SpotLight on Dynamics of Individual Learning

Roussel Rahman (rahmar2@rpi.edu) &
Wayne D. Gray (grayw@rpi.edu)
Cognitive Science Department, Rensselaer Polytechnic Institute
Troy, NY 12180 USA

Abstract

How does one person learn a complex task? Averaging performance over a group of individuals implicitly assumes that there is only one set of methods for accomplishing the task and that all learners acquire those methods in the same sequence. We maintain that the *average subject* is a mythical beast that does not exist. Hence, rather than profiling a mythical "average subject", we use *SpotLight* – a tool for analyzing changes in individual performance as skill is acquired in a complex task. Specifically, in this report, SpotLight uses 35 features and measures (some collected at millisecond level, others collected once per game), to investigate the skill acquisition of 9 players each of whom spent 31 hours learning the complex task of Space Fortress (SF). SpotLight enables us to uncover the evolution of individual strategies and the iterative efforts of individuals to create, discover, and explore new ways to improve their current performance. To our surprise, these players seem to have followed a common 'design for the weakest link' rule, in which after the current weakest link was strengthened a player's attention turned to the next weakest link. While this rule served our performers well, an often imposed constraint on the rule – 'while retaining existing strengths' – sometimes led the odd performer to suboptimal plateaus.

Keywords: Individual Learning; Part-task; Plateaus; Dips; Leaps; PDL; SpotLight; Relative entropy; Power Law

Introduction

Studies of skill acquisition often proceed by averaging data collected over large groups of individuals. Such methods are fine if we wish to measure the average effect of a treatment administered in different ways, but they fail to achieve our goal of understanding how individuals acquire complex skills.

Here we adopt the *Plateaus, Dips, and Leaps* (PDLs) approach advocated by Gray and Lindstedt (2017) and use the *SpotLight* tool (Rahman & Gray, in preparation) which enables us to identify the PDLs in individual performance. Our results show some commonalities in individual strategies amidst vast differences. For example, after varying numbers of hours of practice, all players adopted an optimal, but effort-expensive strategy. However, the most striking commonalities are not in the gameplay strategies *per se*, but the ways and means in which these strategies were modified. Specifically, we observe that the changes in each player's gameplay strategies pivoted around part-tasks in which the player was performing well. Relative to these pivots, gameplay strategies were recurrently modified to address the weakest parts of gameplay. Based on these findings, we propose that task execution strategies were recurrently updated by a common *optimize strategies for the weakest links* rule.

By focusing on one subgoal at a time, this rule provides checkpoints towards devising optimal strategies for the whole task. However, an excessive focus on subgoals may lead performers to lose sight of the overall goal (maximize total returns) and adopt strategies that maximize the current subgoal at the expense of the overall goal. For example, as we elaborate later, our worst performing player, in attempting to reinforce his skills in one subtask, adopted a suboptimal strategy that basically contradicted the entire point of the game and led to a plateau of stable suboptimal performance (Fu & Gray, 2004) from which there was no path forward. His best alternative would have been to discard the results of a long branch of exploration and strategy development, and revert to a much earlier set of strategies; which he did not do and fell victim to a type of sunk cost fallacy (Sweis et al., 2018).

Learning in Simple vs Complex Tasks

The research literature seems bifurcated between simple and complex tasks. For simple tasks, the power law of practice does a fine job of modeling learning (Newell & Rosenbloom, 1981). Those who follow in this tradition have complicated the world a bit (very reasonably) by proposing revisions that incorporate strategy-specific power laws (Rickard, 1997; Delaney, Reder, Staszewski, & Ritter, n.d.; Donner & Hardy, 2015). However, the more complex the task, the more the number of subtasks, and/or the more alternative ways of implementing a subtask, the less we would expect one individual's choices to resemble another's.

A complex task encompasses a hierarchy of subtasks (Simon, 1962), where higher level subtasks consist of and serve as goals for lower level ones. Task complexity exceeds the sum of separate subtasks because of intermediate associations in the hierarchy, which also implies a number of alternative routes in the hierarchy to reach from bottom to top. Consequently, even if practice alone suffices to maximize performance for simple tasks, more complex tasks require identifying optimum strategies from many alternative strategies.

The number of subtasks and the sets of possible strategies for each subtask raises new questions as to how learning progresses with practice. Does the individual attempt to optimize all parts of the task? Considering the limited cognitive and physical resources available to performers, it is reasonable to expect at least some parts to be *satisficed* (Simon, 1947). How choices are made as to which parts are satisficed or optimized and how such choices affect the goals and ultimately performance, are questions that directly relate to the dynamics of individual learning. To explore answers to these questions, we put the SpotLight on individual performance in the complex task of SF (Mané & Donchin, 1989).

The SpotLight Tool

The SpotLight tool (Rahman & Gray, in preparation) reveals changes in the execution of individual tasks and subtasks by detecting the PDLs in individual performance. The tool is instrumented with relative entropy (denoted by *RE* in Equation 1), an information-theoretic measure of the difference between a target probability distribution (p_i) and a reference distribution (p_r); in other words, it measures the difference between two states of uncertainty (Vedral, 2002). Whereas the scope of comparison in relative entropy is limited to two distributions, the scope is extended in the SpotLight to a finite number of distributions (see Equation 1). First, longitudinal records of performance (univariate or multivariate) are discretized into n consecutive phases and converted into phase-specific probability distributions. Then, a stable phase (i.e., its corresponding distribution) of performance is set as the common reference, relative to which relative entropy of each target distribution in each phase is calculated. Therefore, the output from the SpotLight is a relative entropy curve consisting of n points. This way, information of systematic changes in performance is retained in the relative entropy curve as differences from the stable reference. For details and demonstrations of the SpotLight, please refer to Rahman and Gray (in preparation).

$$RE(p_i||p_r) = \int_X p_i(x)\ \log_2\left(\frac{p_i(x)}{p_r(x)}\right)\ dx \tag{1}$$

Indices of targets: $i = 1, 2, 3, \ldots, n$

Index of common reference: $r\ (1 \leq r \leq n)$

SpotLighting at Different Levels of Granularity

In the relative entropy curve (e.g., in Figures 2 and 4), general improvement of performance with practice is captured by a continual decrease of relative entropy, and the periods of PDLs are identifiable as exceptions from this general trend. Specifically, during plateaus – periods of non-improvement with practice – relative entropy remains steady; during dips, relative entropy temporarily increases; and during leaps of performance, relative entropy sharply drops. Because the SpotLight models performance recorded by any measure through a single variable, relative entropy, individual performance in a complex task can be compared and investigated across levels of granularity. Therefore, a strategy change affecting the higher-level measures of performance (e.g., the Total score in SF) can be investigated further in lower levels (e.g., number of Fortress kills, use of resources, spatial locations of player's ship) to identify the subtasks associated with the strategy change.

Relative Entropy versus More Common Measures

To explain the choice of relative entropy over other more common measures (e.g., moving average, cumulative sum or coefficient of variation), its relativity property mitigates random noise from analysis (Rahman & Gray, in preparation). That is, random noise present in both the target and the reference distributions is eliminated in relative entropy. Moreover, relative entropy compares entire probability distributions, enabling more efficient use of the information present in the data. Finally, the probabilistic approach also enables future works in other rigorous frameworks (e.g., Bayesian updating or Kolmogorov equations for stochastic processes) to explore evolution of probability distributions with individual learning.

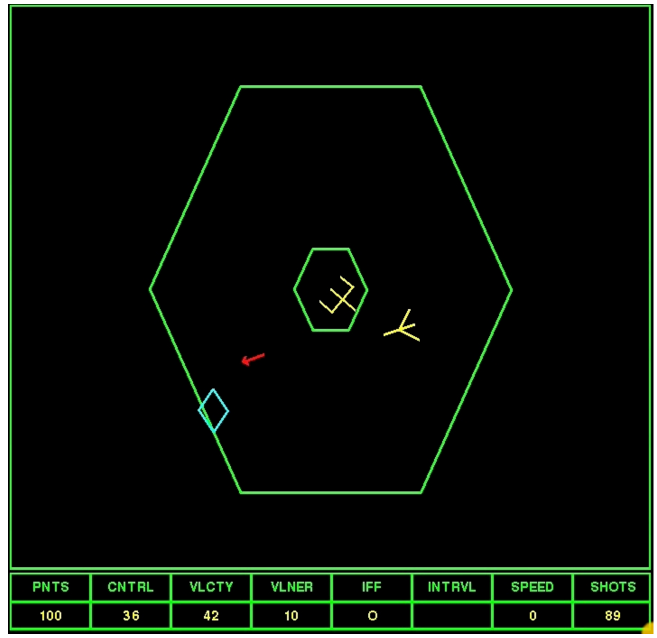

Figure 1: Space Fortress 4 (Destefano, 2010). Screenshot showing the Space Fortress in the center, the player's OS having recently fired a missile (red) at a mine (blue diamond).

Space Fortress: A Complex Task

Each game of SF lasts 5 minutes, where the player battles the Fortress. The player flies a ship ('Own Ship' or OS) carrying a limited number of missiles in a frictionless environment (Figure 1). The Fortress, fixed at the center, can rotate to shoot at OS. The mines (minions of the Fortress) periodically spawn to home in on OS. The mines are of two types which are only distinguishable by a three-letter code shown once at the start of each game. After a necessary identification step, one missile hit kills a mine. Contrastingly, killing the Fortress has two steps. First, 10 missile strikes make it vulnerable (with an inter-strike interval > 250 ms, failure to maintain the intervals results in full recovery); then, a double-strike (with an inter-strike interval < 250 ms) kills it. Conversely, OS is destroyed after 4 hits from either the Fortress and/or the mines. After being destroyed, OS or the Fortress immediately regenerates and the battle continues. At random intervals, the player receives opportunities to choose between receiving a bonus of 100 points or 50 missiles. The time to notice and to act to receive the bonus is limited. If OS' arsenal is empty, the player can gain more missiles at the cost of 3 points for each one.

The objective of the game is to maximize the Total score, consisting of four subscores – Points, Speed, Control and Velocity – measuring performance in different subtasks. In turn, each subscore consists of even lower-level measures (e.g., speed of killing mines, flying OS inside the large hexagon). For details of scoring rules, please see Destefano (2010).

Review of Relevant Works

Mané and Donchin (1989) developed SF as a common task for psychologists to use in comparing the effectiveness of different training paradigms for skilled performance. For example, in the *emphasis change* study by Gopher, Weil, and Siegel (1989), the experimental group was instructed to prioritize parts (OS control, OS velocity and mine handling) while training in the whole task. In contrast, Frederiksen and White (1989) adopted a direct part-task training approach by building up from small to more integrated subtasks. Despite treatment differences, both experimental groups benefited from specialized training and scored significantly more in post-test than the control groups.

More recently, Boot et al. (2010) employed Variable Priority Training (VPT), a variant of training with emphasis change, and found results consistent with the earlier findings in terms of accelerated learning. Lee et al. (2012) combined part-task training and VPT in a Hybrid Variable priority Training (HVT) regimen, to also show accelerated learning. Interestingly, again using HVT, Lee et al. (2015) showed that training strategy can compensate for intelligence differences within a group of individuals. Together, these works indicate that learning is aided by complexity reduction through training or emphasizing various parts of the whole task.

Finally, Destefano and Gray (2016) provide a prequel to this paper in that they used the PDL framework to uncover previously unknown individual strategies that even the designers might not have foreseen.

Methodology

We use the dataset from Destefano (2010), that contains highly detailed records (~ 40 measures) of nine players over 31 hours. Each individual played 8 games in each 1-hr session per day for 31 days, resulting in total 248 games per player. Experimenter instructions included rules and objectives of the game and some general suggestions of optimal gameplay. We exclude the 8 games from the first day, as the players needed time to get familiar with the complex rules. Therefore, the final dataset contains 240 games per player.

Due to space constraints, we demonstrate the SpotLight analyses of the Total scores of two example players (Figures 2 and 4) and provide a summary of lower level analyses of all nine players. For the Total score, we use a sliding window approach (span = 20 games) to discretize each player's performance into 221 windows and convert measures in each window to a normal distribution. The span of 20 games is chosen to estimate distributions reliably with sufficient samples (more would be better, but that means less number of windows). We use the last window (of games 229-248) as our reference, because it is the most stable phase according to the power law.

A drawback of the sliding windows approach is that each game is included in a number of successive windows, therefore, the changepoints shown in the relative entropy curve may shift within a range of [0, window span/2]. We use the sliding window approach for most but not all of our analyses. For example, for several low-level measures (e.g., spatial locations of OS or OS velocity), 9000 samples were collected at 30 Hz frequency from each 5-min game. Therefore, the sliding window approach is not necessary, and the SpotLight analysis is performed by fitting normal distributions to each game's data and taking the last game as the reference.

Strategy Shifts of the Best Performing Player

Figure 2 shows the relative entropy curve (red line) of the Total score (blue line) for Player 7. Note the two periods of dip+leap in the Total score (in the shaded regions in Figure 2); both dip periods are indicated by increased relative entropy (green- and gray-shaded) and each leap period is indicated by rapid drops of relative entropy (red- and yellow-shaded). A dip followed by a leap indicates performance improvement from shifting to a new strategy that implements better goals with the dip revealing a temporary performance decrement as the new strategy is learned (Gray & Lindstedt, 2017).

Importantly, the Total score is the aggregate of all performance measures; to identify the details of strategy shifts, performance in lower-level subtasks was investigated in the same manner (not included here). We found that Player 7's strategies were centered on flight-related tasks. Here we discuss our findings of the two strategy modifications that had the largest impact on Player 7's Total score.

The first dip+leap shown in Figure 2 stems from Player 7 adopting a strategy of *flying in small circles around the Fortress* at the 81^{st} game (Figure 3d). The tightness of the flight path in Figure 3e vs Figure 3d shows the rapid improvement Player 7 made across just 7 games. Once adopted, this strategy was maintained (with minor improvements) to the last game (Figure 3f).

Destefano (2010) and Towne, Boot, and Ericsson (2016) separately observed expert players to adopt these circular

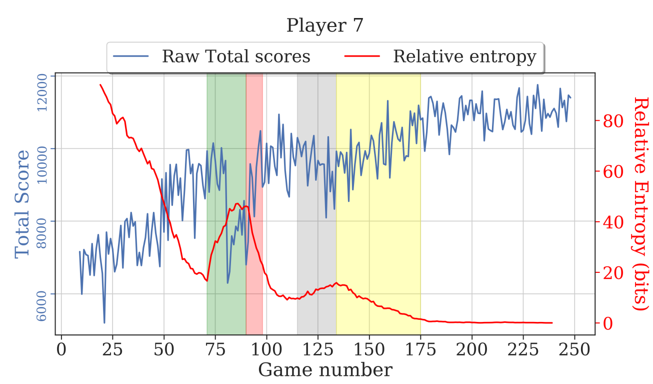

Figure 2: Performance of our best player, Player 7, through Total score and its relative entropy curve. Green- and gray-shaded regions denote two dip periods; red- and yellow-shaded regions show the two leaps that follow the dips.

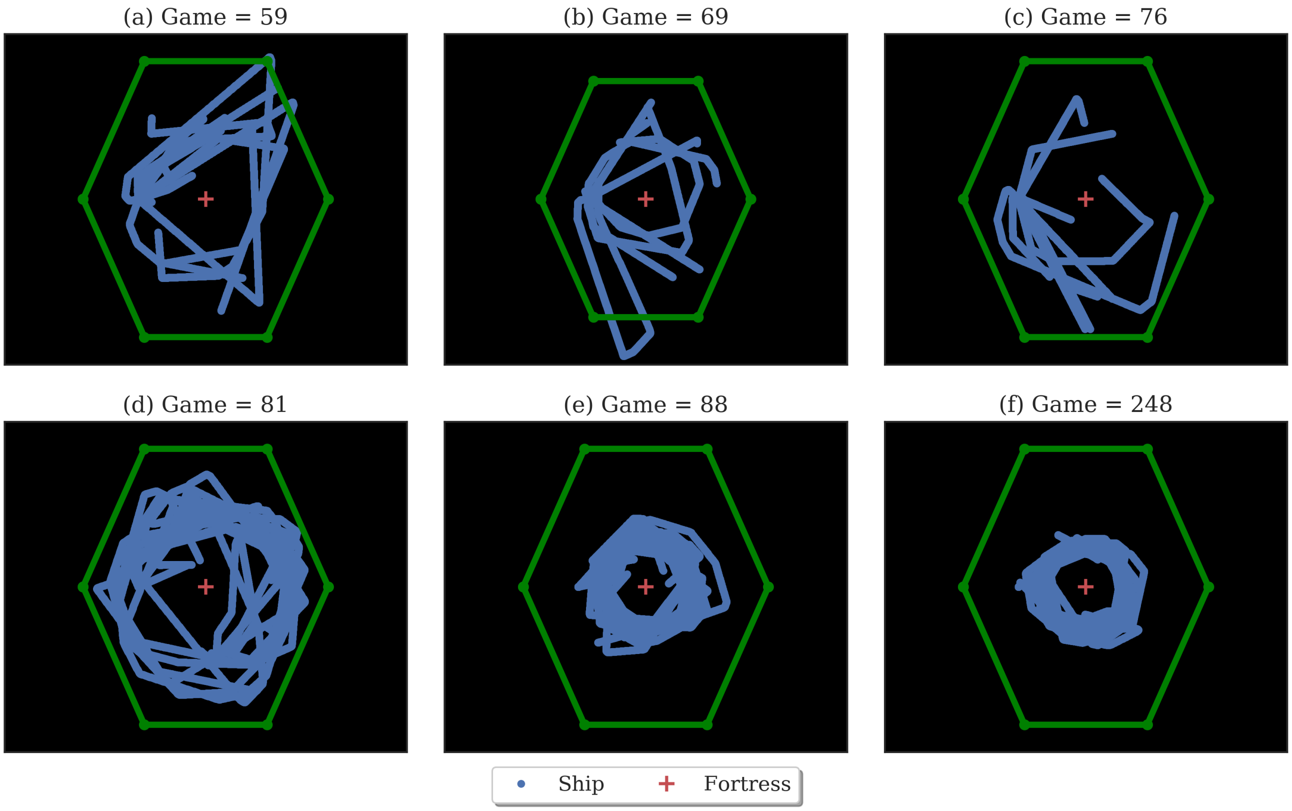

Figure 3: Distributions of OS location in six example games of Player 7: (a)-(c) illustrate explorations of optimal flight path and velocity; (d) shows the 81^{st} game, where Player 7 shifted to a strategy of circular paths around the Fortress; (e) shows the 88^{th} game, illustrating fast improvements within 7 games; (f) shows that the strategy was maintained till the end of practice.

paths. Flying in circles is beneficial as it maximizes opportunities to attack the Fortress and increases predictability of the Fortress' behavior (Rahman & Gray, in preparation). However, maintaining circles requires precise synchronization between acceleration and rotation of OS. Therefore, before benefitting from the circular path strategy, Player 7 needed to master "circular flying".

Importantly, rapidly decreasing relative entropy before the first dip+leap started (Left of the green-shaded region in Figure 2) indicates that the player was improving quite fast even before changing strategy. On the other hand, impact of the strategy change was enormous; for example, one subscore dropped by 98.9% (from 3424 to just 39 in 80^{th} and 81^{st} games, respectively). What the Total score does not and cannot show us, is that the player extensively explored different flight paths (Figures 3a-c) in the ~ 30 game period prior to the green-shaded period (in Figure 2). Presumably, Player 7 had realized flight patterns being a weakness in his otherwise strong game, before investing effort to perfect it and restructuring other aspects of gameplay accordingly.

At the second dip+leap (gray- and yellow-shaded regions), Player 7 tweaked the circular flight path strategy by adding *flying OS slower* to it. A low velocity is especially helpful for aiming at the moving targets (i.e., mines) and for making tiny movements to evade hits from mines to OS without swaying too far from the circles.

Strategy Shifts of the Worst Performing Player

Figure 4 shows the relative entropy curve (red line) of the Total score (blue line) for Player 2. Notably, unlike Player 7, Player 2 shows no major dips in performance. Rather, the two biggest points of discontinuity in the relative entropy curve (asterisked in Figure 4) denote the start of two leaps of performance. Absence of dips before leaps indicates that the costs of adopting new strategies were not high enough to cause dips (Gray & Lindstedt, 2017). SpotLight analyses of Player 2's performance in lower-level subtasks (not included here) reveal that the player's strategies pivoted around killing mines. Interestingly, Player 2 flew in circles around the Fortress since the beginning of practice, but possibly without realizing the benefits or acquiring the skills to utilize the strategy.

At the first point of discontinuity, Player 2 adopted a strategy of *flying OS slower* (same as Player 7's) that improved the player's mine killing performance. However, unlike Player 7, the strategy did not aid Player 2 much in protecting OS, as it was the Fortress causing the most damage. The player

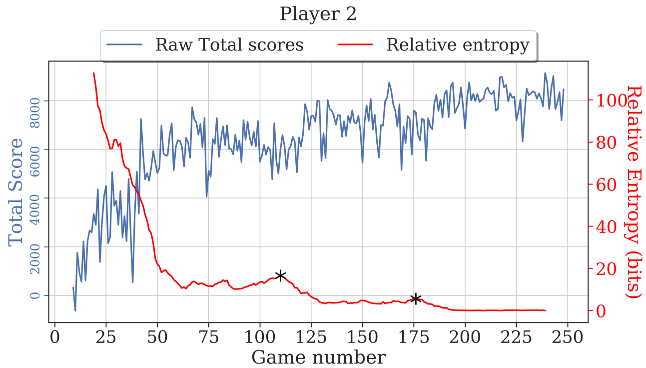

Figure 4: Performance of our worst performing player, Player 2, through Total score and its relative entropy curve. The two asterisks denote the start of two leaps intermediate in practice. (Note: Scales are different from Figure 2)

Table 1: Impact of Player 2's shift to *lazy* strategy.

Measure	Before*	After*
Fortress kills	13.6 (2.8)	4.8 (1.4)
Mine kills	30.6 (2.7)	35.9 (2.3)
Missiles fired	318 (54.9)	99.3 (27.7)
Missiles fired with penalty	154 (80)	8.3 (12.4)
OS destroyed	2.7 (1.3)	0.8 (0.7)
Total Score	7091 (728)	8067 (517)

* Mean (SD) in 50 game blocks

was shooting a lot of missiles at the Fortress, almost half of which were wasted. Consequently, the player was spending more than anyone else in the group to buy necessary missiles.

The second point of discontinuity corresponds to a strange strategy Player 2 adopted to address these weaknesses: *Minimize exchanges with the Fortress (!), save missiles, and kill as many mines as possible*. This *lazy* strategy, despite its extreme ingenuity, contradicts the whole point of the game. The player adopted a flight pattern of bigger circles (i.e., away from the Fortress) to get more time to move away from the Fortress' line of fire. In the process, the player lost a big source of points (100 points/Fortress kill). Nevertheless, the strategy markedly improved the player's ability to protect OS and manage OS' arsenal (Table 1). Additionally, the strategy helped the player to focus resources on the strength of killing mines. Therefore, the differential gain from the *lazy* strategy was positive, and the Total score improved by almost 1000 points.

Common Strategies among All Nine Players

Experimenter instructions included the suggestions of flying slowly in small circles. Therefore, it is not surprising that all nine players adopted the circular paths. However, only five adopted the strategy within the first 50 games; the rest experimented with flight patterns deep into practice, with one player taking as late as the 160th game to adopt flying in circles. The players were more varied in terms of smallness of circles and slowness of flying, possibly because these suggestions were less objective. The flight-related scores – Velocity and Control – are prone to ceiling effect, therefore do not portray improvements in these two respects beyond a certain point. But, generally, slower velocity in smaller circles around the Fortress resulted in higher Total scores.

We observe another common strategy in optimizing arsenal management. The game starts with 100 missiles in OS. There are two options to get more missiles: (1) Bulk option at 2 points/missile – choose 50 missiles instead of 100 points when bonuses are available, and (2) Retail option at 3 points/missile – fire a missile with penalty when arsenal is empty. Though the bulk option is better, it may result in surplus missiles if taken late in the game. Therefore, the optimal strategy is to switch to the retail option for missiles and take bonus points (instead of missiles) towards the end of a game. Six of the 9 players discovered this strategy on their own.

None of the players discovered the strategy of switching from taking bonus missiles to bonus points before adopting the circular flight paths. We believe this is due to the fact that gains from optimum arsenal management are measured in hundreds of points, whereas gains from regulating the flight pattern are measured in thousands of points. Hence, the gain in points from switching from bonus missiles to bonus points is harder to notice than gains from changing the flight pattern.

The six players who discovered the optimal strategy of managing arsenal, also show similar steps in progressing towards the optimal strategy. Each of them first relied on the retail option only, before switching to the bulk option only and then, finally reaching the optimum balance between the two. The reason can be understood by walking through the possible steps in a player's learning. At the start, a player is weak in every aspect of the game and the main focus is to just learn how to kill the Fortress and mines. Because even taking bonuses is not quite simple in SF, each player initially relies on the retail option. This does not become a big issue until a player becomes very good at killing the Fortress and mines, and needs more missiles. At this point, the primary choice becomes the bulk option. Finally, when the player has maximized returns from larger resources of points, not losing points through unused missiles comes to the fore.

'Design for the Weakest Link' Rule

'Design for the weakest link' is a principle often adopted in engineering design problems. The concept is to specify design parameters to address the weakest point in a machine. An analogy to this concept applies to our players. Note that the instructions included suggestions of optimal play, but the individuals themselves had to decide on the order they would learn the numerous subtasks and update strategies to realize these suggestions. Each new strategy a player adopted, addressed the weakest link of performance; by weakest, we mean the scope in gameplay with maximum potential for improvement. However, an additional constraint we find is that the new strategy must reinforce existing strengths; by strengths, we mean the parts of gameplay closer to being optimal. It is possible that performers use the subgoals of addressing weakest links and reinforcing existing strengths in part-tasks as checkpoints towards the global optimum of the whole task.

A successful proponent of the rule is our best performer, Player 7. Despite improving fast, the player reorganized their entire gameplay around a strategy of flying in small circles around the Fortress. Though the reward system may not show it, all tasks are not equally influential in the game. For example, skills in killing the Fortress crucially depends on flight pattern, but not the other way around. It is likely that the player realized that determining the best flight pattern is crucial and strove to make it a strong point. Once acquired, the player maintained this strategy, but made smaller refinements to address other weaknesses.

Similar to the best player, our worst player (Player 2) also pivoted strategies around his strengths to address the weakest links in gameplay. But, in effort to reinforce strengths, the player adopted a suboptimal strategy that worked well in the short-term, but would never lead to maximal performance even after an infinite amount of practice. This strategy demonstrates that excessive emphasis on the subgoal of reinforcing strengths can lead performers to local optima, instead of the global one; that is, to performance plateaus rather than performance asymptotes.

The 'design for the weakest link' rule extends to the whole group. First, the players followed the same order, without exception, in adopting the two optimal strategies – respectively for flight pattern and arsenal management. This order fits into the rule nicely, that the players simply addressed the weakest links first. Second, even in terms of managing arsenal only, the players went through the same steps to reach the optimum. All requiring several steps indicates that the players focused on a part only until it was not the weakest, but not necessarily optimal. In other words, the players were satisficing in part-tasks, with 'not the weakest' as the criterion of sufficiency. However, Player 7 does provide one exception, as the player optimized – not just satisficed – the weakest link in flight pattern and made it the strongest before moving on. Even then, it is quite possible that satisficing observed in players' gameplay are actually static points in the dynamics of reaching the optimum.

Summary and Conclusions

In this work, we put the SpotLight on the commonalities in individual learning of a complex task that underlie vast differences in performance. We observe that our players progress towards optimal strategies by recurrently applying the rule of 'design for the weakest link', while simultaneously reinforcing existing strengths. More comprehensively, the rule stands to be: *optimize strategies for the weakest links, but relative to existing strengths*. A resultant of adopting this common rule is that the individuals' very different routes to expertise tended to converge towards the same strategies. Therefore, a possible explanation for the rule is that optimizing strategies for the weakest links serves as checkpoints towards the globally optimum strategies that maximize the overall or ultimate goal. Although the rule served the performers well, we also observe that the constraint of *relative to existing strengths* on the rule may lead to local optima of strategies – instead of the global optimum – and therefore, to stable suboptimal performance.

The 'design for the weakest link' rule provides a simple explanation as to how individuals may progress in learning a complex task, and what may cause them to plateau. But, we do not claim that it to be an absolute general rule, especially with the scope of study being only one task. Rather, it serves as a demonstration of how the PDLs and strategies uncovered by the SpotLight, can aid in finding common patterns in the dynamics of individual learning. These patterns, in turn, would be useful to discover the laws that govern individual learning and finding ways of overcoming suboptimal plateaus to accelerate learning. Finally, our experimental paradigm of SF emulates an important characteristic of complex real-world tasks – numerous, interconnected elements resulting in many alternative strategies. Therefore, a promising direction for future research is to apply and test the SpotLight tool in investigating learning of complex real-world tasks to progress towards the general laws of individual learning.

Acknowledgments

Correspondence should be sent to Roussel Rahman, Cognitive Science Department, Rensselaer Polytechnic Institute, Troy, NY 12180. Email: rahmar2@rpi.edu. The work was supported, in part, by grant N00014-16-1-2796 to Wayne Gray from the Office of Naval Research, Dr. Ray Perez, Project Officer.

References

Boot, W. R., Basak, C., Erickson, K. I., Neider, M., Simons, D. J., Fabiani, M., . . . others (2010). Transfer of skill engendered by complex task training under conditions of variable priority. *Acta Psychologica*, *135*(3), 349–357.

Delaney, P. F., Reder, L. M., Staszewski, J. J., & Ritter, F. E. (n.d.). The strategy-specific nature of improvement: The power law applies by strategy within task. *Psychological Science*, *9*(1), 1–7.

Destefano, M. (2010). *The mechanics of multitasking: The choreography of perception, action, and cognition over 7.05 orders of magnitude*. Unpublished doctoral dissertation, Rensselaer Polytechnic Institute.

Destefano, M., & Gray, W. D. (2016). Where should researchers look for strategy discoveries during the acquisition of complex task performance? The case of Space Fortress. In *Proceedings of the 38th Annual Conference of the Cognitive Science Society* (pp. 668–673).

Donner, Y., & Hardy, J. L. (2015). Piecewise power laws in individual learning curves. *Psychonomic Bulletin & Review*, *22*(5), 1308–1319.

Frederiksen, J. R., & White, B. Y. (1989). An approach to training based upon principled task decomposition. *Acta Psychologica*, *71*(1-3), 89–146.

Fu, W.-T., & Gray, W. D. (2004). Resolving the paradox of the active user: Stable suboptimal performance in interactive tasks. *Cognitive Science*, *28*(6), 901-935.

Gopher, D., Weil, M., & Siegel, D. (1989). Practice under changing priorities: An approach to the training of complex skills. *Acta Psychologica*, *71*(1-3), 147–177.

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

Lee, H., Boot, W. R., Baniqued, P. L., Voss, M. W., Prakash, R. S., Basak, C., & Kramer, A. F. (2015). The relationship between intelligence and training gains is moderated by training strategy. *PloS One*, *10*(4), e0123259.

Lee, H., Boot, W. R., Basak, C., Voss, M. W., Prakash, R. S., Neider, M., ... others (2012). Performance gains from directed training do not transfer to untrained tasks. *Acta Psychologica*, *139*(1), 146–158.

Mané, A., & Donchin, E. (1989). The space fortress game. *Acta Psychologica*, *71*(1-3), 17–22.

Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. *Cognitive skills and their acquisition*, *1*(1981), 1–55.

Rahman, R., & Gray, W. D. (in preparation). Strategy changes revealed by plateaus, dips and leaps.

Rickard, T. C. (1997). Bending the power law: A CMPL theory of strategy shifts and the automatization of cognitive skills. *Journal of Experimental Psychology: General*, *126*(3), 288.

Simon, H. A. (1947). Administrative behavior.

Simon, H. A. (1962). The architecture of complexity. *Proceedings of the American Philosophical Society*, *106*(6), 467–482.

Sweis, B. M., Abram, S. V., Schmidt, B. J., Seeland, K. D., MacDonald, A. W., III, Thomas, M. J., & Redish, A. D. (2018). Sensitivity to "sunk costs" in mice, rats, and humans. *Science*, *361*(6398), 178+.

Towne, T. J., Boot, W. R., & Ericsson, K. A. (2016). Understanding the structure of skill through a detailed analysis of individuals' performance on the space fortress game. *Acta Psychologica*, *169*, 27–37.

Vedral, V. (2002). The role of relative entropy in quantum information theory. *Reviews of Modern Physics*, *74*(1), 197–234.