Proceedings Paper:

Exploration and Skill Acquisition in a Major Online Game

Tom Stafford
Department of Psychology, University of Sheffield
t.stafford@sheffield.ac.uk

Sam Devlin
Digital Creativity Labs, University of York
sam.devlin@york.ac.uk

Rafet Sifa
Fraunhofer IAIS, Germany
rafet.sifa@iais.fraunhofer.de

Anders Drachen
Digital Creativity Labs, University of York
andersdrachen@gmail.com

Abstract

Using data from a major commercial online game, Destiny, we track the development of player skill across time. From over 20,000 player record we identify 3475 players who have played on 50 or more days. Our focus is on how variability in elements of play affect subsequent skill development. After validating the persistent influence of differences in initial performance between players, we test how practice spacing, social play, play mode variability and a direct measure of game-world exploration affect learning rate. These latter two factors do not affect learning rate. Players who space their practice more learn faster, in line with our expectations, whereas players who coordinate more with other players learn slower, which contradicts our initial hypothesis. We conclude that not all forms of practice variety expedite skill acquisition. Online game telemetry is a rich domain for exploring theories of optimal skill acquisition.

Keywords: learning; games; skill acquisition; expertise; game analytics

Introduction

Computer games afford a rich data set for the investigation of skill acquisition. Players invest tens, hundreds or even thousands of hours on individual games, and — unlike offline domains of expertise — details of every action during practice can be unobtrusively recorded. The present analysis uses data from the online shooter video game Destiny, which has over 30 million active users as of 2016 (Nunneley, 2016). Using data on players' performance we trace their skill acquisition over time and relate it to their practice habits. Specifically we are interested in how *variability* in practice relates to learning.

The power law of learning is justly well-known in cognitive science (A. Newell & Rosenbloom, 1981; Ritter & Schooler, 2001), both for being a dependable regularity in skill acquisition data (Rosenbaum et al., 2001) and for expressing a truth we know from personal experience: when we first begin learning something new progress is often rapid, but later it slows or stalls. Nonetheless the presentation of Power Law learning curves based on averages masks the variability that occurs both within and between individuals (Gallistel et al., 2004; Gray & Lindstedt, 2016). This is important for two reasons. Firstly because individual variability is interesting as an outcome. We wish to know why some individuals learn more rapidly, and achieve greater eventual levels of performance (and why some individuals are hindered in their learning). Secondly, variability is interesting as a driver of learning. Previously it has been suggested that greater initial variability in practice may drive higher subsequent performance (Stafford et al., 2012; Stafford & Dewar, 2014), a result which accords with computational accounts of how learning must balance exploration and exploitation of options (Sutton & Barto, 1998; Humphries et al., 2012).

In addition to looking at *how* a skill is practised, there are also results which suggest an effect on skill acquisition of *when* a skill is practised — the issue of practice spacing (Stafford & Haasnoot, 2017; Delaney et al., 2010; Cepeda et al., 2008) — as well as an effect of variability in how different components are practised (Magill & Hall, 1990). From this perspective, variability is as much an engine of learning as consistency (Schmidt, 1975; Van Rossum, 1990; K. M. Newell & McDonald, 1992; Ranganathan & Newell, 2010). This raises the question of exactly which kinds of variability, and in what quantities, support optimal skill acquisition.

Previous work has looked at skill learning in a simple online game (Stafford & Dewar, 2014; Stafford & Haasnoot, 2017), with the emphasis that even a simple online game contains many fundamental cognitive processes - perceptual, decision making and action implementation. Others have looked at skill development in more complex games (Thompson et al., 2017, 2013), and here we use the opportunity to analyse data from one such game, *Destiny*, to explore issues of how playing style, and particularly variability within play, affects skill development.

Destiny is a science-fiction themed, massively multi-player, online game where players need to defend the Earth from various alien threats, taking on the role of Guardians. Players journey to different planets, complete missions, daily events, and perform a variety of different tasks to build up their characters. Destiny is a hybrid digital game that blends features from a number of traditional game genres including role-playing games and massively multi-player online games but which is first and foremost a shooter (Tammasia et al., 2016). The main components of the gameplay is focused on tactical single-player or small-team combat against players or artificial agents (Drachen et al., 2016).

Thousands of behavioural or performance-based metrics are tracked and stored by Bungie, the developer of Destiny, which in aggregate provides a detailed record of the behavioural history of Destiny players.

The metrics that can be calculated based on such datasets varies, and previous research in game analytics and other domains have seen such behavioural data being used for a variety of purposes (Tammasia et al., 2016; Rattinger et al., 2016; Drachen et al., 2016). For Destiny, a number of these metrics are of key interest in relation to evaluation of player skill and skill evolution.

- *Playtime*: Playtime in the current context simply refers to the amount of time a player spends playing the game per day, across either a single or all characters.
- *Kills, Assists, Deaths*: the shooter-heavy gameplay of Destiny means that traditional skill indicators from shooter games such as Kill/Death Ratio (KDR) form an important means for evaluating player skill.
  For Destiny, a variant of KDR, the Kill-Assists/Death Ratio (KADR) is also used. An *assist* is a common term in esports signifying that a player helped another player take down a specific enemy (or in other ways help another player), without scoring the killing shot/hit on that enemy.
  KADR is thus a more nuanced aggregate measure of performance than KDR. We use KADR-KDR as a measure of a players' propensity for 'social play'.
- *Combat Rating*: The Combat Rating (CR) is a game metric designed to reflect a players' overall skill. How CR is based on the TrueSkill system (Herbrich & Graepel, 2006), a Bayesian model used for player/team ranking. TrueSkill and CR both serve a similar functionality to ELO (Charness, 2005). While the algorithm is confidential, it broadly works by initialising a player at CR 100. If the player is part of a team that wins a match, their CR goes up, more if there is a large difference in the CR between the two teams. Conversely, if they lose, the CR goes down, again in relation to the gap in CR between the two teams. CR is used by the Destiny matchmaking system to configure players into teams and balancing opponents. This means that players will be playing with and against players with similar CR (i.e. they are matched against players of simialr skill-levels).
- *Grimoire Score*: A Grimoire in Destiny is a record of a players experience — new cards are awarded the first time a specific action is taken or challenge overcome. In essence, the Grimoire score is an expression of the degree to which a player has explored the world and content of Destiny.

Working with very large datasets introduces some new opportunities for the cognitive scientist (Goldstone & Lupyan, 2016; Stafford & Haasnoot, 2017). Observational studies, however large, necessarily have reduced power of causal inference compared to experimental studies. Large numbers mean that the data can be 'sliced' to explore if and how potential effects play out throughout the population, as well as allowing matching of individuals on various properties which might confound any effect. With enough data any observable difference can be 'statistically significant'. In experimental studies effort is expended in achieving enough power to make convincing inferences. With large data set it is more important to invest effort in exploring possible confounds and putting observable differences in the context of other effects via calculation of effect sizes.

Our hypothesis is that early variability will be associated with faster skill acquisition. This assumes that players have a tendency to under-explore the space of possible actions, and so, due to this reliance on habit, will be learning sub-optimally. We will test this hypothesis against different indices of variability in early practice: spacing of play, social play, world knowledge (grimoire score), and distribution of play across game modes (event entropy). These metrics are defined further below.

Data and method

Our data comprise low level daily metrics indicating performance and meta information for over 20,000 randomly selected Destiny players. The behavioral telemetry was provided by Bungie.

For each player the data consists of a unique player ID and character IDs for each character the player has. A player is allowed to have at most three characters per account. For each character, the dataset contains daily aggregate player behavior such as number of deaths, completed objectives, weapon usage and average life span, and importantly playtime, each across the six game modes - or ways to play the game.

Our analytic strategy is first to split the data into a development (8682 players) and validation set (12861 players). All exploratory analysis was finalised on the development set, before being run on the validation set to produce the figures presented here. All conclusions presented are unaffected by the minor differences between the development and validation set results. This affords us some protection against discovering false patterns in our data that result from researcher degrees of freedom in analysis. It is inappropriate to make inferences from hypothesis testing p-values for an exploratory analysis such as this, but we report them for completeness where we have done standard analyses. Our main focus is on measures of effect size and confidence estimates around those measures.

Analysis scripts, player summary data, and full reports of both development and validation set results are available at https://osf.io/c59n9/. For commercial con-

fidentiality reasons the full raw dataset is not available at the point of writing.

Analysis

First, we seek to confirm that players improve with practice. Following the method of (Stafford & Dewar, 2014), we first select only players who play some minimum number of games (50). This produces a data set of 3475 longer term players (in the validation set; 1984 in the development set) and then divide by ranking all players according to the average of their three all time best ratings (in terms of CR). Figure 1 shows the average score per game for those who scored in the top third, middle third and lowest third of the high score table. This shows that the learning curve exists for averaged data, and that — in line with (Stafford & Dewar, 2014) — players who end up with the highest scores begin the game with performance already above that of lower scorers (compare (Stafford & Dewar, 2014) Figure 2).
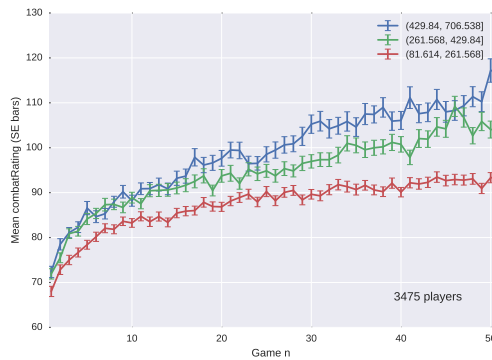

Figure 1: Average performance rating as a function of game number and ranking based on players' highest three ratings. Error bars show +/- 1 standard error.

Note that our learning curves show performance, rather than speed, on the x-axis, and so are inverted relative to the classic 'Power Law of Learning'. None the less they reflect the expected decelerating function of learning with practice amount. Our investigation of other factors must take account this fundamental pattern in how player performance changes over time, as well as the stratification that we observe between players of differing initial performance. To do this, we fit a linear regression for each player's performance against game number. Because this regression produces a slope and an intercept, we are able to subsequently analyse player differences in both level of initial performance and subsequent change in performance (i.e. rate of learning). Henceforth when we refer to "learning rate" we mean the slope of this regression for each player. In order to explore which variables might be related to player learning rate we first visualise players split on some candidate variables against mean combatRating against prac-

tice amount (game number).

Variation in practice timing — spacing

In order to compare practice timing, we calculate the time range over which players recorded their first 25 days of play (obviously this has a minimum of 25 days, and no theoretical maximum). This range correlates positively (Pearson's r= 0.18, 99% confidence interval 0.14,0.22) with learning rate and negatively (Pearson's r= −0.09, 99% confidence interval −0.14,−0.05) with initial performance.

In order to visualise the effect of greater or less spacing, we select players in the top quartile for spacing their first 25 games ('spacers') and those in the bottom quartile for spacing their first 25 games ('groupers') and plot the average performance against game for the two groups. This is shown in Figure 2.
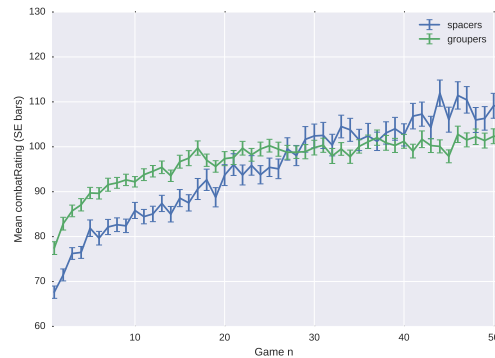

Figure 2: Average score as a function of game number and high and low spacers. Error bars show +/- 1 standard error.

Variation in practice type

Playing style — social play   For each player we have a game by game measure of their 'assists', which are kills made by teammates which they are near to. Variation on this measure allows us to rate players according to a propensity for social play, i.e. a higher rate of assists will reflect a player who coordinates their actions with their team.

This measure correlates negatively (Pearson's r= −0.16, 99% confidence interval −0.20,−0.12) with learning rate and positively (Pearson's r= 0.50, 99% confidence interval −0.47,0.54) with initial performance.

Figure 3 shows the learning curve for players split on the average of their assists over their first 25 games, as an index of players' propensity for "social play". Those in the top quartile of the distribution of assists we term 'social players'. Those in the bottom quartile we term 'lone wolves'.
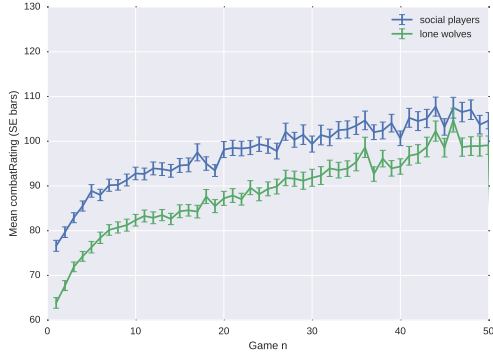
Figure 3: Social play and skill acquisition. Error bars show +/- 1 standard error.

Figure 4: Average score by games for players with high and low grimoire count. Error bars show +/- 1 standard error.

World knowledge — Grimoire score   For each player we are able to see the complete history of their Destiny playing, including how many games they play in total. Each player also has a 'grimoire' score, which is a count of the items they have encountered in the world. Obviously this is higher for players who have played more games, but there is considerable between-player variability, suggesting that some players focus on exploring the world, completing actions and collecting items, whereas others aren't focused on this aspect of the game. In order to compare grimoire scores between players who have complete different numbers of games, we calculate a normalised (Z) score for each player based on the distribution of grimoire scores among players who have completed the same number of games.

This measure does not correlate with learning rate (Pearson's r= 0.04, 99% confidence interval 0.00,0.09) and correlates positively, but weakly (Pearson's r= 0.13, 99% confidence interval 0.10,0.18) with initial performance.

Figure 4 shows the average score, in terms of CR, against game for players whose grimoire Z scores are in the top and bottom quartiles of the distribution.

Playing style — mode entropy   The play modes in Destiny are:

- *Strikes*: 3 player cooperative events.
- *Raid*: 6 player cooperative missions, requiring high level skills to complete.
- *Story*: the main single-player game mode, which can be played cooperatively by up to 3 players.
- *Patrol*: a single-player exploration mode.
- *PvP*: all the player-vs-player (PvP) game modes of Destiny

Note that due to the aggregation into daily sets, it is possible for players to have played multiple sessions of Destiny within the same 24 hour cycle. Because Destiny
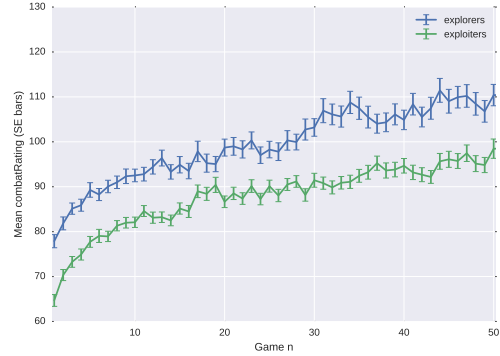
has six different main game modes, it is of interest to evaluate how a player spends his or her time across those game modes. In order to quantify the measure of heterogeneity in terms of how a player splits their time between game modes, we use Shannon's entropy [see e.g. (Lessne, 2014; Algoet & Cover, 1988)] which is defined as:

$$H = -\sum_i p_i \log_2(p_i) \tag{1}$$

where p_i denotes the probability of the player's activity across game modes i. For game mode p_i is calculated as the amount of time spent in specific game mode i divided by the total time spent playing all game modes that day.

Event entropy over the first 25 games for each player does not correlate with learning rate (Pearson's r= -0.02, 99% confidence interval $-0.06,0.03$) and correlates positively (Pearson's r= 0.22, 99% confidence interval 0.17,0.26) with initial performance.

Figure 5 shows performance against game for those in the top and bottom quartiles for event entropy calculated over the first 25 games.

Statistical model   Hitherto, we have explored our data using visualisation of different groups and reported bivariate correlations. By entering all factors into a regression model we can check whether how all factors combine to explain variation in the learning rate. This is an essential complement to the visualisation. It allows us to confirm that patterns visible in the data are statistically significant. As well as the four measures described above — spacing, social play, grimoire score and event entropy — we include maximum numbers of games a player plays as a measure of overall motivation. The results of the regression of our five factors against the learning rate are shown in shown in Table 1.
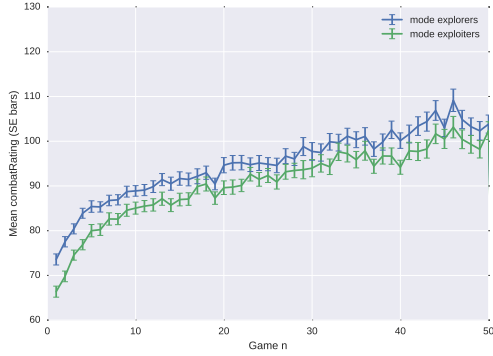
Figure 5: Play mode entropy and skill acquisition. Error bars show +/- 1 standard error.

Table 1: Regression of player behaviours on player learning rate

Factor	B	T	p
Games played	0.044	1.99	0.0465
Spacing	0.199	10.90	0.0001
Assists	-0.172	10.04	0.0001
Grimoire	0.003	0.16	0.872
Event entropy	0.011	0.62	0.537

$R^2 = 0.063, F(5, 3287) = 44.47, p < 0.0001$

Note that only spacing and assists, our measure of social play, are significant. Figure 6 shows the standardised regression coefficients (beta weights) when our five factors are used to predict learning rate (slope) and for the initial performance (intercept) of individual learning functions.

Discussion

Using a complex online game we show that changes in player's performance can be tracked and related to aspects of how they play. We validate the separation of learning curves by initial performance shown by
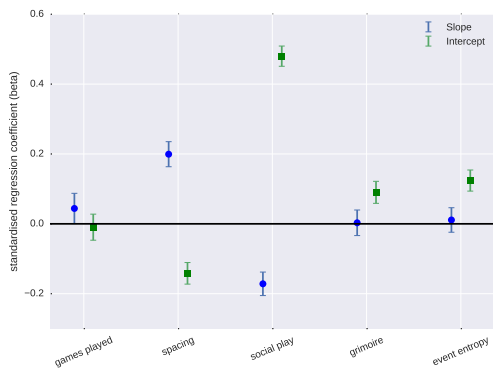

Figure 6: Beta weights for players factors used to predict slope (learning rate) and intercept of individuals' learning functions. Standard error bars shown.)

(Stafford & Dewar, 2014). As with that previous result, players who achieve the eventual highest levels of performance also perform better on their first game. Further, the difference between those with high and low initial performance only grows as more practice is completed.

Two other factors influence rate of learning — spacing, and social play. The effect of spacing matches that found in experimental studies of skill acquisition, as well as previous analysis of a different, simpler, game (Stafford & Haasnoot, 2017). The differences between players who space their practice and those who don't is striking, such that the high-spacing players, on average, perform less well initially, but because they learn at a faster rate their average rating exceeds that of the low-spacing players by game 50. The effect of social play was not predicted: those who play more socially, as measured by their assist rate, learn slower — perhaps because the demands of team coordination distract from skill honing. Two other direct measures of exploration are not found to relate to rate of learning, in contrast to earlier results (Stafford et al., 2012; Stafford & Dewar, 2014). This suggest that curiosity alone is not sufficient to enhance skill acquisition.

Destiny, and online games in general, represent a rich test-bed for theories of skill acquisition. Games are played for reasons of intrinsic motivation and so represent an important contrast to lab studies which are completed for financial incentives or as part of a course requirement. In addition they represent an opportunity to collect large data sets, which allow confidence in the estimates of the effects of the factors analysed. Overcoming statistical uncertainty allows researchers to move swiftly to wrestling with interpretative uncertainty.

In this case, although variability in practice timing — spacing — enhances skill acquisition, we failed to demonstrate that our measures of other kinds of practice variability can be related to enhanced skill acquisition. This leaves open the possibility that the exploration which supports skill acquisition is not captured by our measures, or that more complex skills, such as Destiny playing, require an equal match of habitual practice and exploratory variability.

References

Algoet, P. H., & Cover, T. M. (1988). A sandwich proof of the shannon–mcmillan–breiman theorem. *Annals of Probability*, 899-909.

Cepeda, N. J., Vul, E., Rohrer, D., Wixted, J. T., & Pashler, H. (2008). Spacing effects in learning a temporal ridgeline of optimal retention. *Psychological science*, *19*(11), 1095–1102.

Charness, M. K. R. R. E. V. E., N.; Tuffiash. (2005). The role of deliberate practice in chess expertise. *Applied Cognitive Psychology*, 151-165.

Delaney, P. F., Verkoeijen, P. P., & Spirgel, A. (2010). Spacing and testing effects: A deeply critical, lengthy, and at times discursive review of the literature. *Psychology of learning and motivation*, *53*, 63–147.

Drachen, A., Green, J., Gray, C., Harik, E., Lu, P., Sifa, R., & Klabjan, D. (2016). Guns and guardians: Comparative cluster analysis and behavioral profiling in destiny. In *Proc. of the ieee cig.*

Gallistel, C. R., Fairhurst, S., & Balsam, P. (2004). The learning curve: implications of a quantitative analysis. *Proceedings of the national academy of Sciences of the united States of america*, *101*(36), 13124–13131.

Goldstone, R. L., & Lupyan, G. (2016). Discovering psychological principles by mining naturally occurring data sets. *Topics in cognitive science*, *8*(3), 548–568.

Gray, W. D., & Lindstedt, J. K. (2016). Plateaus, dips, and leaps: Where to look for inventions and discoveries during skilled performance. *Cognitive Science*.

Herbrich, R., & Graepel, T. (2006). *Trueskill: A bayesian skill rating system*. Microsoft Research Tech. Rep. MSR-TR-2006-80.

Humphries, M. D., Khamassi, M., & Gurney, K. (2012). Dopaminergic control of the exploration-exploitation trade-off via the basal ganglia. *Frontiers in neuroscience*, *6*, 9.

Lessne, A. (2014). Shannon entropy: a rigorous notion at the crossroads between probability, information theory, dynamical systems and statistical physics. *Mathematical Structures in Computer Science*.

Magill, R. A., & Hall, K. G. (1990). A review of the contextual interference effect in motor skill acquisition. *Human movement science*, *9*(3), 241–289.

Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. In J. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1–55). Lawrence Erlbaum.

Newell, K. M., & McDonald, P. V. (1992). Searching for solutions to the coordination function: Learning as exploratory behavior. In G. E. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior* (Vol. 2, pp. 517–532). North-Holland.

Nunneley, S. (2016). Activision q1: Destiny has nearly 30m players; 55m monthly players for all ip. *VG24/7*. Retrieved from `www.vg247.com/2016/05/05/activision-q1-destiny-has-nearly-30m-players-online-player-community-hits-55m/`

Ranganathan, R., & Newell, K. M. (2010). Motor learning through induced variability at the task goal and execution redundancy levels. *Journal of motor behavior*, *42*(5), 307–316.

Rattinger, A., Wallner, G., Drachen, A., Pirker, J., & Sifa, R. (2016). Integrating and inspecting combined behavioral profiling and social network models in destiny. In *Springer lncs vol. 9926* (p. 77-89).

Ritter, F. E., & Schooler, L. J. (2001). The learning curve. In N. Smelser & P. Baltes (Eds.), *International encyclopedia of the social & behavioral sciences* (pp. 8602–8605). New York: Elsevier.

Rosenbaum, D. A., Carlson, R. A., & Gilmore, R. O. (2001). Acquisition of intellectual and perceptual-motor skills. *Annual review of psychology*, *52*(1), 453–470.

Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological review*, *82*(4), 225–260.

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.

Stafford, T., & Haasnoot, E. (2017). Confirming and quantifying sleep consolidation in skill learning: a field study using an online game. *Topics in Cognitive Science*.

Stafford, T., Thirkettle, M., Walton, T., Vautrelle, N., Hetherington, L., Port, M., ... Redgrave, P. (2012). A novel task for the investigation of action acquisition. *PloS one*, *7*(6), e37749.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge University Press.

Tammasia, M., Raffe, W., Sifa, R., Drachen, A., Zambetta, F., & Hitchens, M. (2016). Predicting player churn in destiny: A hidden markov models approach to predicting player departure in a major online game. In *Proc. of the ieee cig.*

Thompson, J. J., Blair, M. R., Chen, L., & Henrey, A. J. (2013). Video game telemetry as a critical tool in the study of complex skill learning. *PloS one*, *8*(9), e75129.

Thompson, J. J., McColeman, C. M., Stepanova, E. R., & Blair, M. R. (2017). Using video game telemetry data to research motor chunking, action latencies, and complex cognitive-motor skill learning. *Topics in Cognitive Science*.

Van Rossum, J. H. (1990). Schmidt's schema theory: The empirical base of the variability of practice hypothesis: A critical analysis. *Human Movement Science*, *9*(3), 387–435.

Acknowledgements