Testing theories of skill learning using a very large sample of online game players

Tom Stafford (t.stafford@sheffield.ac.uk)
Department of Psychology, University of Sheffield
Western Bank, Sheffield, S10 2TP, UK

Michael Dewar (michael.dewar@nytimes.com)
The New York Times R&D Lab, 620 Eighth Avenue
New York, NY 10018, USA

Abstract

We analyse data from a very large ($n = 854064$) sample of players of an online game involving rapid perception, decision-making and motor responding. This data set allows us to connect full details of training history with measures of performance, for participants who are engaged for a sustained amount of time in effortful practice. We show that lawful relations exist between practice amount and subsequent performance, and between practice spacing and subsequent performance. This confirms results long established in the literature on skill acquisition. Additionally, we show that higher initial variation in performance is linked to subsequent higher performance, a result we link to the exploration-exploitation trade-off from the computational framework of reinforcement learning. We discuss the benefits and opportunities of behavioural datasets with very large sample sizes and suggest that this approach could be particularly fecund for studies of skill acquisition.

Keywords: skill acquisition; learning; game.

Introduction

The investigation of skill learning suffers from a dilemma. One horn of the dilemma is this: experts in real-world skills can be brought into the lab and their performance tested, but it is difficult to reliably recover comprehensive details of their training. This makes it impossible to be certain of exactly how features of the history of their practice are related to the skilled performance you can observe. The other horn of the dilemma is this: you can test different training regimes rigorously, but you are restricted to measuring performance on trivial or unnatural skills, and often without extended training of the order that experts in complex real-world skills engage in.

Computer games offer a partial resolution to this dilemma. Even simple computer games are not trivial in terms of the cognitive abilities which they test. In fact, these abilities are often the staples of cognitive science: perception, decision making and motor responses. Computer game playing is a real-world skill in which many people choose to become expert, devoting hundreds of hours of practice. Unlike most skills, computer games allow a potential record every action in the history of that practice — allowing for the first time detailed investigation of the connection between features of practice and level of final performance. This is what the current investigation sets out to do. We take detailed records of practice activity from an online game and relate amount of practice and features of practice to levels of eventual performance. In doing this we are able to confirm and quantify established findings from experimental studies of learning. In addition we provide a confirmation of a recent result based on the theoretical framework of reinforcement learning (Stafford et al., 2012). Use of online games to collect very large samples offers a new method for the investigation of skill acquisition, we argue, and the work here showcases just some of the possibilities opened up by this approach.

Practice amount and spacing

We first consider two well established results against which we will validate our data set as a model of skill acquisition: the effects of practice amount and of practice spacing on performance. Studies of learning have shown a lawful relation between practice amount and performance. If performance is gauged in terms of some measure of efficiency (e.g. time taken to make cigars by experienced cigar manufacturers Crossman, 1959), then it is possible to express the relation between practice extent and performance in a power law of learning (Newell & Rosenbloom, 1981; Ritter & Schooler, 2001).

For practical reasons studies of the effect of extensive practice have typically looked at different learners possessing differing amounts of practice rather than the same learners at different stages (i.e. cross-sectional rather than longitudinal designs). Experimental studies of learning which do follow learners longitudinally have predominantly focussed on lab-based tasks which can be mastered in one or a small number of sessions (although there are, of course, honourable exceptions such as the work looking at the automatisation of visual search performance (e.g. Neisser, Novick, & Lazar, 1963; Czerwinski, Lightfoot, & Shiffrin, 1992).

Highlighting the importance of practice quantity in skill development, Ericsson and colleagues stress that the highest levels of performance are never reached without an amount of practice on the order of ten thousand hours (Ericsson, 2006; Ericsson, Krampe, & Tesch-Rmer, 1993). Additionally, they report that the nature of that practice matters — effortful, directed, 'deliberate' practice is what distinguishes elite performers, even among those who appear to have performed similar quantities of practice.

Experimental studies of learning have focussed on another factor which defines the nature of practice — spacing. The distributed practice effect denotes the finding that if time devoted to practice is separated out rather than massed, or if the spacing is larger rather than smaller, retention improves

(Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006; Delaney, Verkoeijen, & Spirgel, 2010). The distributed practice effect is surely one of the most solid findings in learning and memory research. It holds for both motor skill and declarative learning (Adams, 1987). Due to the limitations of experimental methods there is a dearth of evidence on longer spacing intervals (Cepeda et al., 2006), a dearth which we hope the present study offers a method of addressing.

Next we review an area where the approach adopted in this paper affords particular traction for looking at how the history of skill acquisition affects performance.

Exploration versus exploitation

The computational framework of reinforcement learning (Sutton & Barto, 1998), outlines a fundamental trade off in decision making: every decision forces us to choose between taking the action which we estimate will yield the best long term consequence (highest 'value'), or trying out an action of unknown or less certain value. This is known as the 'exploration—exploitation dilemma'. Every choice is an opportunity to receive the outcome from only one action, and so also to update our estimate of the value of only one option. Too much exploitation leads an agent to rely on suboptimal actions, seldom discovering better valued actions. Too much exploration, on the other hand, leads to an agent wasting time exploring the space of actions without garnering the reward of frequently choosing the highest known-valued action. The implications for skill learning are that non-maximising performance during early practice may allow superior subsequent performance. Indeed we might even expect that 'expert learners' would adopt an early exploration strategy in order to maximise final performance.

We have already found evidence for this in humans and rats using an experimental task (Stafford et al., 2012). There is other evidence that variability in practice conditions can aid final performance (Roller, Cohen, Kimball, & Bloomberg, 2001), as well as generating benefits in learning which cross-task (Seidler, 2004) (which has been termed 'structural learning' by some). This is somewhat in tension with accounts which emphasise the need for transfer-specificity in skilled performance (e.g. Logan, 1988). There is not a direct contradiction, merely we are emphasising the benefit of training off the to-be-tested skill.

Method

Game designers Preloaded produced a game for the Wellcome Trust called 'Axon', which can be played here `http://axon.wellcomeapps.com/`. They inserted tracking code which recorded a machine identity each time the game was loaded and kept track of the score and date and time of play. The game was played over 3.5 million times in the first few months of release (Batho, 2012).

The game involved guiding a neuron from connection to connection, through rapid mouse clicks on potential targets. A screenshot can be seen in Figure 1 (see figure caption for description of game dynamics). Cognitively the game involved little strategic planning, testing rapid perceptual decision making and motor responses.

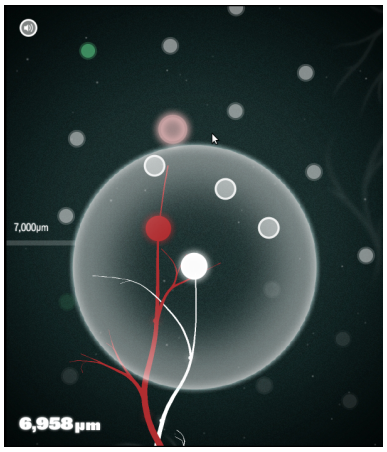

Figure 1: Screenshot of the game Axon. Players control the axonal branching of the white neuron. At each point, possible synaptic contacts (the other dots) are those within the zone of expansion (the larger transparent circle), which shrinks rapidly after each new contact is made. Non-player neurons (in red here) compete for these synaptic opportunities. Score is total branch length in micrometers (shown bottom left).

The analysis was approved by the University of Sheffield, Department of Psychology Ethics Sub- Committee, and carried out in accordance with the University and British Psychological Society (BPS) ethics guidelines. The data was collected incidentally and so did not require any change the behaviour of game players, nor impact on their experience. No information on the players, beyond their game scores, was collected and so the data set was effectively anonymised at the point of collection. For these reasons the institutional review board waived the need for written informed consent from the participants.

Because the data we record is indexed by machine identity, which is derived from the web browser used to access the game, it is not possible to guarantee that a single individual is responsible for all the scores recorded against an single identity. Nor is it possible to guarantee that a single individual is responsible for only one set of scores. These uncertainties add noise to our analysis, but the data set is large enough to accommodate this. It is not clear what, if any systematic distortions these caveats would introduce. For the remainder of this paper we will use the term 'player(s)' to refer to the set of scores associated with a single machine identity.

The data was extracted from Google Analytics using a Python library by Clint Ecker (2009). Data from between 14th of March and 13th of May 2012 was downloaded and compiled into the source data set for the analyses presented here. This data set comprised a total number of 854064 players. Most played only a small num-

ber of times (the modal number of plays is 1), but some played up to 1000 times. The data and code for producing the analysis and plots presented here are available from `https://github.com/tomstafford/axongame`.

Results

Practice amount

On average, scores are higher with each consecutive play for up to 100 plays (Figure 2). At around 80 plays the levels of variation between scores, combined with the drop off of number of players reaching that number of attempts, begin to be seen in the loss of the smooth curve and larger error bars.

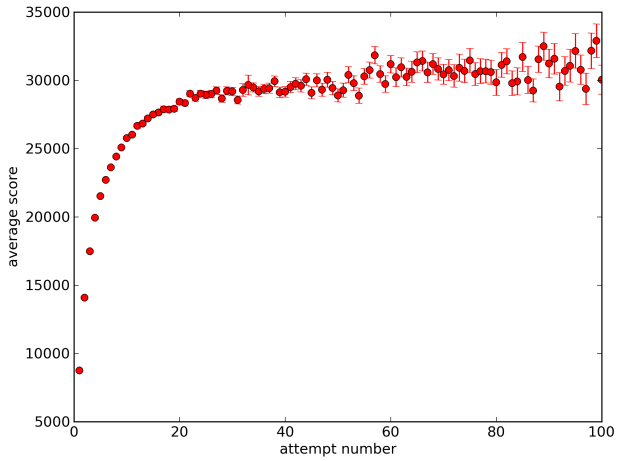

Figure 2: Average score for each play attempt. Standard errors shown (n.b. some error bars not visible at this scale).

Taking only those who played more than 9 times ($n = 45672$), we can calculate a 'high score' for each players (i.e. the highest score they achieved, irrespective of which play it occurred on). The criterion of 9 or more plays for subset selection is arbitrary, an attempt to balance size of subset (which drops with a higher criterion) against likelihood that practice effects will be reliable (which should be greater for higher criterion values). For this, and all other analyses presented in this paper, the results are not contingent on the particular values used to divide up the data (i.e. here we get similar results if greater than 8, 10, 5 or 20 plays are used as the criterion. To confirm this we invite interested readers to run the analysis with altered parameters themselves, by visiting the data and analysis code repository referenced above).

From this subset players are then grouped into 5 groups based on the percentile ranking of their high score, and the average score is calculated for each attempt for all players in each percentile group. This shows that the difference between higher and lower scorers is not merely the amount of practice. The difference in average score is present from the very first plays (Figure 3).

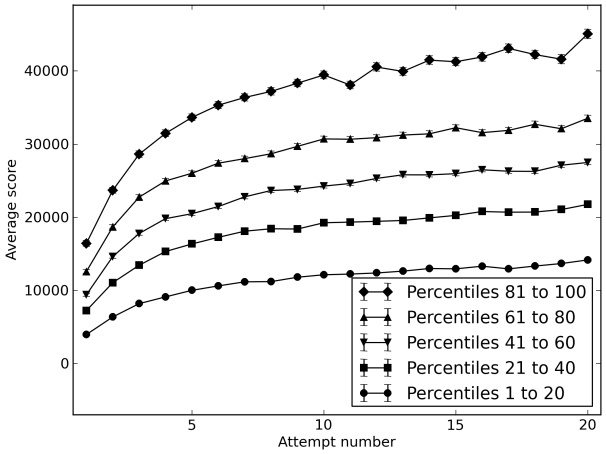

Figure 3: Average score against attempt number for different groupings according to maximum score. Standard errors shown.

Practice spacing

Taking only those who played more than nine times, we divide players into percentile groups according to their highest score, regardless of on which play it was obtained. We also calculate the separation in time between their first and last play. The result shows a clear upward trend (Figure 4, red dots), with players who score most highly spreading their first and last plays further apart. This is unsurprising, however, since even if there was no relation between practice and scoring, and scores were simply random on each attempt, those players who played had more attempts would tend to collect higher scores and have first and last attempts which were more separated in time. We use bootstrapping to estimate confidence intervals as if this were the case. Keeping the number of players and the number and time of the attempts constant, we generate 2000 simulated datasets, sampling with replacement at random from the total record of all scores for all players. The observed data falls below this bootstrap data for low maximum score percentiles and above for high maximum score percentiles, suggesting that the scores really are distributed non-randomly and according to the spread in time of participant's plays (Figure 4).

It is possible to interrogate this result further by a finer slicing of the data. Taking only players who played more than 14 times (n=21575), we calculate the spread in time between the first play (or second play where this data was missing) and their tenth play (or ninth, where this data was missing). We also identify their best score on plays 11 to 15. We then divide them into two groups, those who played their first ten times within a 24 hour period ("goers"), and those who split their first ten plays over more than 24 hours ("resters"). Resting between first and tenth plays appears to have a benefit on your subsequent performance (Figure

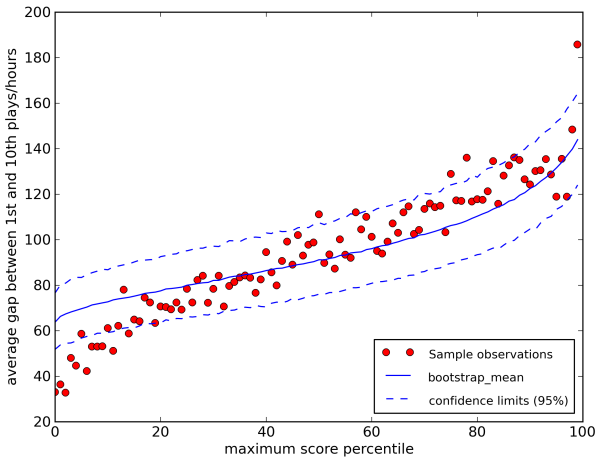

Figure 4: Players graded according to their maximum score percentile against the delay between their first and last plays. Standard errors shown.

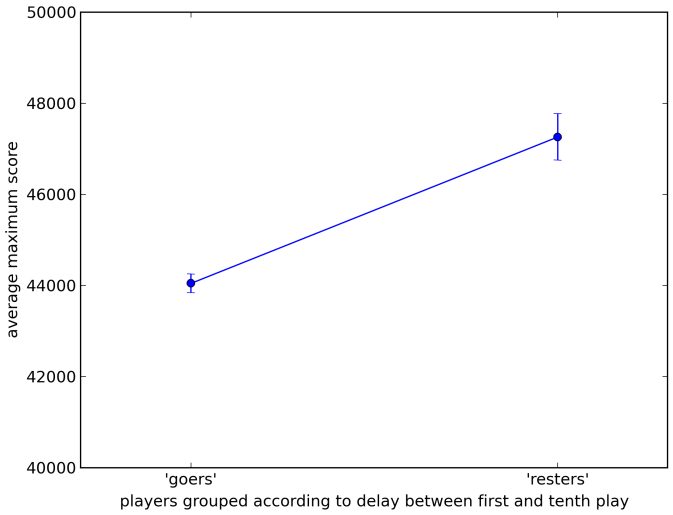

Figure 5: Average maximum score following first ten plays, for those who group their first ten plays within one day ('goers') and for those who split their first ten plays over two or more days ('resters'). Standard errors shown.

5). The difference between the groups is highly significant ($t(20354) = 6.219, p < 0.00001$), albeit for a small effect size (Cohen's d = 0.11).

Exploration versus exploitation

The variance of scores for each player in the first five plays was calculated, and this statistic for each player ranked and so percentile groups created. The same was done for the average on plays six to ten. Plotting one against the other we see a clear correlation - with higher early variance associated with higher subsequent performance (Figure 6, the very high number of individuals made a scatterplot impractical at this scale, so we present a heatmap).

The Pearsons's r correlation coefficient was 0.59 and significantly different from zero at a high probability ($p < 0.0001$). Randomising the scores for each attempt within the structure of the number of players and the number of attempts per players, it is possible to generate a bootstrap data set which gives a confidence interval for this correlation - in other words, answers the question "to what extent is a correlation between high early variance and high late scoring inherent in the distribution of scores and the structure of how players accumulate scores from that overall distribution". These bootstrapped confidence intervals, at the 95% level were 0.009 to -0.009. Thus we can conclude with a high degree of confidence that the correlation is both significantly different from zero and not a trivial consequence of the distribution of scores. Instead, the correlation results from the particular way individual player's early scores are related to their later scores.

Discussion

These results confirm, but also quantify, results from experimental psychology regarding the effects of practice quantity and quality on performance. As players practice their average score improves. Dividing the players into percentile groups according to high scores appears to show that practice alone does not allow most players to achieve the highest scores. The best players have an advantage from the very first plays. This advantage is consolidated with practice, in that not only do they score more on their first plays, but their rate of improvement is faster. This is in marked contrast to some popular (e.g. Gladwell, 2008) and academic (e.g. Ericsson et al., 1993) accounts of high performance which have denigrated the importance of talent with respect to practice. We regard this result as provisional. It needs to be replicated with another data set so we can assess if it generalises to other skills. Replication would also assuage worries that some specific confound of the present data set has produced the result. For example, we have no way of controlling for the prior game experience or hardware set-up of the players of the Axon game. It is possible that it a certain amount game experience is required for individuals to get high learning rates with this specific game (we thank an anonymous reviewer for pointing out this potential 'thresholding of performance improvement by prior experience' confound).

The analysis of practice spacing confirms the wisdom from experimental studies of learning and memory that distributed practice is better than massed practice. It remains to be seen if there is an optimal amount of spacing, as has been reported for semantic knowledge (Cepeda, Vul, Rohrer, Wixted, & Pashler, 2008), or an optimal timing of spacing (Goedert & Miller, 2008).

The exploration-exploitation result confirms a preliminary result from a recent experimental study (Stafford et al., 2012). Although bootstrapping confirms that this finding is

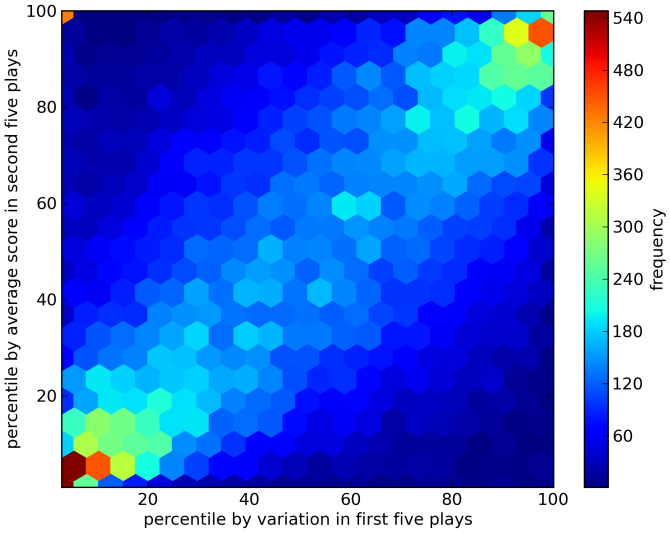

Figure 6: Heatmap made from scatterplot of variance of scores on first five plays versus average score on plays six to ten.

not an incidental result of the distribution of scores, it still isn't clear if the level of exploration (operationalised as score variance on early plays) *per se* causes the higher level of performance ('exploitation', characterised as score average on later plays). It is doubtful that low scoring attempts in themselves cause higher subsequent performance. Rather low scores may be the impetus for players to shift their playing style or tactics in ways which unlock higher subsequent performance (similar to the postulated freeing and freezing of degrees of freedom which have been thought to characterise changes in motor skill (Berthouze & Lungarella, 2004; Bernstein, 1967). The ultimate test exactly if and how early exploration affects subsequent performance will be to intervene to make players explore and see how this affects later scores. In other domains there have been suggestions that introducing guided mistakes or deliberate failure into early training may have benefits for overall performance (something for which there is some evidence: Lorenzet, Salas, & Tannenbaum, 2005).

Games

Games are a great opportunity for the cognitive science of learning. They provide participants in high numbers who are engaged willing to undertake extensive practice. Games can provide large amounts of detail on training conditions and actions, in ways that other paradigms cannot. In the future it may even be possible to introduce experimental manipulations into engaging games through partnership with games designers.

'Big Data'

The method of study adopted here means we lose experimental control over the factors involved in learning. However, advantages stem from the very large sample size we are able to collect. Some of the emphasis on the importance of experimental control in cognitive science is due to the loss of statistic power than can result from uncontrolled measurement. With large sample sizes, loss of statistical power is not an issue. We need only concern ourselves with the ways in which lack of experimental control introduces systematic confounds into our data set. As well as large statistical power, very large sample sizes mean we can interrogate data in new ways. One of these is 'slicing' by which we mean identifying individuals who meet certain conditions and comparing within that group. This is a substitute for the conventional experimental method of creating individuals that meet certain conditions in low numbers. In experimental design you control potential confounds in advance (by attempting to remove them). With slicing you attempt to account for potential confounds *post hoc* by selecting multiple different sub-datasets, each of which controls statistically for a potential confound - and thus by a process of elimination gathering support for your hypothesised causal variables. This is a less powerful method than experimental control, but it does offer some advantages.

Bootstrapping provides a way of testing observed patterns against sophisticated null hypotheses. Both bootstrapping and slicing are illustrated in this paper in the analysis of spacing effects.

Two modern crises of psychology are the apparent low replicability of effects (Pashler & Wagenmakers, 2012) and the use of inappropriate statistics (Wagenmakers, Wetzels, Borsboom, & Maas, 2011; Simmons, Nelson, & Simonsohn, 2011). Very large sample sizes can side-step both of these. With a large enough sample size you do not need to use inappropriate statistical techniques - small effects are easy to find. Furthermore, you have enough data to use techniques such as cross validation to guard against false-positives.

Analysed in detail, very large data sets provide an observational playground in which we can not just detect effects, but compare the size of different effects against each other. For example, in the present data set it can be seen that the benefit of 24 hours spacing is about 3000 points (Figure 5). This is comparable to about 5 plays, in the 10-15 play range (Figure 2), or equivalent to an extra 50% practice at this stage of experience.

Obviously, nothing will replace the controlled experiment in terms of causal inference. For hypothesis testing the controlled experiment must remain the the gold-standard. However, there is space within the scope of investigation for studies with purposes other than theory-driven hypothesis testing (Rozin, 2009). This paper has focussed on characterising the data and confirming effects discovered in traditional controlled experiments. We believe the approach illustrated here can be complementary to experimental studies, and has the potential to open up new avenues for investigation in the study of skill acquisition.

Acknowledgements

Thanks to Tony Barnes for an introduction, to Phil Stuart, Charles Batho and Cameron Yule at Preloaded, to the Wellcome Trust for allowing the data from their game to passed to us, to Stuart Wilson for help with python, Ashvin Shah for discussion of reinforcement learning and four anonymous reviewers. Special thanks to Edith Mary Cameron, whose late arrival and post-birth disposition allowed TS to carry out the bulk of the analysis and writing for this paper.

References

Adams, J. A. (1987). Historical review and appraisal of research on the learning, retention, and transfer of human motor skills. *Psychological Bulletin*, *101*(1), 41–74.

Batho, C. (2012). *Axon — a game for science.* http://preloaded.com/blog/2012/07/05/axon-game-science/.

Bernstein, N. A. (1967). *The co-ordination and regulation of movements*. Oxford: Pergammon.

Berthouze, L., & Lungarella, M. (2004). Motor skill acquisition under environmental perturbations: On the necessity of alternate freezing and freeing of degrees of freedom. *Adaptive Behavior*, *12*(1), 47–64.

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological bulletin*, *132*(3), 354–380.

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.

Crossman, E. (1959). A theory of the acquisition of speed-skill. *Ergonomics*, *2*(2), 153-166.

Czerwinski, M., Lightfoot, N., & Shiffrin, R. M. (1992, July). Automatization and training in visual search. *The American Journal of Psychology*, *105*(2), 271–315.

Delaney, P. F., Verkoeijen, P. P. J. L., & 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.

Ecker, C. (2009). *Python library for Google Analytics API.* https://github.com/clintecker/python-googleanalytics Accessed 17 May 2012.

Ericsson, K. A. (2006). The influence of experience and deliberate practice on the development of superior expert performance. In N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The cambridge handbook of expertise and expert performance* (p. 683-703). Cambridge University Press.

Ericsson, K. A., Krampe, R. T., & Tesch-Rmer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*(3), 363–406.

Gladwell, M. (2008). *Outliers: The story of success*. Little, Brown.

Goedert, K. M., & Miller, J. (2008). Spacing practice sessions across days earlier rather than later in training improves performance of a visuomotor skill. *Experimental Brain Research*, *189*(2), 189197.

Logan, G. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*(4), 492–527.

Lorenzet, S. J., Salas, E., & Tannenbaum, S. I. (2005). Benefiting from mistakes: The impact of guided errors on learning, performance, and self-efficacy. *Human Resource Development Quarterly*, *16*(3), 301–322.

Neisser, U., Novick, R., & Lazar, R. (1963). Searching for ten targets simultaneously. *Perceptual and Motor Skills*, *17*(3), 955–961.

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

Pashler, H., & Wagenmakers, E.-J. (2012, November). Editors introduction to the special section on replicability in psychological science a crisis of confidence? *Perspectives on Psychological Science*, *7*(6), 528–530.

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

Roller, C., Cohen, H., Kimball, K., & Bloomberg, J. (2001). Variable practice with lenses improves visuo-motor plasticity. *Cognitive brain research*, *12*(2), 341–352.

Rozin, P. (2009). What kind of empirical research should we publish, fund, and reward?: A different perspective. *Perspectives on Psychological Science*, *4*(4), 435–439.

Seidler, R. (2004). Multiple motor learning experiences enhance motor adaptability. *Journal of Cognitive Neuroscience*, *16*(1), 65–73.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*(11), 1359–1366.

Stafford, T., Thirkettle, M., Walton, T., Vautrelle, N., Hetherington, L., Port, M., et al. (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.

Wagenmakers, E.-J., Wetzels, R., Borsboom, D., & Maas, H. L. J. van der. (2011). Why psychologists must change the way they analyze their data: The case of psi: Comment on Bem (2011). *Journal of Personality and Social Psychology*, *100*, 426–432.