Gaze Transitions as Clues to Expert-Novice Strategy Differences in a Dynamic Video Game

Wayne D. Gray, Ryan M. Hope, Matthew-Donald Sangster, & John K. Lindstedt

Cognitive Science Department – Rensselaer Polytechnic Institute

Abstract—Many important and vital tasks entail dynamic decision making over prolonged periods of time in situations where "even hesitating requires a decision to hesitate" [1]. Although many such tasks are difficult to study without interfering or modifying, there is a category of these tasks, namely video games, which can be brought into the lab without harm to persons or to realism. In this report, we find significant differences in gaze transition probabilities between expert and novice Tetris™ players and present a novel visualization for two sets of transitions that suggest strategy differences in how the game is played.

Index Terms—Expertise, Tetris, Regions of Interest, ROI, AOI, fixations, saccades

1 Introduction

We study expertise in tasks requiring real-time human interaction with dynamic task environments – tasks entailing the integration of cognition, perception, and action where "even hesitating requires a decision to hesitate" [1]. Unfortunately, many such tasks, for example, air traffic control, laparoscopic surgery, driving a car, are difficult to study without interfering or modifying the task, or even dangerous to study under real conditions. Fortunately, there is a category of these tasks; namely, video games, which can be brought into the lab without harm to persons or to realism.

Here we focus on the video game Tetris™. Although easy to learn, it is diffcult to master. Although a task which people play in the real-world, it is easy to bring into the laboratory [e.g., see 3].

Our past research on Tetris has examined complementary actions [4] [aka epistemic actions, 5]. Our current work uses AI agents to help us understand how different solutions to placing one Tetris piece (i.e., zoid) on the same board can vary in goodness [6], [7], whether player expertise can be predicted based on where players place their first 2, first 10, first 100, or all zoids of a game [8], [9], and whether experts and novices differ in their patterns of gaze transitions from one region of interest (dynamic or static ROIs) to another.

Here we present our first, albeit short, paper on the latter topic; namely, whether differences in expert versus novice strategies can be revealed by *eye transition data*. The work presented here is based on:

- 34,589 fixations from 4 games of Tetris for 10 Novices and 10 Experts (where the fixation count excludes successive fixations to the same ROI).
- A *base model* which assumes that the probability of a transition between two ROIs simply reflect the probabilities that an ROI is fixated.
- A log-linear analyses which tests and rejects that base model.

As the base model is rejected, in this short report, we provide two examples in which Novices and Experts differ in their patterns of transitions.

A Challenge to Cognitive Science Although Tetris is a simple game played by millions, when we look at Tetris what we see presents a challenge to Cognitive Science. First, the *cognitive control* required to excel at Tetris requires switching among substeps to: (a) monitor, (b) guide, and (c) place the current zoid, while (d) planning the placement of the next zoid. Second, the actions that we see for the current zoid suggest the: (a) dynamic adjustment of placement plans, (b) a continual evaluation of the pile, (c) and some sense that our experts have that we do not yet have of strategic goals and tactics for Tetris. Hence, we turn to detailed analyses of fixation transitions to gain some insight as to how these challenges might be addressed by cognitive theory.

Manuscript received 26 Jul. 2015; accepted 1 Aug. 2015; date of publication xx Aug. 2015; date of current version 25 Oct. 2015. For information on obtaining reprints of this article, please send e-mail to: grayw@rpi.edu.

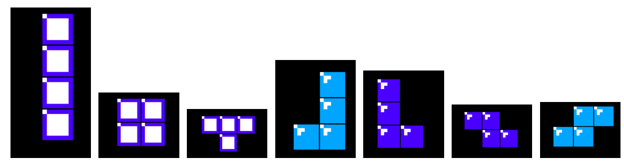


Fig. 1 – The Tetris pieces are called *zoids*. Each zoid is composed of four square blocks. There are seven (7) different zoids which are commonly called the: I, Square, T, J, L, Z, and S.

A Very Short History of Research on Gaze Transitions The transition of gaze fixations between various regions of interest (ROIs) has been researched at least since Fitts, Jones, and Milton [10] famous studies of fighter pilots during the Korean War era. Our approach to the analysis of transition frequencies has a less ancient but still old foundation in the pioneering work of Ellis and Stark [11] who applied χ^2 analyses to determine if the number of transitions between various ROIs were simply proportional to the number of fixations to each ROIs. Our use of log-linear analyses follows the more modern recommendations of Holmqvist, Nyström, Andersson, *et al.* [12]. In contrast to the χ^2, for each cell of our "from–to" table, the log-linear analysis provides us with the adjusted residuals expressed as z-scores. In standard statistical analyses, a z-score of ±1.96 is considered to show a two-sided significance at the $p < 0.05$ level. For the results shown here, we highlight z-scores with a two-sided significance of $p < 0.001$; namely, those with z-scores greater than ±3.3.

2 Methodology

All eye data reported here were collected on an *SMI 500 Red™* at 500 hz and were processed using the GazeTool package [2]. All players were individually run and all had prior experience with Tetris. All used *MetaT* [3], which logged and timestamped all system and user events, to play Tetris for an hour as the first session of a longer study.

Table 1 – A two-event transition matrix showing similarities and differences between Expert (top half) and Novices (bottom half) in shifts of attention from ROI "A" to ROI "B." The numbers represent significance values in terms of their *z-score*. For example a $z > 1.96$ represents a two-sided probability of $p < 0.05$ of finding that value by chance. The table highlights z-scores which are greater (green) or less (brown) than ±3.3. The two-sided probability of finding such differences by chance would be $p < .001$

From / To	PBox	BTW	Zoid	CurDest	NextDest	Pile
EXPERT	**EXPERT**	**EXPERT**	**EXPERT**	**EXPERT**	**EXPERT**	**EXPERT**
PBox	0.0	**5.7**	**3.4**	−1.6	1.8	−2.6
BTW	**3.6**	0.0	**5.4**	**5.3**	1.1	**-4.2**
Zoid	**5.2**	**16.9**	0.0	−2.7	−2.4	**-13.4**
CurDest	0.7	**-11.6**	**8.5**	0.0	**-3.3**	**3.3**
NextDest	−0.8	**-7.0**	−1.9	−0.3	0.0	**5.6**
Pile	−0.3	−1.8	**-16.2**	2.6	**4.2**	0.0
NOVICE	**NOVICE**	**NOVICE**	**NOVICE**	**NOVICE**	**NOVICE**	**NOVICE**
PBox	0.0	−0.2	−0.4	−2.9	−3.2	−2.0
BTW	−1.1	0.0	0.6	**-7.1**	**-4.4**	0.5
Zoid	−0.7	**5.7**	0.0	−0.4	2.7	**-10.0**
CurDest	**-3.6**	**-11.9**	**5.6**	0.0	**-5.7**	**8.7**
NextDest	**-3.4**	−2.2	**3.9**	**-6.1**	0.0	**8.3**
Pile	−1.9	**8.2**	**-12.7**	**8.1**	**8.8**	0.0

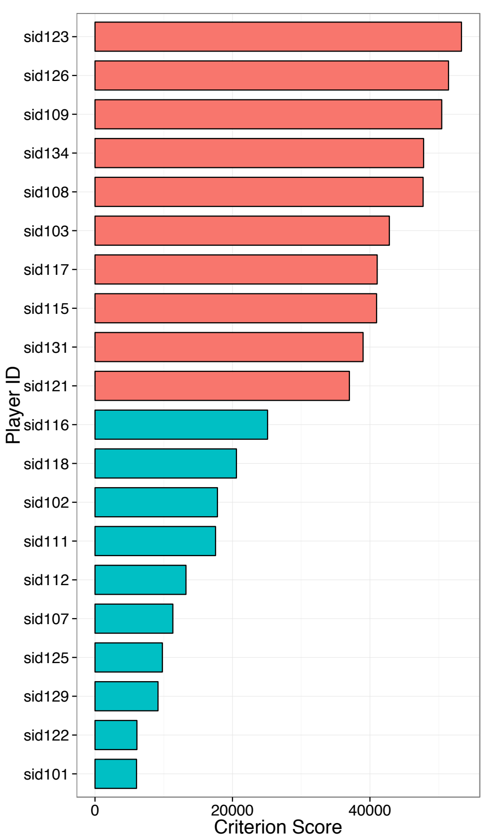


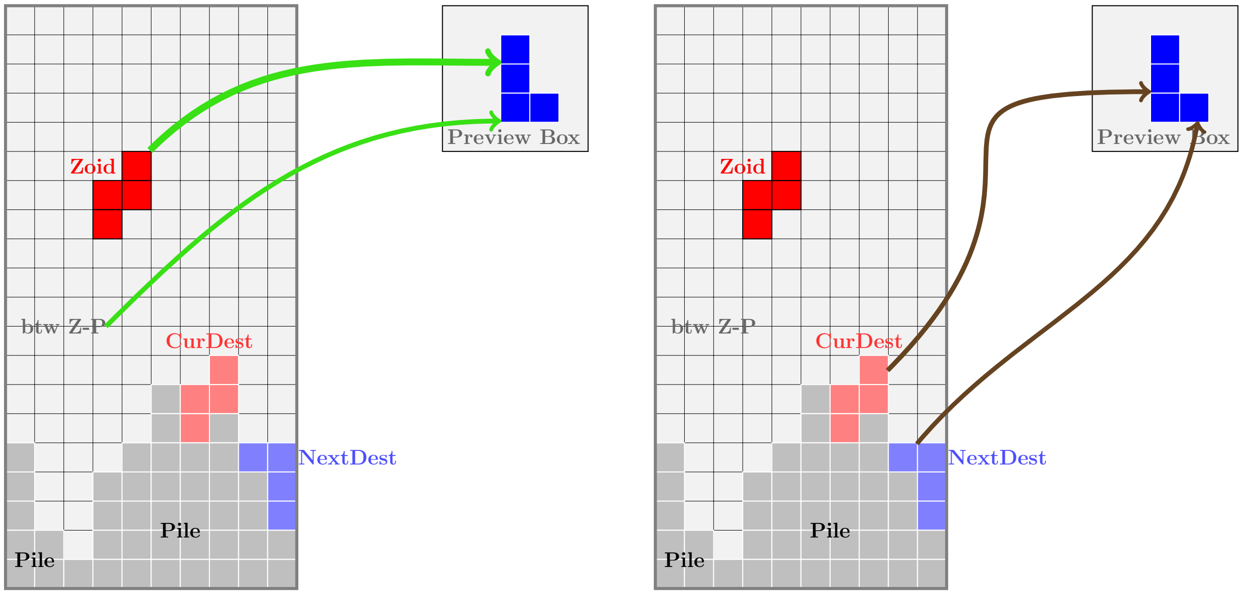
Fig. 2 – Criterial Score for each of the 10 Expert and 10 Novice Players. The criterion score is the mean of the highest 4 games played during the first hour of a Tetris study. Players in the Intermediate group were excluded so as to widen the contrast between Experts and Novices.

To rate player expertise, we averaged the score of the four highest games played in our lab in a one-hour session. This number is the *criteria score* for each player. We applied the ddendro algorithm [13] to the criteria score to classify players into 8 groups. The top three players were enough different from our Expert players as to be in a separate group but as three did not seem sufficient for a sample, we exclude them from the current analyses and focus here on our 10 expert and 10 novice players. The jump between the lowest expert and highest novice, shown in Figure 2, was occupied by our intermediate players.

2.1 ROIs

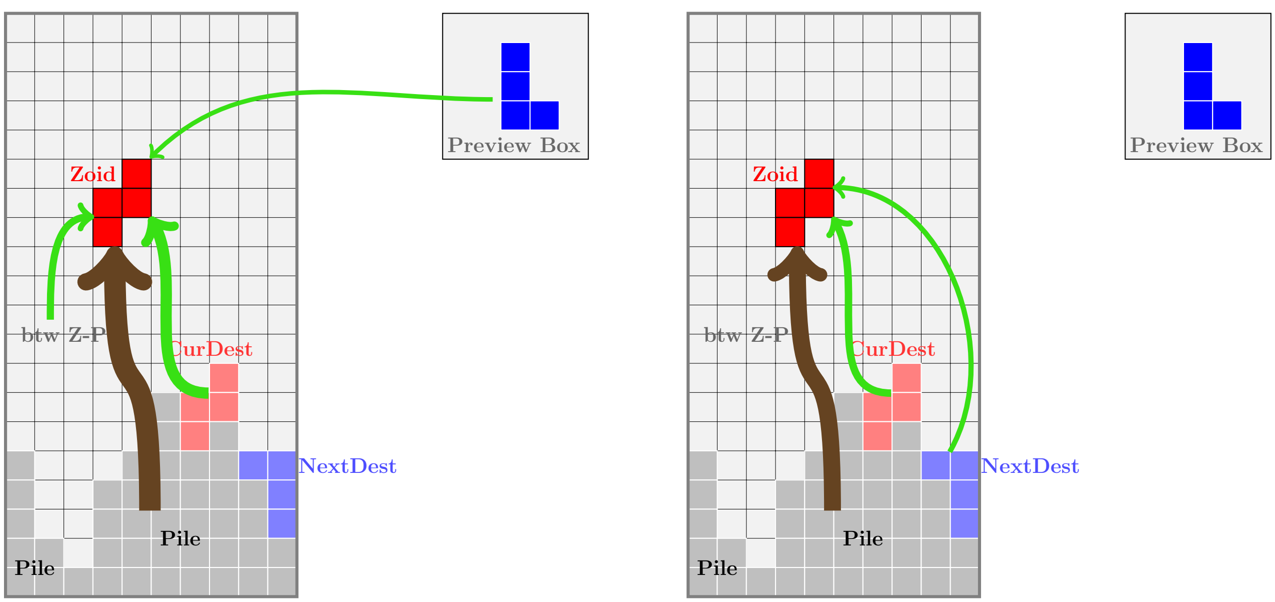
We used the dynamic ROI (dROI) feature of GazeTools [2] to assign fixations to one of six regions of interest (see the labeled ROIs in Figures 3 and 4): Preview Box (PBox), BTW (the area *between* the Zoid and the Pile), Zoid, CurDest, NextDest, and Pile). Fixations to other areas were few and were excluded from the current analysis.

Each Tetris *episode* entails the movement and placement of one zoid (one of 7 possible Tetris shapes, see Figure 1). During the episode only the zoid is dynamic. However, on each episode the zoid ends up in one spot – the *current destination* or "CurDest" – dependent on the twin factors of "opportunity" and "choice". This terminal spot, the CurDest, is also a dROI as it varies from episode to episode. Likewise, on each episode the zoid that will be placed during the *next episode* appears in the aptly named "Preview Box". The spot where this next piece is placed, on the next episode, is also a dROI which we refer to as the *next destination* or "NextDest".



(a) Expert **(b)** Novice

Fig. 3 – *To the Preview Box* – Differences in transition frequency of eye movements made by Expert (a) vs Novice (b) Tetris Players from other ROIs to the Preview Box. Green transitions are more likely than chance ($> +3.3sd$), brown transitions are less likely ($< -3.3sd$), with the width of the line being proportional to the $\pm$ of the sd.



(a) Expert **(b)** Novice

Fig. 4 – *To the zoid* – Differences in transition frequency of eye movements made by Expert (a) vs Novice (b) Tetris Players from other ROI to the zoid Box. Green transitions are more likely than chance ($> +3.3sd$), brown transitions are less likely ($< -3.3sd$), with the width of the line being proportional to the $\pm$ of the sd.

3 Results

Our approach to the analysis of transition frequencies was pioneered by Ellis and Stark [11] who performed χ^2 analyses to determine if the found number of transitions exceeded the number expected by the base model. Our use of log-linear analyses follows the recommendations of Holmqvist, Nyström, Andersson, *et al.* [12]. In contrast to the χ^2, for each cell of our "from–to" table, the log-linear analysis provides us with the adjusted residuals expressed as a z-score. In standard statistical analyses, a z-score of ± 1.96 is considered to show a two-sided significance at the $p < 0.05$ level. For the results shown here, we highlight z-scores with a two-sided significance of $p < 0.001$; namely, those with z-scores greater than ± 3.3.

Table 1 shows the *adjusted residuals* from a log-linear analysis of the *transition matrix* comparing Expert (top) and Novice (bottom) Tetris players. Left-to-right, each row shows the transition "from" the ROI on the left-most column "to" each ROI in column 2-7.
The model uses *structural zeros* to zero out what would otherwise be self-transitions (i.e., two successive fixations in the same ROI), as per [12]. The adjusted residuals are reported as z-scores which become higher or lower than zero to the extent to which the *from-to* transitions are higher or lower than would be expected simply by chance. To be clear, if Tetris players show no systematicity in their eye transitions, then we would expect the residual score shown for saccading from, say, the zoid to the CurDest, would be close to zero. Although Table 1 is complete, it is less than intuitive when it comes to "seeing" its implication for expert versus novice strategy differences. For that task, visualizations such as we use in Figures 3 and 4 are more helpful.

3.1 Transitions to the Preview Box.

Figure 3 zooms in on the green and brown numbers in Column 1 of Table 1 to highlight differences in expert versus novice transitions to the Preview Box. The green arrows for the experts show that they make transitions from the zoid to the Preview Box and from the area between the zoid and the Pile to the Preview Box more than would be expected by the null hypothesis that transitions are simply proportional to the number of fixations to a ROI. In contrast, the Novice plot on the right of Figure 3 suggests that Novices are less likely to transition from either the CurDest or the NextDest than would be expected by the null hypothesis.

Our current interpretation of these data is that Experts make strategic use of the time during which the current zoid is dropping to gain information about the next zoid so as to begin planning their next placement. In contrast, Novices seem to be doing something completely different. As suggested by the absence of green lines from the zoid and the BTW to the Preview Box, Novices do not use the drop time for planning but rather to monitor the dropping zoid to ensure it ends up in the targeted location. Likewise, the brown lies from the CurDest and NextDest locations to the Preview Box, suggest that Novices are less likely than expected to plan for the next zoid while the current zoid is falling, than are the Experts.

3.2 Transitions to the zoid.

Figure 4 zooms in on the green and brown numbers in Column 3 of Table 1 to highlight similarities as well as differences in expert versus novice transitions to the zoid. As can be seen in the Figure, both groups are more likely than expected to transition from the CurDest to the zoid and less likely than expected to transition from somewheres else in the Pile to the zoid. These patterns suggest that both experts and novices tend to visually guide or, at least, monitor the descent of the zoid to its current destination.

This contrast between more likely than expected (i.e., by our default hypotheses) transitions from the CurDest to zoid and less likely than expected transitions from the Pile to the zoid also supports our decisions to treat the CurDest as its own ROI and, therefore, distinct from either the Pile or the BTW regions.

These contrasts highlight the utility of our dROI measure as, without knowing the regions where the current zoid and the next zoid are placed, both of those areas would be classified as part of the BTW area. However, we see that when CurDest and NextDest are subtracted from the BTW area that Novices and Experts vary in their likelihood of transitioning from the NextDest to the zoid as well as from the BTW area to the zoid. For the present, however, we will refrain from over interpreting these data and await the analyses of a greatly enlarged dataset of Players.

4 Discussion

The results presented here are preliminary but suggestive. They are preliminary in that we are currently processing a much large set of data and hope to have four times as many experts and novices as at present. They are suggestive in that the differences between experts and novices seem reasonable and seem to support the position that the eye data are revealing strategic differences in how the two groups approach Tetris.

5 Summary & Conclusions

Our analysis of eye data is part of a larger effort to study the acquisition of extreme expertise in a dynamic task entailing real-time decision making. Other parts of this effort are attempting to classifying differences in moves made by novices and experts [8], [9], comparison of human players across their full range of expertise with Artificial Intelligence models that vary in their higher level goals [6], [14], and in differential use of *complementary actions* [4] across the spectrum of Tetris expertise.

In the near future, we hope to use eye data to test, understand, and interpret some of the results from these other parts of our effort. We also plan to focus on differences in the patterns of eye transitions made by Expert players as they play through the low levels of Tetris (when the zoids fall slowly) to reach the higher levels of the game that expose the limits of their expertise. Although the current short report has focused on eye transition data, we will also examine fixation durations and hope to have enough data to provide a meaningful test of differences in fixation duration in a ROI as a function of the preceding ROI. Such data might suggest that the purpose of a fixation on a given ROI varies as a function of the ROI of the preceding fixation.

Tasks entailing dynamic decision making over prolonged periods of time provide a challenge to the cognitive science community to build integrated models of cognitive systems. They also provide a challenge to trainers to provide training that takes domain experts beyond plateaus of stable but suboptimal performance [15], [16] to extreme expertise.

Acknowledgments

The work was supported, in part, by grant N000141310252 to Wayne D. Gray from the Office of Naval Research, Dr. Ray Perez, Project Officer.

6 References

[1] S. J. Lec, *Unfrisierte Gedanken (Unkempt thoughts)*, J. G. (translator), Ed. New York: St. Martin's Press, 1962.

[2] R. M. Hope. (2014). Gazetools: a collection of functions for processing and classifying eye gaze data., [Online]. Available: `http://ryanhope.github.io/gazetools/`.

[3] J. K. Lindstedt and W. D. Gray, "MetaT: Tetris as an Experimental Paradigm for Cognitive Skills Research," *Behavior Research Methods*, pp. 1–23, 2015. DOI: `10.3758/s13428-014-0547-y`.

[4] M. Destefano, J. K. Lindstedt, and W. D. Gray, "Use of complementary actions decreases with expertise," in *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*, L. Carlson, C. Hölscher, and T. Shipley, Eds., Austin, TX: Cognitive Science Society, 2011, pp. 2709–2014.

[5] D. Kirsh and P. Maglio, "On distinguishing epistemic from pragmatic action," *Cognitive Science*, vol. 18, no. 4, pp. 513–549, 1994. DOI: `10.1016/0364-0213(94)90007-8`.

[6] C. Sibert, W. D. Gray, and J. K. Lindstedt, "Tetris™: exploring human performance via cross entropy reinforcement learning models," *Topics in Cognitive Science*, submitted.

[7] C. Sibert, "Improving Novice Tetris Players with Feedback from AI Model Based Tutors," Master's thesis, Rensselaer Polytechnic Institute, Troy, NY, 2015.

[8] J. K. Lindstedt and W. D. Gray, "Extreme expertise: Exploring expert behavior in Tetris," in *Proceedings of the 35th Annual Meeting of the Cognitive Science Society*, M. Knauff, M. Pauen, N. Sebanz, and I. Wachsmuth, Eds., Austin, TX: Cognitive Science Society, 2013, pp. 912–917.

[9] ——, "Elements of expertise in Tetris: Asking "How do the moves of good players differ from those of novices?"" *to be submitted by 2015.10.19*, 2015.

[10] P. M. Fitts, R. E. Jones, and J. L. Milton, "Eye movements of aircraft pilots during instrument-landing approaches," *Aeronautical Engineering Review*, vol. 9, no. 2, pp. 24–29, 1950.

[11] S. R. Ellis and L. Stark, "Statistical dependency in visual scanning," *Human Factors*, vol. 28, no. 4, 421–438, 1986, ISSN: 0018-7208.

[12] K. Holmqvist, M. Nyström, R. Andersson, R. Dewhurst, H. Jarodzka, and J. V. D. Weijer, *Eye tracking: A comprehensive guide to methods and measures*. New York: Oxford University Press, 2011.

[13] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning with Applications in R*. Springer, 2014.

[14] C. Sibert, W. D. Gray, and J. K. Lindstedt, "Tetris: Exploring Human Performance via Cross Entropy Reinforcement Learning Models," in *Proceedings of the 38th Annual Conference of the Cognitive Science Society*, 2015.

[15] J. M. Carroll and M. B. Rosson, "Paradox of the active user," in *Interfacing thought: Cognitive aspects of human-computer interaction*, J. M. Carroll, Ed. Cambridge, MA: MIT Press, 1987.

[16] W.-T. Fu and W. D. Gray, "Resolving the paradox of the active user: Stable suboptimal performance in interactive tasks," *Cognitive Science*, vol. 28, no. 6, pp. 901–935, 2004. DOI: `10.1207/s15516709cog2806_2`.