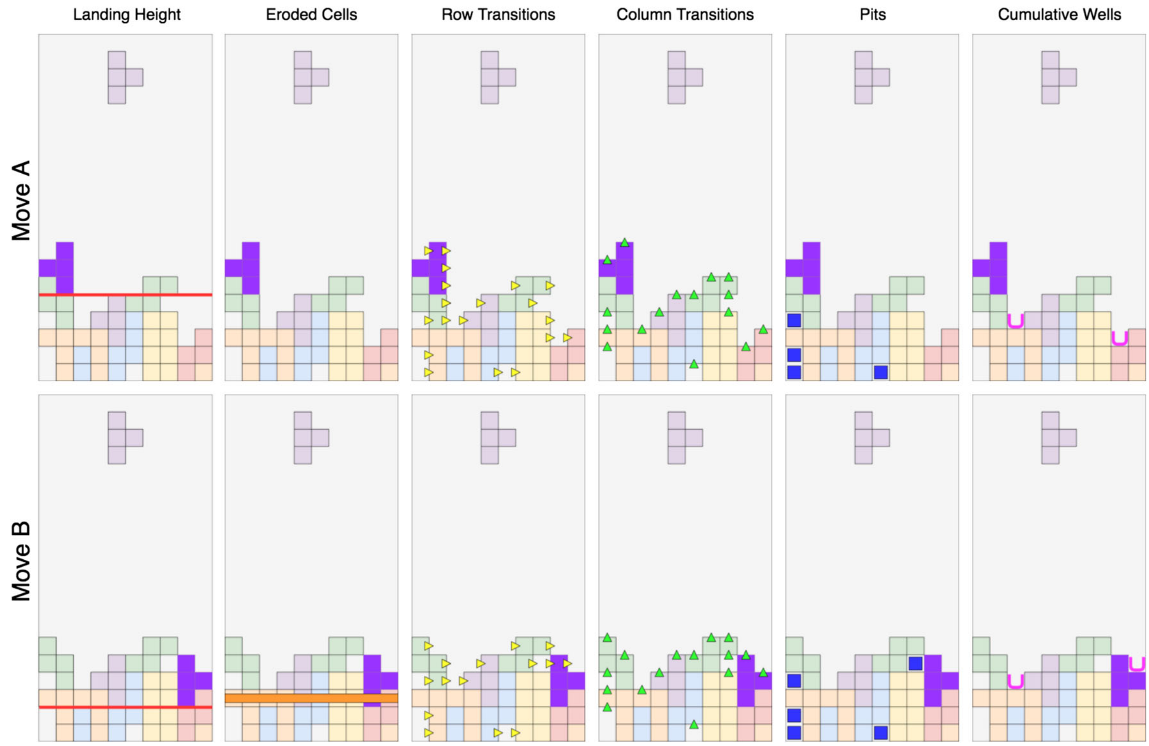

Fig. 2 Two alternative placements (top and bottom) for the same zoid on the same Tetris board. Each column provides a visual example of how the value assigned to that feature would vary across the two sample placements. Note that for each board position, there are between 9 and 34 possible zoid placements. See text for an overview discussion of how the two placements affect the value for each feature in this example

example, placement choices may require the consideration of up to 34 alternatives.[2] It is also the case that this set of feature weights represents just one of the 1,771,561 MPs tested by our studies.

Previous Tetris Modeling Work

In Sibert et al. (2017), we used CERL (Szita and Lorincz 2006; Thiery and Scherrer 2009b) and imposed a rough approximation of human limits by capping the length of each game to 506 zoids. This seemingly arbitrary number was chosen because it was the length of the longest game played by a human in our laboratory. The goal of that work was to explore how the objective function changed the behavior of the resultant models. For these 506 zoid games, models given the objective function of maximizing the number of lines cleared adopted the endurance strategy: clearing no more than one or two lines at any point in the game. For this strategy, the scores were consistently lower than the other strategies we tested. Models that were encouraged to maximize score adopted the escalation strategy: attempting to set up and execute higher scoring three- or four-line clears (remember that as Fig. 1 shows, one four-line clear yields 7.5 times as many points as do four one-line clears). The escalation strategy scored higher than the endurance strategy in nearly every game, but its range of scores was much wider.

Current Tetris Modeling

In the current studies, we manipulate two factors of the Tetris task environment: (a) game length and (b) zoid sequence. Each study evaluates and compares 11^6 decision strategies on each combination of these factors. In each study, each decision strategy is used in a short game and in

Table 1 Feature weights and feature values for each potential move

Feature	Weight	Move A value	Move B value
Landing height	− 1	5	2
Eroded cells	1	0	10
Row transitions	− 1	19	15
Column transitions	− 1	16	16
Pits	− 4	4	5
Cumulative wells	1	2	2
Total move score	0	− 54	− 41

[2] The number of alternative possible placements for a zoid varies with the zoid type and with the current state of the board; however, for an empty board, there are 9 positions where the square zoid can be placed, 17 positions where the Z-, S-, and I-beam can be placed, and 34 positions where the L-, J-, and T- can be placed. Hence, the "average" zoid in an unconstrained board can be placed in any one of 23 locations.

a long game. We refer to these MPs, respectively, as our hares and our tortoises. All MPs in study 1a use one random seed and all MPs in study 1b use another random seed but are otherwise identical. In study 2, the decision strategies used by each of our 11^6 MPs are tested by playing the same set of ten random seeds (all of which differ from the random seeds used in study 1), that is, each of the 11^6 MPs is run ten times, each time with one from a fixed set of ten random seeds. For each of these 11^6 models, we report the average score for each set of ten seeds.

Game Length Our game length manipulation of the task environment asks whether the best placement for a given zoid varies depending on the expected length of the game. Our tortoise MPs play as long a game as they can with the longest game > 100,000 zoids. In contrast, our hare MPs are limited to a game length of no more than 506 zoids.

Zoid Sequence Tetris has 7 differently shaped zoids and the sequencing of these zoids is determined by a random seed (and is with replacement). Do different methods for deciding where to place zoids on the Tetris board emerge when different seeds are used between models? Do different placement methods develop when multiple seeds (i.e., a more varied task environment) are used to train each model? If so, does this imply that models trained on a single seed are fragile and that ones trained on multiple seeds are more robust?

In both studies, we ask whether any of our manipulations affect the behavior of the best scoring model(s). Specifically, do the proportion of one-line, two-line, three-line, and four-line clears vary across our manipulations and do some patterns of line clears resemble human patterns more so than others?

MindModeling.org

To fully explore the performance differences between models on long and short games, we used "MindModeling." The MindModeling project uses volunteered computing time to run cognitive science simulations (Glendenning et al. 2016; Gluck and Harris 2008). For our purposes, it allowed us to define a desired search space, and the system performed a grid search, using a single volunteer computer to play one game, thousands of computers at a time, and returned each game's results. This method allowed a search that would have taken months or years on a single computer to be completed in a matter of days.

Common Procedures

Our goal for both studies was to find the best models on long games and on short games, and to compare their performance and their behavior. We defined a large search grid, and had MindModeling test every combination of feature weights within that space (see Fig. 2, Table 1, and the surrounding discussion).

In summary, for these studies, we vary 11 weights on the same set of 6 features to produce 11^6 models. The models in our long conditions are our tortoise models and each of these runs until it dies. The models in our short conditions are our hare models and each of those runs until it dies or until it plays 506 Tetris episodes, whichever comes first. In study 1, we run two sets of models by using two different random seeds. In study 2, we run one set of models but we run each of the 11^6 models ten times, each time using the same set of ten random seeds. For this study, we report the mean performance of the model whose mean performance across all ten seeds is the highest.

Study 1

For study 1, MPs performed two sets of identical grid searches: one set with one random seed (study 1a) and the second set with a second random seed (study 1b). For each substudy, each seed was run twice, once for the tortoise search and once for the hare.

Study 1a

As expected, the best performing model in the long game condition was not the same as the best performing model in the short game condition.

Study 1a Results The best tortoise model scored outrageously high (34,847,635,540 points), clearing a total of 125,829 lines. This model scored nearly ten billion points higher than the next highest scoring model (27,572,380,920 points, 107,934 lines). However, this score is an inflated score, that is, a score of this magnitude is only possible because, according to Tetris rules, the amount of points awarded with each line clear increases with the game level at which that line is cleared. Hence, the inflated score is intended to reward the difficulty for humans of clearing lines at later, and faster, stages of the game.

Our MPs, however, are unconstrained by the speed of the game, and where the human game is considered unplayable past level 30, MPs play well beyond those limits. The best tortoise model discussed here reached level 12,582. To reduce the sense of score inflation, we will calculate an additional score, called the base score, where each one-line, two-line, three-line, and four-line clear is awarded its level 0 point value, regardless of when the line clear was completed. Even without the inflated level bonuses, the best tortoise model achieved an impressive 5,527,820 points (For the rest of this paper, we will report only the "base score" for models playing long games and both the base score and the inflated score for models playing short games. Any human scores reported are inflated scores).

In the short game (i.e., "human length") condition, the highest performing hare model scored 240,900 points (inflated score), clearing 199 lines using 506 pieces (The maximum number of lines possible to clear with 506 pieces appears to be 202). The score is close to the highest human score we have collected in our dataset of 300 Rensselaer students playing an hour each of Tetris under laboratory conditions (see Lindstedt and Gray 2018, for more details). Without the inflated score bonus, the best hare model's base score was 18,900 points. While eliminating the inflated score makes the performance of long models easier to grasp, for short games played on a more human time scale, the inflated score (with the level bonus) is useful for comparisons with human performance.

The models were then compared by their performance in the alternate game length condition. The highest scoring tortoise model did reasonably well on a short game (8720 points base score, 92,700 points inflated score); this score is close to the scores obtained by intermediate-level HPs in our laboratory. When allowed to run until it dies, the best hare model scored above average on a long game (68,440 points base score) but was nowhere close to the best performing long model. A summary of these results is shown in Table 2.

Although the stamina of tortoise models is impressive, our key focus in these analyses is behavior, not length. Specifically, our focus is on differences in the proportion of one-line, two-line, three-line, and four-line clears. Compared to hare models (see Table 3), tortoise models show a higher percentage of one-line clears, and much lower proportions of three-line and four-line clears. This pattern of tortoise line clears is consistent with the endurance strategy. In contrast, hare models adopt the riskier but higher scoring escalation strategy, indicated by a more U-shaped function which, compared to the tortoise model, shows a higher percentages of three-line clears and a much higher percentage of four-line clears.

Study 1a Discussion In this work, we are interested primarily in MP behavior and, only secondarily, in MP score. We find it interesting that the disparate techniques of model selection (used here) versus model training (used in Sibert et al. 2017) converged on the same behavior, namely, that the models that score the highest in short games of Tetris, our hare models, are also the models that exhibit the more human-like behavior, that is, they are the ones which adopt the riskier but higher scoring escalation strategy of three-line and four-line clears. However, as this conclusion is based on the random selection of one seed, study 1b attempts to determine if the conclusions drawn here are limited to that particular random sequence of Tetris zoids.

Table 2 Best study 1a model points, scores, and lines cleared

Model	Length	Points	Scoring	Lines
Best tortoise	Long	5,527,820	Base	125,829
Best hare	Long	68,440	Base	2243
Best tortoise	Short	92,700 8720	Inflated base	202
Best hare	Short	240,900 18,900	Inflated base	199

Table 3 Study 1a: model behavior—differences in the percentage of types of line clears

Line clear type	Best tortoise	Best hare
1-line	73.61	47.74
2-line	24.26	25.13
3-line	2.01	9.05
4-line	0.11	18.09

Study 1b

Except for the choice of a different random seed, the methods for study 1b were identical to those for 1a. Of prime interest for study 1b was whether the best model would have the same parameter set as the best study 1a model.

Study 1b Results For both hares and tortoises, the best performing model on the new seed was not the same as the best 1a model. The best 1b tortoise model scored 3,137,000 (base) points, clearing 59,678 lines (see Table 4). This model cleared only half as many lines as the best 1a model and scored less than 1/3 as many points. The best 1b hare model scored 206,340 inflated points or 17,820 base points by clearing 202 lines. Hence, the best 1b hare model showed a performance level close to the best hare 1a model. Though these best 1b models had a different sequence of zoids than the best 1a models, comparing Table 5 to Table 3 shows that their behavior, as defined by types of line clears, is similar. Compared to the hare model, the tortoise model clears more one-line clears, very few three-line clears, and has almost no four-line clears.

Although the best 1a models were not the absolute best when performing on a different game seed, we might expect

Table 4 Best study 1b model: points, scores, and lines cleared The tables were renumbered as well as the citations. Please check.

Model	Length	Points	Scoring	Lines
Best tortoise	Long	9,491,319,440 3,137,000	Inflated base	59,678
Best hare	Short	206,340 17,820	Inflated base	202

Table 5 Study 1b: model behavior—differences in the percentage of types of line clears

Line clear type	Best tortoise	Best hare
1-line	52.57	38.61
2-line	3.75	32.67
3-line	8.21	14.85
4-line	1.71	13.86

extremely successful models to display at least high levels of performance. To that end, in addition to finding the absolute best models in 1b, we also looked at the performance of the best 1a models in the second game condition (see Table 6). The best tortoise 1a model scored 128,720 base points, clearing only 2377 lines on the alternate seed. This score is far below its performance on 1a (5,527,820 base points and 125,829 lines), as well as being far below the best performing model on this particular seed.

The best hare 1a model scored 122,320 escalating points (12,320 points Base Score) by clearing 200 lines on the seed used for 1b. This score is about half the score the model reached in 1a (240,900 points, 199 lines), and is well below the best scoring hare model for this particular seed.

Study 1b Discussion Study 1b replicates our behavioral finding from study 1a, namely, that the best hare model exhibits more human-like behavior than the best tortoise model in that it "adopts" more of the riskier but higher scoring escalation strategy that gains more three-line and four-line clears. However, for both the hare and tortoise conditions, we were surprised by how much the second seed affected performance.

Study 2

Running our best study 1a model on the game seed for study 1b showed that the 1a model was merely the best model for its original game seed, not the best Tetris model, period. Hence, in our initial quest for a "best" model, we had instead found a model highly optimized to a single task environment. Changing that task environment by sampling a different random seed resulted in different "best model." This reversal in fortunes for 1a models running with 1b seeds held true for models in both length conditions, though much more dramatically for our tortoise models. For study 2, we explored whether a better definition of "best model" would be one based on a model's mean performance across multiple game environments (i.e., across multiple game seeds).

Table 6 Scores and line clears when the best study 1a model plays the study 1b seed

Model	Length	Points	Scoring	Lines
Best tortoise	Long	128,720	Base	2377
Best hare	Short	122,320 12,320	Inflated base	200

Finally, although the behavioral findings of differences in distributions of line clears between our tortoise and hare models were similar in studies 1a and 1b, rerunning the models across multiple seeds provided an opportunity to confirm these findings.

Study 2 Methods

We again used the MindModeling platform and the same feature weight search space, but for each combination of possible feature weights, the model played one game each of a set of ten different game seeds. After all ten games were completed for all 1,771,561 models, we examined the average scores of the models for both escalated and base scores. All discussion of study 2 model scores will be based on the mean score across all ten seeds (The model defines the decision strategy whereas the seeds and game length define the task environment.).

Study 2 Results

For short games, across the ten seeds, the best study 2 hare model averaged 155,516 points in inflated score (see Table 7). This score is about equivalent to an advanced, but not expert, HP. However, despite a hare game score that is fairly consistent with human numbers, across the sample of ten games (each using a different random seed) model scores were much more variable than humans generally are. The difference between the highest scoring game and the lowest scoring game was 214,540 points (inflated score). The best tortoise model was even less consistent. Though the average score was very high, 11,381,153 points (base), across the ten seeds, the SD was even higher, 12,502,997 (base).

In terms of behavior (see Table 8), the distributions of line clears between the best tortoise and best hare models mimicked that found in studies 1a and 1b, namely, more three-line and four-line clears for hares than for tortoises. For all ten model runs (each with a different random seed), the mean number of one-line, two-line, three-line, and four-line clears are shown in Table 8.

Study 2 Discussion

In contrast to study 1, in study 2, the best model was the one that returned the highest performance averaged across ten games, each with a different random seed. Despite the differences shown in number of lines cleared and scores when performance is optimized across ten seeds, study 2 replicates the important constancy found across studies 1a

Table 7 Best study 2 model scores and lines cleared for study 2. Each entry represents the mean score for the 10 seeds used in the best study 2 model. (*CV*, coefficient of variance)

Model	Length	Mean points	SD points	CV	Scoring	Mean lines	SD lines
Best tortoise	Long	239,723,642,513	353,343,343,239	1.47	Inflated	233,511	256,201
		11,381,153	12,502,997	1.10	base		
Best hare	Short	155,516	64,770	0.42	Inflated	184	45
		14,686	5168	0.35	base		

and 1b, namely, comparing Table 8 to Table 3 and Table 5; we see that the behavior of the best long model compared to the best short model follows the pattern of more one-line and two-line clears and fewer three-line and four-line clears. As in study 1, for long games, the endurance strategy produces superior performance; whereas, for short games, the escalation strategy triumphs.

Human Players Change Tempo to Balance Speed with Performance

In the above studies, we conducted model searches independent of human data and chose the best models based on numerical score. Although the shape of the distribution of one-line, two-line, three-line, and four-lines clears for hare models was roughly comparable to human gameplay, in this section, we try to gain a better understanding of how our MPs resemble HPs across varying levels of human skill.

Confronted with the evidence from our MPs, how well do our HPs do? Do they resemble our tortoises or our hares, that is, do some HPs always use the escalation strategy and do others always use the endurance strategy? What, if anything, do we know about human behavior that can answer these questions?

Our prior work (Sibert et al. 2017) used CERL models to compare model performance with performance-based measures of human expertise across 67 human players. Using a very conservative matching criteria, we found that the models were able to exactly predict human moves between 35 and 65% of the time (the predictive power was correlated with the expertise of the player). The hare-like models were considered to better fit the human data, but the overall predictive power of both model types was very similar, and we did not find that certain human players were predicted better or worse by particular models. This suggests, as we have predicted above, that human players do not strictly adhere to either the endurance or the escalation strategy, but instead use a mixture of the two.

Table 8 Study 2: model behavior – for the best tortoise and best hare model, the table shows the mean percentage, standard deviation (SD), and coefficient of variance (CV) for each of the four types of line clears averaged across all 10 seeds

Line clear type	Best tortoise	SD	CV	Best hare	SD	CV
1-line	61.36	0.46	0.01	47.71	8.15	0.17
2-line	31.97	0.34	0.01	27.73	4.48	0.16
3-line	6.00	0.31	0.05	14.25	5.35	0.38
4-line	8.00	0.41	0.05	10.40	7.36	0.71

In this section, we leave aside the models used by Sibert et al. (2017) and re-examine these 67 human players and their games. For each game, we look at the number of Tetrises the HP made at each level they played. Rather than examining a single player's overall tendency to use a particular strategy, we instead examined the number of four-line clears (i.e., the number of Tetrises) they made at each level. As clearing ten lines advances the level by one, it is possible to obtain a maximum of three four-line clears in one level by using filled lines left over from earlier levels. However, in our sample of players, the maximum number of Tetris per level any player cleared was two (i.e., two four-line clears). More importantly, this scrutiny of individual games reveals a tendency of human players to make regular and significant changes in strategy throughout the course of most games. These results are displayed in Fig. 3.

Figure 3 conveys a qualitative sense of changes in tempo with level of Tetris. There are five subplots, one for each rank of player skill. Rank 1 represents the most expert player and rank 5 represents the most novice players. Each line represents one game as played by one HP. For ranks 2, 3, and 4, there are often many HPs overlapping with each other. That information is partially conveyed by the darkness of the circle—circles which are light gray represent one or a few HPs and the darker the circle the more games and the more HPs it represents.

Although we realize that for many purposes, the representation used in Fig. 3 would be too informal, we used it here as it provides a striking illustration of the switch, by individual HPs as well as the group of 67 as a whole, between the escalation and endurance strategies as the tempo of Tetris increases. At the same time, whereas all HPs use the endurance strategy as the game becomes too fast, Table 9 reveals that our most novice players (rank 5

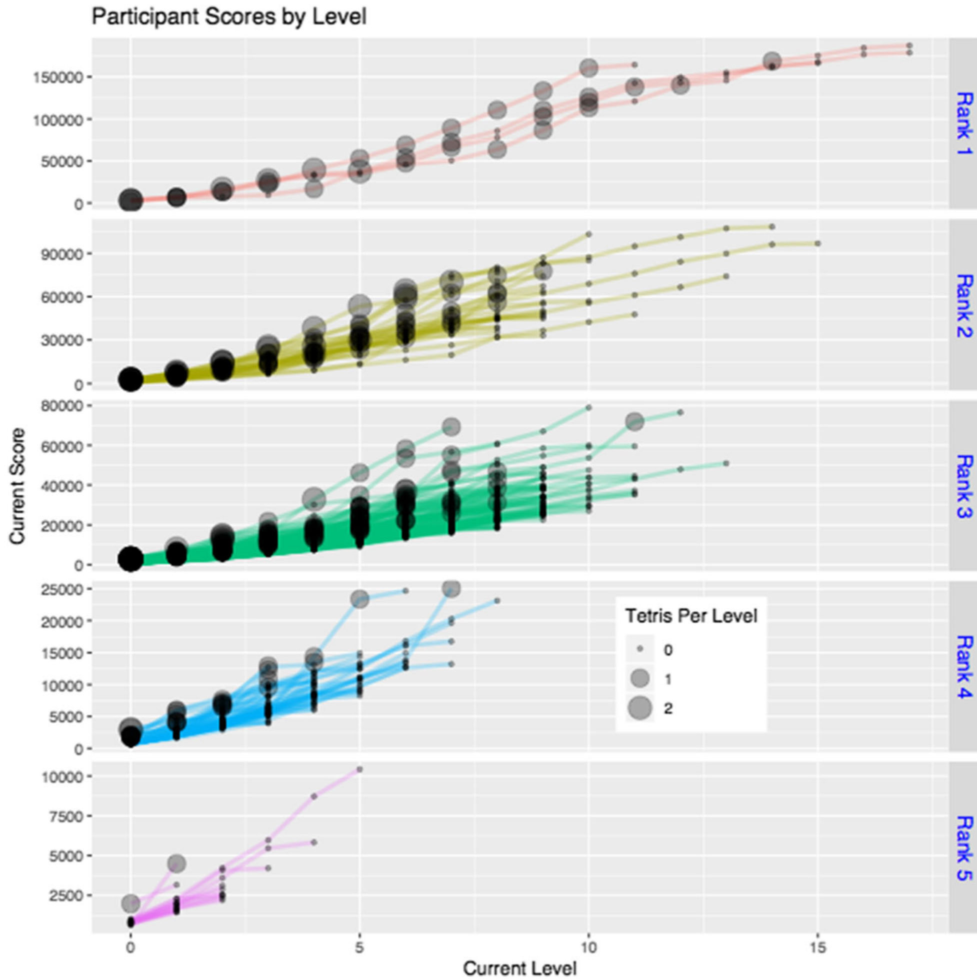

Fig. 3 Number of 4-line clears (i.e., Tetrises) per level for each of 67 HPs across each level and game that each HP played. The larger the dot, the more Tetris per that game at that level. The darker the dot, the more HPs represented at that level. Note that the range of levels played is plotted as the *x*-axis whereas the score reached in each level of play is shown on the *y*-axis. Also, note that the range of the *y*-axes vary by rank level with the shortest range being for rank 5 (novices) and the longest range for rank 1 (experts). See text for more details

and rank 4) use the most extreme version of the endurance strategy, the one-line clear, for the majority of their play. Indeed, our rank 2 players clear almost as many lines with one-line clears as rank 1 does with four-line clears.

At each of the five ranks of human expertise, as the pieces fall faster with increasing levels, the more likely individual HPs are to switch from an escalation to an endurance strategy. In this regard, although our MPs can teach us a lot about the utility of these tortoise and hare approaches, they do not capture the flexibility of HPs who have an ability to modulate their tactics based on game tempo. Presumably, skill at modulating tactics does not reflect skill at sequential decision-making but does represent planning which, as mentioned in our introduction, is another skill that Tetris taps into.

Table 9 Human behavior: percentage of total line clears by skill level

Line clear type	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5
1-line	26.63	33.31	46.87	51.82	62.37
2-line	19.93	22.28	31.63	30.55	24.75
3-line	4.41	14.05	10.81	10.20	9.85
4-line	49.02	30.36	10.69	7.43	3.03

Summary and Conclusions

Slow and steady wins the race, but the best performers may be those who know when to change their strategies. The differences in scores between our best tortoise and best hare model players (MPs) may be the most dramatic outcome of this work, but it is far from the most important. We see our contribution in terms of six bullets which we list in order of their increasing relevance to human behavior, namely,

1. the ability of small changes in the test environment (i.e., different random seeds for generating sequences of 7 numbers) to produce large changes in performance,
2. the failure to find one model (i.e., one set of six feature weights) that could override the importance of random number generation on zoid sequence (see Appendix A1),
3. the emergence of two different strategies for sequential decision-making, endurance for the tortoise models and

escalation for the hares, and the consistency of those strategies across multiple environments,
4. the identification of endurance and escalation strategies in human behavior,
5. the finding that the endurance strategy is favored by less experienced HPs and the escalation strategy favored by our most experienced HPs (Table 9), and
6. the additional finding that even our very best human players switch to an endurance strategy when they reach the limits of their expertise (Fig. 3).

Though the results of these studies revealed many things we did not expect, the most important finding confirms that small changes to the task environment can have a significant impact on behavior (Sibert et al. 2017). Indeed, these studies began as a validation of the methods of that prior study where, due to the technical constraints of the modeling method, we were only able to analyze models developed over short games.

Although we found that we were able to induce significant behavioral differences by varying the objective function of the models, we worried that by training on short games, the resultant models did not develop to their full potential.

In the current study, models allowed to play games of unlimited length obtained very high scores, but they did so by adopting an endurance strategy. Models evaluated on a shorter game length adopted an escalation strategy. However, unlike our human players, no model developed that was able to switch between these two strategies. The rigidity of these models implies that the ability to change strategy as conditions change is a significant factor in human Tetris play.

Something that we did not expect was the extreme fragility of the highest performing models. In our previous studies, only looking at short games, we found the tortoise models to be more consistent in their scores, and hare models more variable. In this report, when the game length was unlimited, we found that it was the tortoise models that were the most inconsistent. Part of this was likely due to the method of model search—in these studies, we selected the highest performing model, rather than use a search method that develops the best model. This meant that the selected model was only the best model on the particular game seed being tested. While we expected some variation from this method, we were not expecting the magnitude of variation that we observed. In study 2, we tried to address this issue by evaluating models based on their average performance over ten seeds, rather than a single seed, but rather than finding one model with consistently high performance across all seeds, we found models with extremely high scores on one of the ten seeds and average or even poor performance on the rest of the set. Evaluating a set of seeds may produce a better result than a single seed, but additional methods may be needed to produce the kind of high level consistency observed in human players. This conclusion is somewhat tempered by our finding that the spread across the ten models is less for the best hare model than for the best tortoise with the coefficient of variance (CV) for the former being 0.35 (base score) and the latter, 1.10 (base score) (see Table 7).

The tortoise models are the most prone to the kind of overfitting that we observed here in part because of the "strategy" employed.[3] One of the features of Tetris that makes it ideal for this kind of complex task analysis is that every game of Tetris eventually ends, even for a model. For humans, this loss is most often due to speed, but can also occur when the player makes a series of bad or less than optimal choices that raise the pile too high to execute the planned moves. This kind of board state often results during HP pursuit of the escalation strategy. The endurance strategy, by contrast, never builds up the pile in anticipation of a higher-order line clear, choosing to clear a line whenever possible. When the game length is unlimited, this strategy gets very close to being able to play infinitely.

The only reason that this is not possible is because there exist sequences of zoids that will fill the board without allowing any lines to be cleared (Brzustowski 1992; Burgiel 1997). For an effective tortoise model, the limiting factor has little to do with the model's choices, and more to do with when the seed will produce an unplayable string of zoids.

While the tortoise model's use of the endurance strategy is focused on long-term gameplay, such that it avoids risky and high scoring moves and can only score high over very long periods of time, the hare's reliance on the escalation strategy may lead to early losses by building toward moves that rely on waiting for particular pieces that may never arrive. For HPs, Fig. 3 suggests that strategy switches usually occur just once per game—when the escalation strategy is no longer viable. Table 9 suggests that the HPs who begin games with the escalation strategy and switch, as drop speed increases, to the endurance strategy are players of higher skill. These players implement the escalation strategy early and maintain it until the increasing drop rate approaches the limits of their ability to decide on good zoid placements while allowing enough time to move the zoid to one of those placements. Hence, our comparison of MPs with HPs leads us to believe that the speed of higher level players is not due simply to motor movement time but to decision-making time. Indeed, this more nuanced view of speed also emerges from a large-scale study of human performance in Tetris (Lindstedt and Gray 2018, under review) and is consistent with work on Hick's Law (Hick 1952) especially work concerned with individual differences (Jensen 1987).

[3] Note that the authors disagree among themselves as to whether this use of the term "strategy" is strictly appropriate; however, at least for now, we cannot find a better one.

Challenges for the Future

We plan to extend our modeling work in several directions, each aimed at exploring individual aspects of the Tetris task. Our current models provide insights into some of the individual decisions made throughout a game, but they are limited to considering a single zoid at a time. All human players have access to at least one upcoming zoid (i.e., the one in the "Next" or Preview Box, see Fig. 1a, upper right of the game board), and evidence suggests that higher level players take this into account when making decisions about zoid placements, that is, these players seem to optimize a continuously updated two-move sequence.

The current models are also limited in their zoid evaluation, in that they do not consider advanced moves that involve sliding one zoid into gaps or row transitions (see Fig. 2). Such moves often involve last instant rotations and slides to position the zoid inside a well and then move it under a left- or right-side overhang. The ability to recognize and make such movements is a characteristic of strong human players and are used by all players who make the playoffs at the annual Classic Tetris World Championships (see https://thectwc.com/). We hope that by imbuing our models with the capacity to choose such placements, we will get a more accurate estimation of the models' predictive powers.

We also intend to try isolating other skills which we believe contribute to high performance. We know how MPs evaluate piece placements using certain board features, and we know HPs are taking advantage of the same or similar information to come to similar decisions, through perceptual learning. By comparing HP decisions to MP decisions, we hope to get a better idea of which features are important, and how they interact to identify a certain board position as "good" or "bad."

These analyses rely not on finding the places where MPs agree with HPs, but instead, finding the places where they disagree. Such disagreements may indicate an error, either in judgment or execution. But there is a third option that a mismatch occurs when the human player deliberately makes a suboptimal decision to serve a longer term goal. The models, by design, are single-piece optimizers, and long-term performance arises from making the best individual move at each decision point. Human players are not as single-minded, and human experts consider a number of factors in their decisions, including the probability of getting a key zoid based on how recently that zoid has occurred in the sequence. We should not, then, automatically consider every disagreement between model and human to be an error, and we hope to be able to distinguish between a mistake and a strategic deviation.

Finally, we believe (as mentioned earlier) that placement speed has a significant decision-making component, but we also believe that movement time, per se, limits human choice to those positions that the HP can reach in the time available. Hence, the final skill that we mention is the motor time to move a zoid to a location. For future models, we plan on exploring model performance not simply based on the "best placement possible" but on the "best placement that the current zoid can be moved to in the time available."

With careful analysis of human data pointing the way toward the skills that are required to succeed in Tetris, we hope to use these modeling tools to further explore the individual contribution of each of those skills, and the effect they have on performance. These studies suggest that Tetris is a worthy and challenging domain in which to study human performance in its own right, as well as providing methods that could be used to analyze human performance in other complex tasks.

Support Information Wayne Gray received grant (N00014-17-1-2943) from the Office of Naval Research, Dr. Ray Perez, Project Officer.

Author Contributions Sibert and Gray conceived and designed the study and wrote the manuscript. Sibert did the modeling and data analyses. Gray and Sibert discussed the results and elaborated the theoretical implications of the results.

References

Anderson, J. R. (1991a). Is human cognition adaptive? *Behavioral and Brain Sciences, 14*(3), 471–517.

Anderson, J. R. (1991b). The place of cognitive architectures in a rational analysis. In K. VanLehn (Ed.). Architectures for intelligence (Chap. 2, Vol. 22, pp. 1–24). Hillsdale, NJ: Lawrence Erlbaum Associates.

Baccherini, D., & Merlini, D. (2008). Combinatorial analysis of Tetris-like games. *Discrete Mathematics, 308*(18), 4165–4176. https://doi.org/10.1016/j.disc.2007.08.009.

Brzustowski, J. (1992). Can you win at tetris? (Master's thesis, University of British Columbia, Vancouver, BC, Canada).

Burgiel, H. (1997). How to lose at Tetris. The Mathematical Gazette, 81 (491), 194–200. Retrieved from http://www.jstor.org/stable/pdfplus/3619195.pdf

Fahey, C. P. (2015). Tetris AI. [Online; accessed 2015-Jan-30]. Retrieved from http://www.colinfahey.com/tetris/

Gabillon, V., Ghavamzadeh, M., & Scherrer, B. (2013). Approximate dynamic programming finally performs well in the game of Tetris. In Advances in neural information processing systems (Vol. 26, pp. 1754–1762). Retrieved from http://media.nips.cc/nipsbooks/nipspapers/paper_files/nips26/881.pdf

Glendenning, K., Wischgoll, T., Harris, J., Vickery, R., & Blaha, L. (2016). Parameter space visualization for large-scale datasets using parallel coordinate plots. *Journal of Imaging Science and Technology, 60*(1), 104061–104068. https://doi.org/10.2352/ImagingSci.Technol.2016.60.1.010406.

Gluck, K. A. & Harris, J. (2008). Mindmodeling@home. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.). Proceedings of the 30th annual conference of the cognitive science society (p. 1422). Cognitive Science Society.

Gray, W. D. (2017). Games-XP: action games as cognitive science paradigm. *Topics in Cognitive Science, 9*(2), 289–307. https://doi.org/10.1111/tops.12260.

Hick, W. E. (1952). On the rate of gain of information. *The Quarterly Journal of Experimental Psychology, 4*, 11–26. https://doi.org/10.1080/17470215208416600.

Holmes, E. A., James, E. L., Coode-Bate, T., & Deeprose, C. (2009). Can playing the computer game "Tetris" reduce the build-up of flashbacks for trauma? A proposal from cognitive science. *PLoS One, 4*(1), e4153. https://doi.org/10.1371/journal.pone.0004153.

Jensen, A. R. (1987). Individual differences in the Hick paradigm. In P. A. Vernon (Ed.), *Speed of information processing and intelligence* (pp. 101–175). Norwood, NJ: Ablex.

Lindstedt, J. K., & Gray, W. D. (2015). Meta-T: Tetris as an experimental paradigm for cognitive skills research. *Behavior Research Methods., 47*, 945–965. https://doi.org/10.3758/s13428-014-0547-y.

Lindstedt, J. K., & Gray, W. D. (2018). Distinguishing experts from novices by the mind's hand and mind's eye. In *Manuscript submitted for publication*.

Martin-Gutierrez, J., Luis Saorin, J., Martin-Dorta, N., & Contero, M. (2009). Do video games improve spatial abilities of engineering students? *International Journal of Engineering Education, 25*(6, SI), 1194–1204.

Mayer, R. E. (2014). *Computer games for learning: an evidence-based approach*. Cambridge, MA: MIT Press.

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*, 1–21. https://doi.org/10.1111/tops.12225.

Simon, H. A. (1991). Cognitive architectures and rational analysis: comment. In K. VanLehn (Ed.), *Architectures for intelligence* (Vol. 22, pp. 25–39). Hillsdale, NJ: Lawrence Erlbaum Associates.

Simon, H. A. (1992). What is an "explanation" of behavior? *Psychological Science, 3*(3), 150–161.

Şimşek, Ö., Algorta, S., & Kothiyal, A. (2016). Why most decisions are easy in tetris—and perhaps in other sequential decision problems, as well. In *Proceedings of the 33rd international conference on machine learning* (pp. 1757–1765). JMLR.

Stafford, T., & Dewar, M. (2014). Tracing the trajectory of skill learning with a very large sample of online game players. *Psychological Science, 25*(2), 511–518. https://doi.org/10.1177/0956797613511466.

Stuart, K. (2010). Tetris and Snake - the biggest games in the world. The Guardian. Retrieved October 20, 2012, from http://gu.com/p/2e2kk/em

Szita, I., & Lorincz, A. (2006). Learning Tetris using the noisy cross-entropy method. *Neural Computation, 18*(12), 2936–2941.

Terlecki, M. S., Newcombe, N. S., & Little, M. (2008). Durable and generalized effects of spatial experience on mental rotation: gender differences in growth patterns. *Applied Cognitive Psychology, 22*(7), 996–1013.

Thiery, C., & Scherrer, B. (2009a). Building controllers for Tetris. *ICGA Journal, 32*, 3–11 Retrieved from http://hal.archives-ouvertes.fr/inria-00418954/.

Thiery, C., & Scherrer, B. (2009b). Improvements on learning Tetris with cross-entropy. *ICGA Journal, 32*(1), 23–33.