Press Pause when you Play: Comparing Spaced Practice Intervals for Skill Development in Games

Colby Johanson, Carl Gutwin, Jason T. Bowey, and Regan L. Mandryk
University of Saskatchewan
Saskatoon, Canada
{firstname.lastname}@usask.ca

ABSTRACT

Games allow players to fulfill the need for competence by providing well-designed, increasingly difficult challenges. To meet these challenges, players repeatedly attempt to achieve objectives—and through this repetition, they improve their game skills. Players are keenly aware of whether they are making progress during these attempts, and they want to get better as quickly as possible. Previous research suggests that one way of improving skill development is by taking breaks between periods of activity (called "spaced practice"). However, there is little knowledge about whether this idea works in games, what the optimal break length is, and whether the effects last. We carried out a study comparing spaced and continuous practice in a *Super Hexagon* clone, using five-minute play intervals and five break lengths (no break, two minutes, five minutes, ten minutes, one day). We found that spaced practice led to significant gains in performance, particularly for novices. This result shows that players can achieve an immediate improvement in skill development, simply by scheduling short breaks in their play session; designers can also make use of this result by building rest periods into the structure of their games. Our study also indicated that breaks are valuable both in the short and the longer term—in a retention test after one day, all of the groups performed similarly, suggesting that even if a player does not use spaced practice initially, taking a break after the play session can still lead to improvements. Our study provides new information that can aid in the design of practice schedules for perceptual-motor tasks in games.

CCS Concepts

•**Human-centered computing** → **Empirical studies in HCI;** *HCI theory, concepts and models;* User studies; •**Applied computing** → *Computer games;*

Author Keywords

Skill Development; Spacing; Practice; Games

CHI PLAY '19, October 22–25, 2019, Barcelona, Spain.

ACM ISBN 978-1-4503-6688-5/19/10 ...$15.00.
http://dx.doi.org/10.1145/3311350.3347195

INTRODUCTION

Repetition of a task is an essential part of many digital games. Repetition in games may inspire memories of classic arcade games on which players willingly spent coin after coin, all for the chance to beat their high score. However, even in modern games, players show a desire to repeat—and improve their performance on—in-game tasks [90, 92]. Consider, for example, the success of games like *Super Meat Boy* [110] or *Cuphead* [109], which require players to repeat, and improve, sequences of game actions; or the appeal of "speedruns", where players practice a level over and over to set a time record.

In most games, these repeated attempts are not just thrashing; instead, players get better with practice. When a player attempts a difficult challenge and fails, they can apply that experience to future attempts [57]. Repetition therefore serves as *practice* that can lead to the development of in-game skills—a desirable outcome for the player. Furthermore, because games are typically designed to become more difficult as the player makes progress [90, 125], players must continuously develop their skills to keep progressing. In-game skill development is a key reason why gamers are intrinsically motivated to play, because developing and improving skills satisfies their need to experience competence [90, 62, 87, 66]. In addition, many games also have mechanisms that provide extrinsic motivation to improve one's performance, such as in-game rewards for success [125], out-of game achievements [23], or leaderboards that allow players to compare themselves to their peers [10]. Because players are so interested in improving their skills, and in doing so as quickly as possible, it is important that we better understand how in-game skills develop and how this development can be accelerated.

The skills that are developed in many games are perceptual-motor skills (requiring coordination of physical actions such as button presses with on-screen events), and in this domain, several strategies based on theories of skill development have been shown to improve performance—for example, in aiming [35, 6], pursuit tracking [2, 9], and mirror tracing [106]. However, although game developers have made use of ideas about skill development in other areas (e.g., introducing mechanics one at time [43, 56], providing clear feedback [112, 55], and allowing players to immediately practice skills they have just learned [125, 49]), there has been little application of theories of motor learning in games.

One theory of skill development that has been widely studied is *spaced practice*—the idea of scheduling periods of rest to break up periods of activity. Theory suggests that these rest breaks allow the brain to generalize and compile the feedback that has been gathered, leading to improvement once the rest is finished [97, 19]. Studies have shown spaced practice to be effective across many tasks, when compared to continuous practice (i.e., no rests) [98, 20]. However, there have been very few investigations of spaced practice in digital games—and in general, research into the theory has been criticized for rarely testing it outside of laboratory settings [16, 21]. Furthermore, the motivational pull of games changes the idea of "practice" in that it is not the chore that it is in other contexts where this effect has been found, and players may want to keep playing without a break.

This is evident by the strong reactions of players at the mere suggestion of taking a break. Many *Wii Sports* [81] players mock the game's suggestion that they should take a break after playing for an hour [115]. Even games that take a more implicit approach to breaks spark debate in the gaming community. For example, players who die fighting a boss in *Dark Souls* [37] are made to traverse the game world to make another attempt and there are players who question this design choice. These players would rather attempt the boss again immediately (e.g., [116, 117, 13]), and they describe the walk back as simply a waste of time; others argue that the relatively easy task of walking back to the boss gives them an opportunity to improve by reconsidering their strategy [116]. This debate shows that a typical behaviour for many people is to play continuously with no breaks—and if pausing could be shown to actually help these players overcome challenges, it could have a broad effect on how people play.

To better understand how spacing practice affects skill development in digital games, we carried out a study using a clone of a commercial game (*Super Hexagon* [14]) that is almost entirely based on perceptual-motor skills. We divided a total play time of 25 minutes into five sessions: four training sessions and one retention session. Participants received different amounts of rest between training sessions: three seconds (i.e., continuous practice), two minutes, five minutes, ten minutes, or one day. The retention session took place one day after the participant's last training session. Our study was designed to answer three questions:

- Are there benefits of spaced practice on performance in digital games?
- Do performance benefits last beyond the training session?
- Is there an optimal rest period for this game?

Our results showed that there were clear and *immediate* benefits of spaced practice. There were significant differences in the rate of improvement in the game, with the continuous-practice group performing worse during training than any of the spaced-practice groups. These results suggest that scheduling short rest breaks can provide an immediate benefit in skill development, particularly for novice players (who may be the group most interested in accelerating their skills). In addition, we found that a longer break (a one-day rest between training and the retention test) was also effective in improving performance, even for the group that played continuously during training – by time of the retention test, all of the groups performed similarly. Among our rest intervals, there was no one interval that worked better than any other. The ten-minute group improved more quickly early on, but the differences between groups tapered off in later sessions, suggesting that the effects of spaced practice could interact with player experience.

Our results show that designers can use spacing to help players improve in the moment—an important factor in perceptions of competence and player experience—and also show that longer breaks can assist players regardless of practice schedule. Our study provides designers with valuable new understanding of how skill development works in games, and suggests new strategies for helping gamers improve as they play.

RELATED WORK

Spaced Practice

Spaced practice means scheduling periods of rest to break up periods of work within a training session [97]. This approach has been shown to improve performance during training, and to improve retention as well [64, 25, 97]. There is no fixed timing for the rest periods relative to work periods [118], and any amount of rest compared to a continuous-practice condition is typically considered spaced practice [97].

The effect of spaced practice has a long history, going back to Ebbinghaus's work on learning lists of nonsense syllables [28] and Snoddy's work on mirror tracing [106]. Some of this past work differentiated between performance during training (temporary performance) and retention after training (permanent performance) [2, 9]. In some studies, improvements due to spaced practice affected only temporary performance (because continuous practice conditions could be affected by increased fatigue, boredom, or failure to sustain attention [118]). Retention tests are therefore seen as an important part of assessing the value of different practice schedules—and studies have shown positive effects of spaced practice on retention after a delay of a day or longer [2, 5].

Meta-reviews of spaced practice studies have showed strong overall effects for the technique [64, 25]. For example, Lee and Genovese found a large mean weighted effect size of 0.96 for training (temporary performance), and a medium effect size of 0.53 for retention [64]. However, Verhoeven and Newell suggest that the meta-reviews do not necessarily provide unequivocal support for the idea that spacing practice enhances learning compared to continuous practice, as there are aspects of practice that moderate the effectiveness of spacing (e.g., differences in the task or the learner) [118]. Additionally, there is little agreement as to what length of rest optimizes the effect [118]—some suggest that longer breaks are more effective than shorter breaks [97], and others suggest that performance follows an inverted U function [16, 25].

Task Factors

Spaced practice has been shown to work for several different tasks, but there are also examples of spaced practice not producing improvements. For example, more complex tasks such as math problems [77] and learning a musical sequence on the piano [127] did not show a benefit for spaced practice. In general, spaced practice appears to be more effective for simple perceptual-motor tasks than for complex tasks [25]. Additionally, Lee [65] found that discrete skills with a definite beginning and end did not benefit from spaced practice.

Individual Differences

In addition to the complexity of the task being performed, the individual's skill level in the task has also been shown to influence the effectiveness of spaced practice. As early as 1926 it was thought that spacing was most effective in the early stages of learning [106], and other work has also shown that spacing is more important in early stages, with continuous practice being better in later stages [59]. Recent work has also showed that an individual's performance after a break was better predicted by the stage of skill development rather than the break itself [111].

Spaced Practice in Digital Games

The amount of research investigating spaced practice in digital games is limited, with the majority focusing on serious games in the context of education and verbal learning (e.g., [29, 95, 96]) rather than on in-game perceptual-motor skills. When limiting our search to this latter topic, we found only two relevant experiments: a 1985 experiment by Metalis [75], and a 1999 experiment by Shebilske et al. [102].

The 1985 experiment [75] used the Apple II+ game *Little Brick-Out*—a *Pong*-inspired single-player game in which the player controls a paddle to hit on-screen bricks with a bouncing ball. The experiment consisted of 10 rounds of the game, with either a 2-minute break or no break between rounds. Each round lasted anywhere from 57 seconds to 232 seconds, depending on how well the participant performed. It was found that spacing practice resulted in better performance at the 10th training sessions than continuous practice. No retention test was performed. However, this study has a number of issues. The participant group in the spaced practice condition had higher initial performance than the continuous group, and because the game's rounds lasted longer if the player performed well, this meant that the spaced-practice group actually trained with the game for far longer than the continuous-practice group (27 minutes compared to 15 minutes). In addition, the 18 minutes of rest made up a large amount of the total training time.

The 1999 experiment [102] looked at spaced practice in the context of a more complex game called Space Fortress [24]. Participants played for much longer (10 hours over either 2 days or 10 days), and also performed a retention test one week after the last day of training. The group that trained over 10 days outperformed the two-day group both at the end of training and on the retention test. The game used in this study involved complex tasks with strong strategic components, and so did not focus primarily on perceptual-motor tasks. In addition, the experiment included substantial coaching of the participants (e.g., participants watched videotaped instructions, were encouraged to try different strategies, were encouraged to try their best, and were instructed on strategies that other players had found effective). Even with these two studies, there is still little information about the effects of spaced practice on motor skill development in digital games.

Esports: Skill improvement as a profession

Whereas many players repeat game objectives simply to make progress in an entertaining game, the recent rise of esports shows that skill improvement and practice effects have now become a job as well as a leisure activity. Several definitions of esports focus on the importance of training (e.g., "sport activities in which people develop and train mental or physical abilities in the use of information and communication technologies" [121]). The esports industry is now large: a 2019 market report estimated the esport economy to be at $1.1 billion with 26.7% year-on-year growth; further, the global esport audience is estimated to be 443 million people [80].

The need for skill development in esports is a clear example of why it is critical to better understand how techniques such as spaced practice affect training. Practice has long been an important part of traditional sports, and many parallels are found between athletes in esports and more traditional athletic pursuits, particularly in the areas of perceptual-motor expertise and the drive for improvement of those skills [101, 44]. Among these similarities we see common patterns involving training and practicing fundamental skills. Development of skills is crucial for esport players—but is also important for any player who wants to perform at a high level. Understanding how skills in a digital game are developed is important for creating effective training programs both for professionals and amateur players.

A FRAMEWORK OF SKILL DEVELOPMENT FOR GAMES

There are several issues that need to be considered when designing or evaluating a practice scheme for digital games. In the following sections we review game skill types, the stages of skill development, and how skill development is evaluated. This framework was used to inform the design of our game and our experiment; and although our research focuses primarily on perceptual-motor skill learning, this framework describes skill learning in general.

Game Skill Types

Games contain many different skills that players must master in order to succeed [83, 82, 54, 52, 119]. A skill refers to the ability to carry out a specific task to achieve a specific goal [33, 114, 97, 30]. Skills must be learned—learning refers to a relatively permanent change in behaviour that occurs as a result of practice, expertise, or experience [113, 124, 94, 114].

Skills can be broadly classified as either *cognitive* or *motor* [114]. Cognitive skills include problem solving, memory, language, and emotional skills, while motor skills include anything that requires body or limb movements in order to make a physical response [33, 114, 67]. Despite these separate classifications, the reality is that many tasks, including digital games [83], require components of both types of skills [114, 67]. Motor skills are usually referred to as *perceptual-motor* or *psychomotor* because of the importance of *perception* and *decision making* in the process [30, 46]. The learner must learn to process stimuli so as to recognize features that require a response [30, 124, 39, 89, 114]. In other words, the learner develops a stimulus-response *coding* that lets them quickly select correct responses [122, 89].

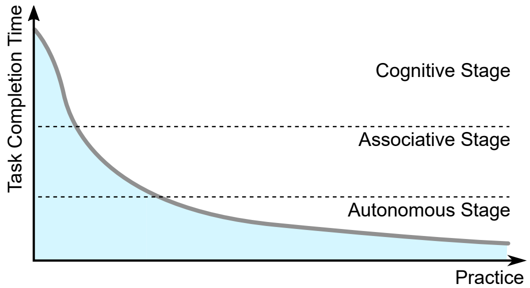

Figure 1: The three stages of learning approximately mapped to the power law of practice. Adapted from [59].

Games That Facilitate Skill Development

Games are considered to be effective learning environments [57, 40]. This is because they have clear goals with strong feedback [112], and are a task that players are willing to give their complete attention to [17]. These factors have been shown to facilitate continued skill development [32].

Feedback acts as a source of motivation for players [55, 84, 90]. It provides them with information regarding how well they are doing [112]—if they are doing well then it can enhance their needs satisfaction of competence, which will in turn make the player more motivated to keep playing [90]. Conversely, if the player is doing poorly, then a motivated player will leverage this information to modify their strategy or response [38, 90].

This feedback is most effective when players are presented with tasks just outside of their comfort zone [32, 57]. Without this, the player may adopt a strategy or response that is acceptable for the difficulty of the task, but is sub-optimal compared to other approaches or when the task becomes more difficult [99, 32]—a phenomenon known as "satisficing" [104]. Games use different approaches to achieve this. When a game's challenge primarily comes from its mechanics or design, the game can be implemented to consistently increase in difficulty as the player makes progress [125, 90]. If the game's challenge primarily comes from competing against other players, a matchmaking system can be applied so that similarly skilled players compete against one another [84].

The Stages of Skill Development

In digital games, players can continue to make performance improvements over dozens or hundreds of hours [8, 63, 48]. For some games (as with esports) the high skill ceiling makes it possible to play a game as a full-time job and continue to see performance improvements [4, 27]. Changes in performance are described primarily by two theories: the power law of practice [106, 79, 98, 114], and the three stages of skill development [98, 36, 34, 89, 88, 59, 58].

The acquisition of skill follows a predictable pattern, described by the *power law of practice* [106, 79, 98, 114] (shown in Figure 1). Simply put, skill improvement occurs as a function of the number of repetitions, with performance improving dramatically at the start, but eventually slowing down. In the case of digital games, special consideration should be given to the fact that many games are designed to increase in difficulty as the player also improves [125, 90].

The *cognitive* stage of skill development is characterized by initial poor performance with many errors, but also by rapid improvement [79, 98, 36, 34, 89, 88, 59]. Learners must give the task their full attention [18, 100] as they use their declarative (verbalizable) knowledge of the task [59] to learn how the task is performed [36, 97, 114]. As they attempt the task, they begin to refine their attempts based on any feedback that is provided [126, 93, 105, 61, 124, 114, 39]. At the end of the cognitive stage, learners start to form the stimulus-response codings [89, 47] and the procedural knowledge [59, 123] that they will use in higher stages; these structures are more robust to decay [59] and allow quicker response times and improved performance over declarative knowledge alone [70, 7].

The *associative* stage is characterized by a dramatic reduction in errors, greatly improved performance, and slower improvements [79, 36, 114, 98]. The learner's attention is on optimizing their performance [114, 97]. There are three main reasons for performance improvements in this stage: first, the learner develops a consistent response in the form of learned patterns [36]; second, their ability to identify relevant stimuli improves [114]; third, the learner no longer needs to use the declarative knowledge of the task [59] and can instead leverage procedural knowledge and "direct stimulus-response associations" [89].

The *autonomous* stage is characterized by few errors, stable expert performance, and little further improvement in the motor domain [97, 36, 79]. The learner performs the skill with coordination, smoothness, and accuracy [114]. The learner can respond to stimuli with automaticity [36, 114, 89, 100], which allows them to direct attention towards improvement through changes in high-level strategies [32]. While continued improvement at this stage is difficult, it can be accomplished by making use of deliberate practice [32] — a type of practice where the learner avoids acting with automaticity and focuses on specific aspects of the task that can be improved.

Evaluating Skill Development

Skill learning cannot be observed directly, and so is inferred by examining the learner's performance [98, 114, 46, 107]. However, the learner's performance during or immediately after training is often insufficient because transient factors such as fatigue may be at play during the training session [98, 107]. Longer-term learning must therefore be evaluated by having the learner perform an additional test with the experimental variable removed, either in a transfer test (where the learner completes a task different from the training task) or a retention test (where the learner completes the same task but after a delay) [98, 107]. It is important to note, however, that in the game domain, we are interested in both immediate improvement (i.e., as the player is repeatedly attempting the objective) as well as longer-term retention of a skill.

A DIGITAL GAME FOR THE STUDY OF PERCEPTUAL-MOTOR SKILL DEVELOPMENT

Given that a typical experiment lasts less than an hour and that games can be played for dozens or hundreds of hours, designing a game that can be used to evaluate skill development in the context of this tight time-frame poses some challenges. We wanted a game that players could begin playing immediately,

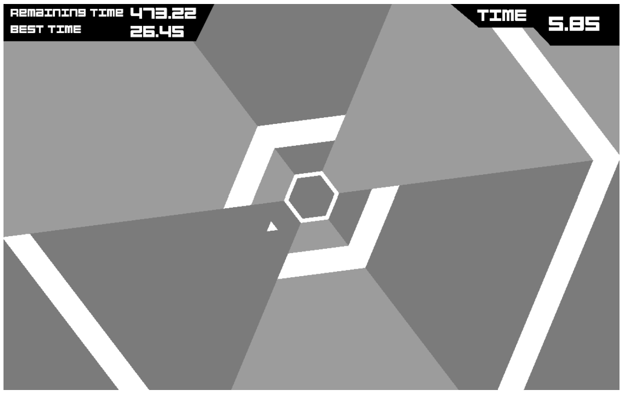

Figure 2: A screenshot of our *Super Hexagon* clone.

with little need for instructions, demonstrations, or guidance to introduce the game mechanics. The game should also present the player with clear goals, and only require minimal learning of declarative knowledge. In addition, players should be able to continuously make performance improvements over several play sessions, and should be motivated to continue to play and improve [24]; the game should be designed to avoid "at-game frustration" [76] and should provide clear feedback [62, 55]. Finally, performance in the game should be straightforward and easily quantifiable [24], and ideally not based on aggregated measures [8].

Super Hexagon

Based on these guidelines, we produced a clone of the game *Super Hexagon* using the Unity game engine. *Super Hexagon* is described by its creator as a "minimal action game" [15]. The player controls a triangle that can rotate radially around a central hexagon, using the left and right arrow keys. In each of the six regions on the screen, obstacles appear at the outside of the screen (white bars) and move inwards towards the center—requiring that the player rotate to avoid the obstacles (see Figure 2). The game's goal is simple—last as long as possible while avoiding the incoming obstacles. The camera rotates continuously to provide some perceptual confusion. As long as the player stays alive, the game gets progressively harder by increasing the camera rotation speed and the rate at which the obstacles spawn and move inwards. The difficulty is reset every time the player fails.

Super Hexagon fits our guidelines because it has minimal controls (only two keys) and a clear goal (avoid the obstacles), meaning that players can begin playing immediately. Performance improvements can continue for a long time (the average completion time for the entire game is around 15 hours [48]), and performance is easily measured as time until failure. Additionally, its commercial success [50, 108] and critical praise [74] suggest that many gamers find the game enjoyable.

The game's challenge comes primarily from interpreting and processing the on-screen stimuli and choosing a correct response: a simple motor movement of pressing one of two keys. The stimuli are the on-screen obstacles that players must observe and process, and the player must decide whether to move or stay still. The human visual perception system takes in far more information than can actually be processed at any point in time [122], so performance improvements in this task are due to the player developing a "coding" [89] of the patterns of the obstacles that they can use to identify what they are seeing. If the player correctly identifies that a response is required, then they have to choose the correct response in terms of the right or left arrow keys and the amount to move. This requires players to develop a mapping between the key and the action, something that also results in improved performance. While this skill is relatively straightforward, it can still be considered to be a high-performance skill similar to those found in many sports and esports, in that there are very clear differences between experts and novices [45, 32] and performance gains continue for a long time [99, 32].

EXPERIMENT

We designed an online experiment to test the effects of spaced practice in a high-performance game environment[1]. We had participants play a total of 25 minutes of our *Super Hexagon* clone, broken into five-minute sessions — four training sessions and one retention session.

The first four sessions served as the training period, and between these sessions participants were given inter-session rests, with those rests varying depending on the group they were randomly assigned to. There were five inter-session rest interval groups: A Continuous Practice Group—operationalized as an inter-session rest interval of 3 seconds—and four Spaced Practice Groups, with inter-session rest intervals of 2 minutes, 5 minutes, 10 minutes, or 1 day. These rest intervals were chosen based on the Donovan and Radosevich [25] meta-review, which found that tasks similar to ours benefited most from a rest interval between 1 and 10 minutes, although we included a 1-day interval to make further comparisons. To evaluate how these different rest periods affected the participants' learning, a final 5-minute retention session was completed one day after the final training session. The delay was chosen to ensure that any temporary effects of spacing practice would not be present [25].

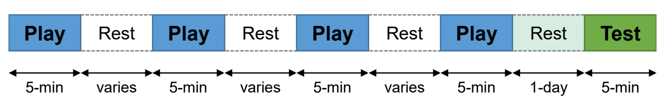

Figure 3: The procedure of the experiment.

To accurately reflect what would happen "in the wild" when players are given rest periods, we instructed the participants assigned to the Spaced Practice Groups that they were allowed to use their computer in any way they wished during the periods of rest, but that they should return to the task promptly when the rest period ended. We notified participants via an audio cue to return to the game, except for the *1-Day* interval group, who were invited back to the task via an email. One day after their last training session, all participants were invited back via email to complete the final five-minute retention session.

[1]The design of this experiment was pre-registered on OSF.io at: https://osf.io/sk2w9

Measures

Questionnaires

We used several questionnaires to get a sense of each participant's interest in the task and ability to complete the task.

Gaming Experience. We expected that individuals who enjoy playing video games regularly might enjoy our game more than those who do not, and so may try harder or improve more quickly. We therefore asked participants a number of questions to evaluate their level of experience with video games, including: "How much do you self-identify as a gamer?", "How many years have you been playing video games?", "How often (on average) do you play video games?", and "If you have played games more often in the past, how often were you playing at peak times?".

Attentional Control. We anticipated that participants who are better at maintaining attention and focus on the task might perform better. We therefore used Derryberry and Reed's Attentional Control Scale (ACS) [22] to measure each participant's attentional control.

Immersive Tendencies. We used Witmer and Singer's Immersive Tendencies Questionnaire (ITQ) [128] to measure participants' tendency to experience presence in virtual environments. The questionnaire consists of three subscales: involvement (propensity to get involved with an activity), focus (ability to concentrate on enjoyable activities), and games (how much they play games and whether they become involved enough to feel like they are inside the game).

Current Motivation. Participants may have different levels of interest in completing our task, so we used Guay et al.'s Situational Motivation Scale (SIMS) [42] to measure the participant's intrinsic motivation, identified regulation, external regulation, and amotivation towards our experiment. We also used Rheinberg et al.'s Questionnaire on Current Motivation (QCM) [91, 120] to measure the participant's task-related anxiety, probability of success, interest, and challenge.

Achievement Orientation. Participants with a competitive nature may invest more effort into the task, so we measured their competitiveness, win orientation, and goal orientation using Gill and Deeter's Sport Orientation Questionnaire [41].

Dependent Measures

We used four dependent measure for the experiment (one for performance, and three for subjective experience). Performance was measured for every session while subjective experience was measured twice: after the last training session and after the retention test.

Average Life Time. For each session, the participant played multiple rounds. Looking at the sessions individually, we took the average of the scores (time) among each session's rounds. If the round was ended prematurely by the timer running out, that round was excluded from the average.

Intrinsic Motivation. We evaluated the participants' intrinsic motivation towards the game by using the Intrinsic Motivation Inventory (IMI) [71]. The IMI measures the participant's interest-enjoyment, effort-importance, and tension-pressure.

Flow. We used Engeser and Rheinberg's Flow Short Scale (FSS) [31] to measure the participant's experience of flow while playing the game. We used the fluency of performance and absorption by activity subscales.

Immersion. We used Jennett et al.'s questionnaire on in-game immersion [51]. The questionnaire measures the participant's attention, temporal dissociation, transportation, challenge, emotional involvement, and enjoyment.

Recruitment and Participants

Our online experiment was conducted on Amazon's Mechanical Turk—a system that acts as a broker between willing workers and requesters, who provide paid human intelligence tasks (HITs). Mechanical Turk has been used for human-computer interaction research in the past (e.g., [53, 10, 103, 60]) and has been proven to be effective, as long as some attention is given to verifying the quality of the data [26, 73, 69, 86, 85].

We were interested in recruiting exclusively *Super Hexagon* novices, so we first posted a screening task, which asked three questions: "What type of input device are you currently using?", "Which hand are you currently using for that input device?", and "Do you have any experience with the game *Super Hexagon*, or its clones?", on a five-point scale from "no experience" to "extremely experienced". The first two questions served to obscure our intent. We had 535 workers complete the HIT. Of those, we gave 397 that indicated they had no experience with *Super Hexagon* a "qualification" that would permit them to participate in our experiment.

The immediate time commitment of the training phase could vary considerably between conditions (from 5 minutes to 50 minutes), so we were transparent in outlining how the the immediate time requirements could vary based on random assignment to condition. Once randomly assigned, the participants MTurk ID became associated with a particular condition and so attempts to be re-assigned to different conditions would be unsuccessful. These steps were taken to avoid self-selection into groups, however, we observed differential dropout in that those assigned the *1-Day* group seemed more likely to not return for subsequent days of practice.

Ethical approval for this study was obtained from the behavioral ethics board of the University of Saskatchewan, and participants were asked to renew their consent at the start of each day's task. To comply with ethical guidelines, the task was only available to workers from the United States who were over 18 years old. Participants were paid $5.50 USD for the training sessions, and an additional $1.50 USD for the retention test (approximately $10 an hour). We had 138 participants complete both the training and the retention test.

Data collected from Mechanical Turk can include low-quality responses [26, 73]; therefore, we filtered out non-compliant participants based on a variety of criteria. We excluded 2 participants due to entering an invalid age (less than 18 or greater than 99) and 13 participants due to taking a rest that was more than 1 standard deviation longer than their peers within their rest group (actual rest periods for each group are reported in Table 1). We also filtered out participants based on their in-game logs indicating low framerate (11 participants) or evidence suggesting that they stopped playing the game (9 participants). This was determined by counting the number of sessions in which they performed worse than their initial session, and excluding them if they performed worse in two or more sessions. In total, 33 participants were excluded (some met more than one criteria for exclusion), leaving 105 participants—52 female and 53 male, with an average age of 37.8 (min 19, max 68). Table 1 lists how participants were distributed among the conditions.

Results

To investigate the Performance differences between groups, we used SPSS to perform a repeated measures analysis of covariance (RM-ANCOVA). We used Session (2, 3, 4, and the Retention Session) as the repeated-measure factor, and inter-session rest interval (3 seconds for the Continuous Practice Group and 2 minutes, 5 minutes, 10 minutes, or 1 day for the Spaced Practice Groups) as the between-subjects factor (with all groups being subsequently being referred to as the "Interval Groups"). In-game Performance (average life time for Session) was used as a dependent measure. Instead of using Session 1 as one of the repeated measures, we used it as a covariate to compensate for individual differences in initial performance. This approach means that we cannot make comparisons to Session 1; however, it has been shown to provide more statistical power under similar circumstances [11], and we wanted to be certain that we accounted for individual differences in *initial* performance. To further acknowledge individual differences between participants, we included further covariates—the results of the subscales from our trait questionnaires of gaming experience, attentional control, immersive tendencies, achievement orientation, and current motivation—in our RM-ANCOVA. All pairwise comparisons were made using Bonferroni corrections. This approach allowed us to answer a number of questions.

Did spaced practice result in differences?

Yes. We found a significant effect of Interval Group on Performance ($F_{4,85} = 3.649$, $p = .009$, $\eta_p^2 = .760$), indicating that there were differences between Interval Groups. We also found a significant interaction between Session and Interval Group ($F_{12,255} = 3.078$, $p < .001$, $\eta_p^2 = .127$), indicating that there were differences in Performance over the Sessions that were due to the Rest Group.

Examining the Sessions individually, there were no significant differences between the Interval Groups for Session 2 ($F_{4,85} = 2.173$, $p = .079$, $\eta_p^2 = .093$), but there were for Session 3 ($F_{4,85} = 3.309$, $p = .014$, $\eta_p^2 = .135$), Session 4 ($F_{4,85} = 5.310$, $p = .001$, $\eta_p^2 = .200$), and the Retention Session ($F_{4,85} = 2.565$, $p = .044$, $\eta_p^2 = .108$).

How did the Interval Groups compare?

Pairwise comparisons for Session 3 showed only that the Continuous Practice Group performed significantly worse than the *10-Minute* Interval Group ($p = .005$), with all other comparisons not being significant ($p \geq .501$). For Session 4 (the final training session), pairwise comparisons showed that the Continuous Practice Group performed significantly worse than the *2-Minute*, *5-Minute*, and *10-Minute* Spaced Practice Groups

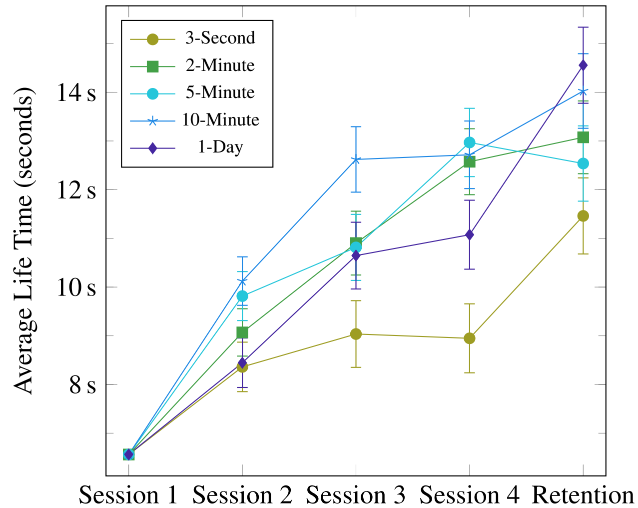

Figure 4: Improvements in average life time across all Sessions and Interval Groups, presented as the estimated marginal means from the RM-ANCOVA. Error bars show standard error.

(all $p \leq .004$), but not the *1-Day* Interval Group ($p = .339$), with all other comparisons being not significant ($p \geq .648$). There were no significant differences between the Interval Groups for the Retention Session ($p = .056$ for the comparison between *1-Day* and *3-Second*, $p \geq .254$ for the others).

How did Performance change over time?

There was a main effect of Session when controlling for Session 1's Performance ($F_{3,255} = 8.189$, $p \leq .001$, $\eta_p^2 = .088$). Pairwise comparisons between the Sessions revealed that every Session was different from the others (all $p < .001$).

It was not possible to determine significance for comparisons to Session 1 Performance due to its use as a covariate instead of a repeated measure. Therefore, to evaluate whether significant improvements were made for each spacing interval, we make our comparisons to Session 2: Session 4 to Session 2 and the Retention Session to Session 2. After the final training session (Session 4), the Spaced Practice Groups (*2-Minute*, *5-Minute*, *10-Minute*, and *1-Day*) had made significant improvements to their Performance (all $p \leq .001$), whereas the Continuous Practice Group (*3-Second*) had not ($p > .999$). After the Retention Session, every Interval Group had made significant improvements in Performance (all $p \leq .001$).

Table 1 shows the Performance gains for each Interval Group over each Session. We observe that the Continuous Practice Group made only marginal improvements during training (Sessions 1 to 4), but were able to make up for it with a significant improvement in the Retention Session. In comparison, the Spaced Practice Groups made significant improvements in Performance on some of the training sessions, but not all, and the only Spaced Practice Group that continued to improve on the Retention Session was the *1-Day* Interval Group.

How did spacing affect subjective experience?

To evaluate differences in subjective experience, we used a separate RM-MANCOVA. We collected subjective measures after the final training session and Retention Session; the

Session Comparison / Measure	2–1 Δ Mean	3–2 Δ Mean	3–2 p	4–3 Δ Mean	4–3 p	Retention–4 Δ Mean	Retention–4 p	4–1 Δ Mean	Retention–1 Δ Mean
3-Second (actual: 3.01 ± 0.06 sec.; $n = 20$)	1.799	0.674	>.999	-0.088	>.999	**2.513**	**.004**	2.386	4.899
2-Minute (actual: 2.20 ± 0.21 min.; $n = 22$)	2.507	**1.834**	**.009**	**1.671**	**.026**	0.501	>.999	6.012	6.513
5-Minute (actual: 6.49 ± 1.69 min.; $n = 21$)	3.253	0.998	.531	**2.155**	**.003**	-0.432	>.999	6.407	5.975
10-Minute (actual: 11.08 ± 1.26 min.; $n = 21$)	3.561	**2.498**	**<.001**	0.095	>.999	1.310	.382	6.154	7.464
1-Day (actual: 24.64 ± 2.29 hr.; $n = 21$)	1.884	**2.200**	**.002**	0.429	>.999	**3.480**	**<.001**	4.513	7.993

Table 1: The Performance difference between the Sessions for each Interval Group. Bold text indicates a significant difference between the two Sessions' Performance. Δ Mean is the difference between the estimated marginal means of the two sessions. Session 1's normalized Performance (average life time) was evaluated at 6.562 seconds, as computed by the RM-ANCOVA. ± on actual rest time is the standard deviation.

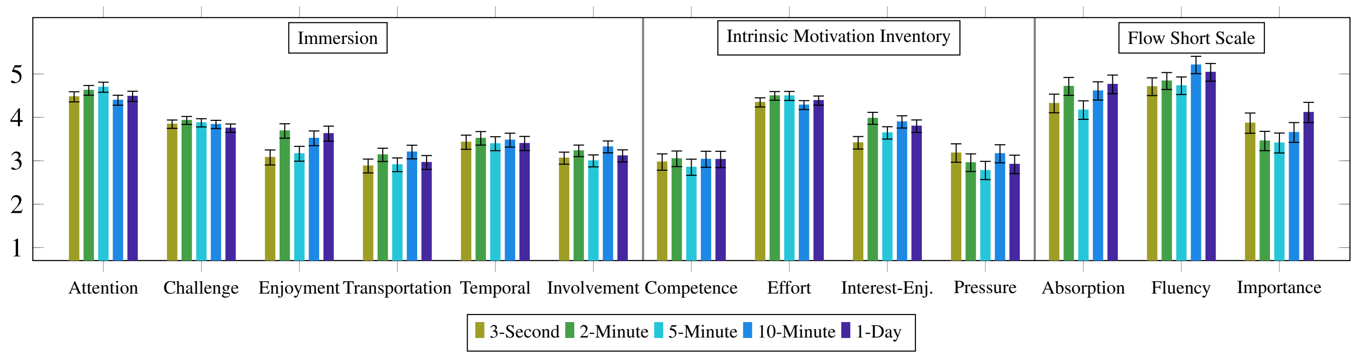

Figure 5: Subjective measures from the subscales of Immersion, Intrinsic Motivation, and Flow. Presented as the estimated marginal means of the combination of both Sessions from the RM-ANCOVA. Error bars show standard error.

results of the subscales for Intrinsic Motivation (IMI), Flow (FSS), and Immersion were used as the dependent measures. We used the same covariates and between-subjects factor as in the previously described RM-ANCOVA.

Only the measures of Enjoyment differed between the Interval Groups, for both the Immersion questionnaire's Enjoyment subscale ($F_{4,85} = 2.733$, $p = .034$, $\eta_p^2 = .114$), and the Intrinsic Motivation Inventory's (IMI) Interest-Enjoyment subscale ($F_{4,85} = 2.490$, $p = .049$, $\eta_p^2 = .105$). All other subscales had no significant differences due to Interval Group ($p \geq .202$). Examining the Sessions individually revealed that the differences in enjoyment were present after the training sessions ($F_{4,85} = 3.473$, $p = .011$, $\eta_p^2 = .140$ for Immersion, and $F_{4,85} = 3.001$, $p = .023$, $\eta_p^2 = .124$ for IMI), but not after the Retention Session (both $p \geq .168$).

We used pairwise comparisons (Bonferroni) to examine the between-group differences after the final training session and found no significant differences between the groups for Enjoyment (Immersion) ($p = .051$ between *3-Second* and *2-Minute*, $p = .060$ between *3-Second* and *1-Day*, and $p \geq .283$ for the others). The only difference between the Interval Groups for Interest-Enjoyment (IMI) was between the Continuous Practice Group and *2-Minute* Interval Group ($p = .022$), with all other comparisons being not significant ($p \geq .143$).

In summary, there were marginal differences in subjective experience, but only for Enjoyment. The Continuous Practice Group enjoyed the experience less than the *2-Minute* Interval Group (mirroring the performance results for these groups).

Summary of Results

Our results indicated that there were differences in Performance as a result of varying the inter-session rest interval. After training (Session 4), the Continuous Practice Group (*3-Second*) performed significantly worse than every Spaced Practice Group, except for the *1-Day* Interval Group. After a one-day break, on the Retention Session, these two groups effectively "caught up" to the performance of the other Interval Groups. In terms of our framework, this indicates that in our scenario, spaced practice had a strong effect on immediate performance, but a relatively weak effect on long-term performance. Even though there are no significant differences between groups after the Retention Session, the Continuous Practice Group performed the worst over all, looking at the changes in performance (see Table 1).

Although the performance essentially equalized by the Retention Session, there were differences in how the Interval Groups arrived at that performance. The *10-Minute* Interval Group's Performance increased faster than any other between Sessions 1 and 2 and 3, but their Performance improvements slowed down between Session 4 and 3 and the Retention Session and Session 4. The *5-Minute* Interval Group's Performance gain between Session 1 and 2 was the second highest, but they failed to make significant improvements between Session 3 and 2, then made up for it with improvements between Session 4 and 3. They then did not improve at all over the Retention Session and were the only Interval Group to perform worse on that Session compared to their Session 4 Performance. The *2-Minute* Interval Group was the only Spaced Practice Group that made gains in Performance consistently over the sessions. They improved third most between Session 1 and 2, and then improved significantly between Sessions 2, 3, and 4. However, they did not improve on their Session 4 Performance in the Retention Session. The *1-Day* Interval Group made the least gains in Performance between Sessions 1 and 2 of all the Spaced Practice Groups, significantly improved their performance between

Sessions 2 and 3, but then did not between Sessions 3 and 4. They made up for it with large Performance improvements between Session 4 and the Retention Session. The Continuous Practice Group made the fewest gains in Performance over the training sessions, improving the least between Session 1 and 2, and with no significant improvements in Performance between Sessions 2, 3, and 4. They did improve between Session 4 and the Retention Session, allowing them to "catch up" to the other group's performance.

Our measures of subjective experience indicated that introducing spaced practice made only minor changes to the players' experience. We found differences on only one measure—enjoyment—and only when comparing the Continuous Practice Group to the *2-Minute* Interval Group after the training sessions. The Continuous Practice Group experienced less enjoyment than the *2-Minute* Interval Group.

DISCUSSION

In the following sections we look back to our three main research questions, consider possible explanations for certain results, outline design implications for game designers, and discuss limitations and future work.

Returning to the Research Questions

Is spaced practice beneficial for game skill development?

Our first research question asked if there are benefits of spaced practice on performance in digital games and our third asked whether performance benefits last beyond the training session. Our results clearly show that taking breaks between practice sessions does improve the development of game skills (in our *Super Hexagon* clone) from session to session—all of the interval groups improved more than the continuous-play group (who did not make significant progress through the entire twenty minutes of practice). Furthermore, we found this result despite the possibility that participants were not using these breaks to actually rest—they could have simply switched to a different task—just as actual gamers might use the break.

However, despite these differences during practice, the results of our retention test suggest that the continuous group's lack of improvement did not mean that they failed to learn how to play the game. After one day, the continuous group saw a large (and significant) jump in performance, and the performance differences between all of the groups largely disappeared. As discussed further below, both of these results can be usefully exploited by designers who want to support improvement in either the short term or the long term.

Is there an optimal rest period for our game?

The idea that a particular task has an optimal rest period is fairly common [25, 16, 72, 12, 68]; however, there was no obvious best interval in our results. The 10-minute interval group made large initial gains, but their rate of improvement decreased at the end of the training period, and their retention-test performance was not significantly better than the other groups. The 2-minute interval is a second candidate for an optimal rest period: this group made consistent improvements, and rated their enjoyment higher; the other advantage of the 2-minute interval is that it may be easier to impose on players (either explicitly, or implicitly through game design). The 1-day interval group saw the largest overall gain from the first session to the retention test—but this is likely too long a break to be feasibly implemented. It is unlikely that players will be willing to stop playing after five minutes and then wait an entire day before returning to the game.

An interesting idea proposed in previous work is that the rest interval should change as the learner gains experience. For example, Snoddy [106] thought that the rest interval in general became less effective with continued practice; and Kim et al. [59] suggest that the interval should be tuned to the stage of learning. In this approach, the learner needs longer breaks during early stages when they are primarily dealing with declarative knowledge (in the cognitive stage), but shorter breaks, or none at all, when they are starting to compile their knowledge into procedures (in the associative stage) [59, 3, 78].

Explanation of Results

Although our results are focused on outcomes rather than explanatory power, there are several issues that can help researchers understand some of our findings:

- Why did the continuous group perform poorly? Learning theories that promote rest intervals suggest that the main problem with continuous practice is that there is no time for the brain to generalize the feedback that has been gathered. This idea seems particularly relevant to games—many players have experienced the feeling of being "stuck in a rut" when trying to solve a particular problem.
- Why did the 10-minute interval group stop improving? As described above, there is debate about whether the rest interval should remain the same at all levels of experience. It is possible that the 10-minute rest group was at a different learning stage by Session 3, and would have continued to improve if their rests had been shorter. More research is needed to explore this issue further.
- Why did performance equalize after the one-day break? We believe that this improvement may simply be an example of the effectiveness of a different kind of break. That is, the one-day break may have allowed players from all the interval groups to generalize their practice experiences and compile declarative knowledge into procedures, leading to improvements on our retention test. Further study should examine the effects of different retention delays.

Finally, our experimental setting—and in particular, the motivation of our players—may have played a role in some of our results. For example, a player's willingness to take breaks (and potentially, the benefit that they receive from a break) may be strongly affected by their motivation to keep playing the game. For example, if one of our interval groups had more highly-motivated gamers who resented the breaks, it could have changed the performance relative to the other groups. (We note, however, that we included some of these traits as covariates in our analyses in order to try and account for their effects—but further analysis of these factors is needed).

Implications for Design

Our results have important implications for game designers and players. First, the during-training improvements seen are valuable because players are highly interested in short-term improvement (i.e., during their current game session). Our results suggest that players can make significantly more progress in learning a game by taking breaks as small as two minutes—and can see these results in a very short time (after one or two five-minute play periods). Second, the delayed improvement shown by the continuous-practice group means that players who dislike pausing can still get the benefits of spacing by taking a longer break after their play session (although the benefits will not be seen right away). This is valuable because some players wish to remain immersed in the game (e.g., as shown by the disagreement over the spawn locations in *Dark Souls*). These results, combined with little to no difference in subjective experience, mean that designers are not constrained to one or the other approach, and can even make alternate paths available for players with different interests.

Although our results show clear benefits of pausing during practice, some players may argue that it is more efficient to practice continuously, because the time spent resting could be used for additional attempts. We did not investigate this comparison directly; however, our data suggests that breaks may be as useful as additional practice, at least for initial learning. In our study, a single break for the 10-minute interval group allowed them to reach the same level of performance as the group that played continuously for 20 minutes. This issue needs to be explored in greater detail, but these early results underscore the suggestion that there are valuable processes going on in the brain during a break.

It is also important to note that despite the differences we saw in training, our retention test suggested that the continuous practice group's lack of improvement did not result in them failing to learn how to play the game compared to the spaced practice groups (i.e., performance for all groups was similar at the retention test). Therefore, the decision whether to play or pause may depend on the player's immediate goals. If the player is making steady performance improvements, then there may be no need to take a break. If, however, the player has reached a point in the game where their performance improvements have stalled, or they are up against a challenge that they cannot overcome, then a short break may be just what they need to continue to improve or overcome that challenge.

Finally, we allowed participants to do whatever they wished to during breaks; however, a game could be designed in such a way as to help a player to take a break from the psycho-motor task, but not leave the game environment. Inventory management, avatar upgrades, side quests, world exploration, loading screens, and cut scenes are all examples of activities that would allow for a break from the challenging task, but allow players to stay in-game. Designers already utilize these activities in various ways and there are opportunities for helping players improve performance through clever pacing switching activities to force breaks and optimize skill development.

Generalization and Future Work

One aspect of this work that could limit generalizablity to other games was our choice to use *Super Hexagon*. In contrast to some digital games with dedicated tutorial sections and complex gameplay mechanics, *Super Hexagon* requires new players to learn only two buttons and understand that the obstacles should be avoided. This difference in required declarative knowledge could affect how well spacing practice works or the underlying mechanisms that result in improved performance (e.g., allowing time to compile declarative knowledge into procedural knowledge [3, 59] or if fatigue is a factor [1]). Previous research on spaced practice suggests that the type of task moderates the effect (e.g., whether it is complex and requires cognitive skill [25], is discrete rather than continuous [65], or requires declarative knowledge [59]). These task factors likely apply to different types and genres of games as well. Therefore, future work could investigate this explicitly, testing games from different genres.

If this effect is robust and occurs in a variety of games, then further research could explore *why* taking a break works to improve performance. For example, a cognitive neuroscientist might be able to identify how neural systems respond differently to spaced versus continuous practice. Another possible approach is to vary what activity participants are doing during the break to understand what it is about the break that results in improved performance. These could easily be informed by game design. Consider, for example, how *Dark Souls* gives players an implicit break after each failure. Instead of forcing the player to step away from or exit the game, the player can be made to engage with the game in a less demanding way — in a role-playing game like *Dark Souls*, the activities could include looting, storytelling, or inventory management.

CONCLUSIONS

Practice to develop skill is an important part of playing a game, and the best way to practice is of interest to both game designers and players. Taking rest breaks during practice has been widely studied in learning research, but there is little knowledge of whether spacing aids the development of game skills. We carried out a study in which players completed five-minute practice blocks with either continuous play or variable spacing between the blocks (from two minutes to 24 hours). Our work provides valuable contributions that can change the way that game designers and game plays consider training:

- We show how spaced practice affects perceptual-motor tasks in a real game (a domain that has seen little research in practice) and in a naturalistic setting.
- We show that players can improve their short-term performance by taking breaks, something that many players essentially never consider, and that game designers can assist player performance simply by introducing a break.
- We show that "just playing through" is a bad approach (something many players do), especially when immediate performance gains would help them make progress.

Our results provide useful information for designers who want to encourage skill development, and for players who simply want to get better at their favourite game.

ACKNOWLEDGEMENTS

We would like to thank SWaGUR for providing funding, the members of the HCI lab for their friendship and support, and Parker Neufeld for developing the *Super Hexagon* clone.

REFERENCES

[1] Jack A. Adams. 1987. Historical review and appraisal of research on the learning, retention, and transfer of human motor skills. *Psychological Bulletin* 101, 1 (1987), 41–74. DOI:http://dx.doi.org/10.1037/0033-2909.101.1.41

[2] Robert B. Ammons. 1950. Acquisition of motor skill: III. Effects of initially distributed practice on rotary pursuit performance. *Journal of Experimental Psychology* 40, 6 (1950), 777–787. DOI:http://dx.doi.org/10.1037/h0061049

[3] John R. Anderson. 1982. Acquisition of cognitive skill. *Psychological Review* 89, 4 (1982), 369–406. DOI:http://dx.doi.org/10.1037/0033-295x.89.4.369

[4] Graham Aston. 2017. What is the Optimum Training Time for Esports Players? *The Esports Observer* (December 2017). Retrieved April 4, 2019 from https://esportsobserver.com/optimum-player-training-time/

[5] A. D. Baddeley and D. J. A. Longman. 1978. The Influence of Length and Frequency of Training Session on the Rate of Learning to Type. *Ergonomics* 21, 8 (aug 1978), 627–635. DOI:http://dx.doi.org/10.1080/00140137808931764

[6] Scott Bateman, Regan L. Mandryk, Tadeusz Stach, and Carl Gutwin. 2011. Target assistance for subtly balancing competitive play. In *Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11*. ACM Press. DOI:http://dx.doi.org/10.1145/1978942.1979287

[7] Roy F. Baumeister. 1984. Choking under pressure: Self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of Personality and Social Psychology* 46, 3 (1984), 610–620. DOI:http://dx.doi.org/10.1037/0022-3514.46.3.610

[8] Walter R. Boot, Anna Sumner, Tyler J. Towne, Paola Rodriguez, and K. Anders Ericsson. 2016. Applying Aspects of the Expert Performance Approach to Better Understand the Structure of Skill and Mechanisms of Skill Acquisition in Video Games. *Topics in Cognitive Science* 9, 2 (nov 2016), 413–436. DOI:http://dx.doi.org/10.1111/tops.12230

[9] Lyle E. Bourne and E. James Archer. 1956. Time continuously on target as a function of distribution of practice. *Journal of Experimental Psychology* 51, 1 (1956), 25–33. DOI:http://dx.doi.org/10.1037/h0039887

[10] Jason T. Bowey, Max V. Birk, and Regan L. Mandryk. 2015. Manipulating Leaderboards to Induce Player Experience. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play (CHI PLAY '15)*. ACM, New York, NY, USA, 115–120. DOI:http://dx.doi.org/10.1145/2793107.2793138

[11] Gerard J.P. Van Breukelen. 2006. ANCOVA versus change from baseline had more power in randomized studies and more bias in nonrandomized studies. *Journal of Clinical Epidemiology* 59, 9 (sep 2006), 920–925. DOI:http://dx.doi.org/10.1016/j.jclinepi.2006.02.007

[12] Ryan Brydges, Heather Carnahan, David Backstein, and Adam Dubrowski. 2007. Application of Motor Learning Principles to Complex Surgical Tasks: Searching for the Optimal Practice Schedule. *Journal of Motor Behavior* 39, 1 (jan 2007), 40–48. DOI:http://dx.doi.org/10.3200/jmbr.39.1.40-48

[13] budda loves you. 2015. Titan Souls General Discussions - Is there any reason why it doesn't just spawn you outside of the boss door when you die. *Steam Community* (2015). Retrieved September 13, 2018 from https://steamcommunity.com/app/297130/discussions/0/618458030655379691/

[14] Terry Cavanagh. 2012. *Super Hexagon*.

[15] Terry Cavanagh. 2012. *Super Hexagon*. Retrieved August 30, 2018 from https://www.superhexagon.com/

[16] Nicholas J. Cepeda, Noriko Coburn, Doug Rohrer, John T. Wixted, Michael C. Mozer, and Harold Pashler. 2009. Optimizing Distributed Practice: Theoretical Analysis and Practical Implications. *Experimental Psychology* 56, 4 (jan 2009), 236–246. DOI:http://dx.doi.org/10.1027/1618-3169.56.4.236

[17] Jenova Chen. 2007. Flow in games (and everything else). *Commun. ACM* 50, 4 (apr 2007), 31. DOI:http://dx.doi.org/10.1145/1232743.1232769

[18] Andy Cockburn, Carl Gutwin, Joey Scarr, and Sylvain Malacria. 2014. Supporting Novice to Expert Transitions in User Interfaces. *Comput. Surveys* 47, 2 (nov 2014), 1–36. DOI:http://dx.doi.org/10.1145/2659796

[19] S. E. Criscimagna-Hemminger and R. Shadmehr. 2008. Consolidation Patterns of Human Motor Memory. *Journal of Neuroscience* 28, 39 (sep 2008), 9610–9618. DOI:http://dx.doi.org/10.1523/jneurosci.3071-08.2008

[20] Peter F. Delaney, Peter P.J.L. Verkoeijen, and Arie Spirgel. 2010. Spacing and testing effects: A deeply critical, lengthy, and at times discursive review of the literature. In *Psychology of Learning and Motivation*. Elsevier, 63–147. DOI:http://dx.doi.org/10.1016/s0079-7421(10)53003-2

[21] Frank N. Dempster. 1988. The spacing effect: A case study in the failure to apply the results of psychological research. *American Psychologist* 43, 8 (1988), 627–634. DOI:http://dx.doi.org/10.1037/0003-066x.43.8.627

[22] Douglas Derryberry and Marjorie A. Reed. 2002. Anxiety-related attentional biases and their regulation by attentional control. *Journal of Abnormal Psychology* 111, 2 (2002), 225–236. DOI:http://dx.doi.org/10.1037/0021-843x.111.2.225

[23] Sebastian Deterding. 2015. The lens of intrinsic skill atoms: A method for gameful design. *Human-Computer Interaction* 30, 3-4 (2015), 294–335. DOI:http://dx.doi.org/10.1080/07370024.2014.993471

[24] Emanuel Donchin. 1995. Video games as research tools: The Space Fortress game. *Behavior Research Methods, Instruments, & Computers* 27, 2 (jun 1995), 217–223. DOI:http://dx.doi.org/10.3758/bf03204735

[25] John J. Donovan and David J. Radosevich. 1999. A meta-analytic review of the distribution of practice effect: Now you see it, now you don't. *Journal of Applied Psychology* 84, 5 (1999), 795–805. DOI:http://dx.doi.org/10.1037/0021-9010.84.5.795

[26] Julie S. Downs, Mandy B. Holbrook, Steve Sheng, and Lorrie Faith Cranor. 2010. Are Your Participants Gaming the System?: Screening Mechanical Turk Workers. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 2399–2402. DOI:http://dx.doi.org/10.1145/1753326.1753688

[27] Stephen Draper. 2017. Esports Terminology: Skill Ceilings and Skill Floors. *Esports Edition* (September 2017). Retrieved April 4, 2019 from https://esportsedition.com/general/skill-ceiling-skill-floor-esports-terminology/

[28] H. Ebbinghaus, H.A. Ruger, and C.E. Bussenius. 1913. *Memory: A Contribution to Experimental Psychology*. Teachers College, Columbia University.

[29] Darren Edge, Elly Searle, Kevin Chiu, Jing Zhao, and James A. Landay. 2011. MicroMandarin. In *Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11*. ACM Press. DOI:http://dx.doi.org/10.1145/1978942.1979413

[30] H C Ellis. 1978. *Fundamentals of Human Learning, Memory, and Cognition*. W. C. Brown Company.

[31] Stefan Engeser and Falko Rheinberg. 2008. Flow, performance and moderators of challenge-skill balance. *Motivation and Emotion* 32, 3 (sep 2008), 158–172. DOI:http://dx.doi.org/10.1007/s11031-008-9102-4

[32] K. Anders Ericsson and Robert Pool. 2016. *Peak: Secrets from the new science of expertise*. Houghton Mifflin Harcourt.

[33] M W Eysenck, M T Keane, and G Mather. 2010. *Cognitive Psychology: A Student's Handbook, 6th Edition*. Routledge Chapman & Hall.

[34] Paul M Fitts. 1964. Perceptual-motor skill learning. In *Categories of human learning*. Elsevier, 243–285.

[35] Paul M. Fitts and James R. Peterson. 1964. Information capacity of discrete motor responses. *Journal of Experimental Psychology* 67, 2 (1964), 103–112. DOI:http://dx.doi.org/10.1037/h0045689

[36] Paul Morris Fitts and Michael I. Posner. 1967. Human Performance. (1967).

[37] FromSoftware. 2011. *Dark Souls*. Namco Bandai Games.

[38] Rosemary Garris, Robert Ahlers, and James E. Driskell. 2002. Games, Motivation, and Learning: A Research and Practice Model. *Simulation & Gaming* 33, 4 (dec 2002), 441–467. DOI:http://dx.doi.org/10.1177/1046878102238607

[39] G.M. Gazda and R.J. Corsini. 1980. *Theories of Learning: A Comparative Approach*. F. E. Peacock Publishers.

[40] James Paul Gee. 2005. Learning by Design: Good Video Games as Learning Machines. *E-Learning and Digital Media* 2, 1 (mar 2005), 5–16. DOI:http://dx.doi.org/10.2304/elea.2005.2.1.5

[41] Diane L. Gill and Thomas E. Deeter. 1988. Development of the Sport Orientation Questionnaire. *Research Quarterly for Exercise and Sport* 59, 3 (sep 1988), 191–202. DOI:http://dx.doi.org/10.1080/02701367.1988.10605504

[42] Frédéric Guay, Robert J Vallerand, and Céline Blanchard. 2000. On the assessment of situational intrinsic and extrinsic motivation: The Situational Motivation Scale (SIMS). *Motivation and emotion* 24, 3 (2000), 175–213. DOI:http://dx.doi.org/10.1037/t12678-000

[43] Arin Hanson. 2011. Sequelitis - Mega Man Classic vs. Mega Man X. (2011). https://www.youtube.com/watch?v=8FpigqfcvlM

[44] Daniel Himmelstein, Yitong Liu, and Jamie L. Shapiro. 2017. An Exploration of Mental Skills Among Competitive League of Legend Players. *International Journal of Gaming and Computer-Mediated Simulations* 9, 2 (apr 2017), 1–21. DOI:http://dx.doi.org/10.4018/ijgcms.2017040101

[45] Nicola J. Hodges, Raoul Huys, and Janet L. Starkes. 2012. Methodological Review and Evaluation of Research in Expert Performance in Sport. In *Handbook of Sport Psychology*. John Wiley & Sons, Inc., 159–183. DOI:http://dx.doi.org/10.1002/9781118270011.ch7

[46] Dennis H. Holding. 1989. Skills Research. In *Human Skills*. John Wiley & Sons, 1–16.

[47] J. Honeybourne, M. Hill, and H. Moors. 2000. *AS Level Textbook*. Nelson Thornes Limited.

[48] HowLongToBeat. 2018. How long is Super Hexagon? (2018). Retrieved September 18, 2018 from https://howlongtobeat.com/game.php?id=9353

[49] Ioanna Iacovides, Anna L. Cox, Ara Avakian, and Thomas Knoll. 2014. Player Strategies: Achieving Breakthroughs and Progressing in Single-Player and Cooperative Games. *Proceedings of the first ACM SIGCHI annual symposium on Computer-human interaction in play - CHI PLAY '14* (2014), 131–140. DOI:http://dx.doi.org/10.1145/2658537.2658697

[50] JC Fletcher. 2012. Super Hexagon has sold approximately 45,000 copies. *Engadget* (2012). Retrieved September 18, 2018 from https://www.engadget.com/2012/09/21/super-hexagon-has-sold-approximately-45-000-copies/

[51] Charlene Jennett, Anna L. Cox, Paul Cairns, Samira Dhoparee, Andrew Epps, Tim Tijs, and Alison Walton. 2008. Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies* 66, 9 (sep 2008), 641–661. DOI: http://dx.doi.org/10.1016/j.ijhcs.2008.04.004

[52] Colby Johanson. 2016. *Scaffolding Novices to Leverage Auditory Awareness Cues in First-Person Shooters*. Master's thesis. University of Saskatchewan, Saskatoon, Canada.

[53] Colby Johanson, Carl Gutwin, and Regan L. Mandryk. 2017. The Effects of Navigation Assistance on Spatial Learning and Performance in a 3D Game. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play - CHI PLAY '17*. ACM Press. DOI: http://dx.doi.org/10.1145/3116595.3116602

[54] Colby Johanson and Regan L Mandryk. 2016. Scaffolding Player Location Awareness through Audio Cues in First-Person Shooters. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 3450–3461. DOI:http://dx.doi.org/10.1145/2858036.2858172

[55] Jesper Juul. 2013. *The art of failure: An essay on the pain of playing video games*. Mit Press.

[56] Tadhg Kelly. 2018. Conveyance. *What Games Are* (2018). Retrieved November 28, 2018 from https://www.whatgamesare.com/conveyance.html

[57] Kristian Kiili. 2005. Digital game-based learning: Towards an experiential gaming model. *The Internet and higher education* 8, 1 (2005), 13–24.

[58] Jong Wook Kim and Frank E. Ritter. 2015. Learning, Forgetting, and Relearning for Keystroke- and Mouse-Driven Tasks: Relearning Is Important. *Human–Computer Interaction* 30, 1 (2015), 1–33. DOI: http://dx.doi.org/10.1080/07370024.2013.828564

[59] Jong W. Kim, Frank E. Ritter, and Richard J. Koubek. 2013. An integrated theory for improved skill acquisition and retention in the three stages of learning. *Theoretical Issues in Ergonomics Science* 14, 1 (2013), 22–37. DOI: http://dx.doi.org/10.1080/1464536X.2011.573008

[60] Aniket Kittur, Ed H. Chi, and Bongwon Suh. 2008. Crowdsourcing user studies with Mechanical Turk. In *Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI '08*. ACM Press. DOI:http://dx.doi.org/10.1145/1357054.1357127

[61] D.A. Kolb. 1984. *Experiential Learning: Experience as the Source of Learning and Development*. Prentice Hall, Englewood Cliffs, NJ.

[62] Raph Koster. 2013. *Theory of fun for game design*. O'Reilly Media, Inc.

[63] HyunKyu Lee, Walter R. Boot, Chandramallika Basak, Michelle W. Voss, Ruchika Shaurya Prakash, Mark Neider, Kirk I. Erickson, Daniel J. Simons, Monica Fabiani, Gabriele Gratton, Kathy A. Low, and Arthur F. Kramer. 2012. Performance gains from directed training do not transfer to untrained tasks. *Acta Psychologica* 139, 1 (jan 2012), 146–158. DOI:http://dx.doi.org/10.1016/j.actpsy.2011.11.003

[64] Timothy D. Lee and Elizabeth D. Genovese. 1988. Distribution of Practice in Motor Skill Acquisition: Learning and Performance Effects Reconsidered. *Research Quarterly for Exercise and Sport* 59, 4 (dec 1988), 277–287. DOI:http://dx.doi.org/10.1080/02701367.1988.10609373

[65] Timothy D. Lee and Elizabeth D. Genovese. 1989. Distribution of Practice in Motor Skill Acquisition: Different Effects for Discrete and Continuous Tasks. *Research Quarterly for Exercise and Sport* 60, 1 (mar 1989), 59–65. DOI:http://dx.doi.org/10.1080/02701367.1989.10607414

[66] Ricardo Lopes and Rafael Bidarra. 2011. Adaptivity challenges in games and simulations: a survey. *IEEE Transactions on Computational Intelligence and AI in Games* 3, 2 (2011), 85–99. DOI:http://dx.doi.org/10.1109/tciaig.2011.2152841

[67] R.A. Magill. 1998. *Motor Learning: Concepts and Applications*. McGraw-Hill.

[68] R. Magill. 2010. *Motor Learning and Control: Concepts and Applications*. McGraw-Hill Education.

[69] Winter Mason and Siddharth Suri. 2012. Conducting behavioral research on Amazon's Mechanical Turk. *Behavior research methods* 44, 1 (2012), 1–23.

[70] RSW Masters, Jon Law, and Jon Maxwell. 2002. Implicit and explicit learning in interceptive actions. *Interceptive actions in sport: Information and movement* (2002), 126–143.

[71] Edward McAuley, Terry Duncan, and Vance V. Tammen. 1989. Psychometric Properties of the Intrinsic Motivation Inventory in a Competitive Sport Setting: A Confirmatory Factor Analysis. *Research Quarterly for Exercise and Sport* 60, 1 (mar 1989), 48–58. DOI: http://dx.doi.org/10.1080/02701367.1989.10607413

[72] Vivienne Robison McClatchy. 1925. The optimal position of a rest period in learning. *Journal of Experimental Psychology* 8, 4 (1925), 251–277. DOI: http://dx.doi.org/10.1037/h0068331

[73] Adam W Meade and S Bartholomew Craig. 2012. Identifying careless responses in survey data. *Psychological methods* 17, 3 (2012), 437.

[74] Metacritic. 2018. Super Hexagon for PC Reivews. (2018). Retrieved September 18, 2018 from https://www.metacritic.com/game/pc/super-hexagon

[75] Spero A. Metalis. 1985. Effects of Massed versus Distributed Practice on Acquisition of Video Game Skill. *Perceptual and Motor Skills* 61, 2 (oct 1985), 457–458. DOI:http://dx.doi.org/10.2466/pms.1985.61.2.457

[76] Matthew K. Miller and Regan L. Mandryk. 2016. Differentiating in-Game Frustration from at-Game Frustration using Touch Pressure. In *Proceedings of the 2016 ACM on Interactive Surfaces and Spaces - ISS '16*. ACM Press. DOI:http://dx.doi.org/10.1145/2992154.2992185

[77] Vanessa D Moss. 1995. The efficacy of massed versus distributed practice as a function of desired learning outcomes and grade level of the student. (1995).

[78] David M Neves and John R Anderson. 1981. Knowledge compilation: Mechanisms for the automatization of cognitive skills. *Cognitive skills and their acquisition* (1981), 57–84.

[79] Allen Newell and Paul S Rosenbloom. 1981. Mechanisms of skill acquisition and the law of practice. *Cognitive skills and their Acquisition* 6 (1981), 1–55.

[80] Newzoo. 2019. Global Esports Market Report 2019. (2019). Retrieved July 30, 2019 from https://newzoo.com/insights/trend-reports/newzoo-global-esports-market-report-2019-light-version/

[81] Nintendo EAD. 2006. *Wii Sports*. Nintendo.

[82] Kent L. Norman. 2011. *Assessing the Components of Skill Necessary for Playing Video Games*. Technical Report. University of Maryland. 1–12 pages.

[83] Kent L. Norman, Yu-Chi Wang, Joseph Barnet, and Reze Mahmud. 2016. *Assessing the Necessary Skill Profiles for Playing Video Games*. Technical Report. University of Maryland. 1–16 pages.

[84] Thomas Pagulayan, Randy J and Keeker, Kevin and Wixon, Dennis and Romero, Ramon L and Fuller. 2002. User-centered Design in Games. In *Human-Computer Interaction: Designing for Diverse Users and Domains*. CRC Press Boca Raton, FL, Chapter 13, 217–236.

[85] Gabriele Paolacci and Jesse Chandler. 2014. Inside the Turk: Understanding Mechanical Turk as a participant pool. *Current Directions in Psychological Science* 23, 3 (jun 2014), 184–188. DOI:http://dx.doi.org/10.1177/0963721414531598

[86] Gabriele Paolacci, Jesse Chandler, Panagiotis G Ipeirotis, and others. 2010. Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making* 5, 5 (2010), 411–419.

[87] Serge Petralito, Florian Brühlmann, Glena Iten, Elisa D Mekler, and Klaus Opwis. 2017. A Good Reason to Die: How Avatar Death and High Challenges Enable Positive Experiences. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2017)* (2017). DOI:http://dx.doi.org/10.1145/3025453.3026047

[88] R.W. Proctor, A. Dutta, and A. Ehrenstein. 1995. *Skill Acquisition and Human Performance*. SAGE Publications.

[89] Robert W. Proctor, T. Gilmour Reeve, and Daniel J. Weeks. 1990. A Triphasic Approach to the Acquisition of Response-Selection Skill. *Psychology of Learning and Motivation - Advances in Research and Theory* 26, C (1990), 207–240. DOI:http://dx.doi.org/10.1016/S0079-7421(08)60055-9

[90] Andrew K Przybylski, C Scott Rigby, and Richard M Ryan. 2010. A motivational model of video game engagement. *Review of General Psychology* 14, 2 (2010), 154–166. DOI:http://dx.doi.org/10.1037/a0019440

[91] Falko Rheinberg, Regina Vollmeyer, and Bruce D Burns. 2001. QCM: A questionnaire to assess current motivation in learning situations. *Diagnostica* 47, 2 (2001), 57–66.

[92] Richard M. Ryan, C. Scott Rigby, and Andrew Przybylski. 2006. The Motivational Pull of Video Games: A Self-Determination Theory Approach. *Motivation and Emotion* 30, 4 (nov 2006), 344–360. DOI:http://dx.doi.org/10.1007/s11031-006-9051-8

[93] Alan W. Salmoni. 1989. Motor Skill Learning. In *Human Skills*. John Wiley & Sons, 197–228.

[94] Alan W. Salmoni, Richard A. Schmidt, and Charles B. Walter. 1984. Knowledge of results and motor learning: A review and critical reappraisal. *Psychological Bulletin* 95, 3 (1984), 355–386. DOI:http://dx.doi.org/10.1037/0033-2909.95.3.355

[95] Florian Schimanke, Robert Mertens, and Oliver Vornberger. 2014. Spaced repetition learning games on mobile devices. *Interactive Technology and Smart Education* 11, 3 (sep 2014), 201–222. DOI:http://dx.doi.org/10.1108/itse-07-2014-0017

[96] Florian Schimanke, Robert Mertens, and Oliver Vornberger. 2017. Designing for Motivation: Design-Considerations for Spaced-Repetition-Based Learning Games on Mobile Devices. *International Journal on E-Learning* 16, 3 (2017), 287–311.

[97] Richard A. Schmidt and Timothy D. Lee. 2005. *Motor Control and Learning: A Behavioral Emphasis*. Human Kinetics.

[98] Richard A. Schmidt, Timothy D. Lee, C.J. Winstein, G. Wulf, and H.N. Zelaznik. 2018. *Motor Control and Learning, 6E*. Langara College.

[99] Walter Schneider. 1985. Training High-Performance Skills: Fallacies and Guidelines. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 27, 3 (jun 1985), 285–300. DOI:http://dx.doi.org/10.1177/001872088502700305

[100] Walter Schneider and Richard M. Shiffrin. 1977. Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review* 84, 1 (1977), 1–66. DOI:http://dx.doi.org/10.1037/0033-295x.84.1.1

[101] Yuri Seo. 2016. Professionalized consumption and identity transformations in the field of eSports. *Journal of Business Research* 69, 1 (jan 2016), 264–272. DOI:http://dx.doi.org/10.1016/j.jbusres.2015.07.039

[102] Wayne L Shebilske, Barry P Goettl, Kip Corrington, and Eric Anthony Day. 1999. Interlesson spacing and task-related processing during complex skill acquisition. *Journal of Experimental Psychology: Applied* 5, 4 (1999), 413–437. DOI:http://dx.doi.org/10.1037/1076-898x.5.4.413

[103] Steve Sheng, Mandy Holbrook, Ponnurangam Kumaraguru, Lorrie Faith Cranor, and Julie Downs. 2010. Who Falls for Phish?: A Demographic Analysis of Phishing Susceptibility and Effectiveness of Interventions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 373–382. DOI:http://dx.doi.org/10.1145/1753326.1753383

[104] H. A. Simon. 1956. Rational choice and the structure of the environment. *Psychological Review* 63, 2 (1956), 129–138. DOI:http://dx.doi.org/10.1037/h0042769

[105] R N Singer. 1980. *Motor Learning and Human Performance: An Application to Motor Skills and Movement Behaviors*. MacMillan.

[106] G.S. Snoddy. 1926. Learning and stability: a psychophysiological analysis of a case of motor learning with clinical applications. *Journal of Applied Psychology* 10, 1 (1926), 1–36. DOI:http://dx.doi.org/10.1037/h0075814

[107] Nicholas C. Soderstrom and Robert A. Bjork. 2015. Learning Versus Performance. *Perspectives on Psychological Science* 10, 2 (2015), 176–199. DOI:http://dx.doi.org/10.1177/1745691615569000

[108] Steam Database. 2018. Super Hexagon - AppID: 221640. (2018). Retrieved September 18, 2018 from https://steamdb.info/app/221640/graphs/

[109] StudioMDHR. 2017. *Cuphead*.

[110] Team Meat. 2010. *Super Meat Boy*.

[111] Caitlin Tenison and John R. Anderson. 2017. The impact of practice frequency on learning and retention. In *Proceedings of the 39th Annual Meeting of the Cognitive Science Society*. 1175–1180.

[112] Line E. Thomsen, Falko Weigert Petersen, Anders Drachen, and Pejman Mirza-Babaei. 2016. Identifying Onboarding Heuristics for Free-to-Play Mobile Games: A Mixed Methods Approach. In *Entertainment Computing - ICEC 2016*. Springer International Publishing, 241–246. DOI:http://dx.doi.org/10.1007/978-3-319-46100-7_24

[113] Edward Lee Thorndike. 1913. *The Original Nature of Man*. Teachers College, Columbia University.

[114] P D Tomporowski. 2003. *The Psychology of Skill: A Life-span Approach*. Praeger Publishers.

[115] u/imaunitard. 2012. r/gaming - Why not take a break? *reddit* (2012). Retrieved July 31, 2019 from https://www.reddit.com/r/gaming/comments/x6exy/why_not_take_a_break/

[116] u/Lopsidation. 2017. r/darksouls - Why are there long walks from bonfires to bosses? *reddit* (2017). Retrieved September 13, 2018 from https://www.reddit.com/r/darksouls/comments/6f8dwm/why_are_there_long_walks_from_bonfires_to_bosses/

[117] u/MetroidAndZeldaFan. 2015. r/truegaming - Do you find the lack of checkpoints in Dark Souls to be fair. *reddit* (2015). Retrieved September 15, 2018 from https://www.reddit.com/r/truegaming/comments/3k1bhp/do_you_find_the_lack_of_checkpoints_in_dark_souls/

[118] F Martijn Verhoeven and Karl M Newell. 2018. Unifying practice schedules in the timescales of motor learning and performance. *Human Movement Science* 59 (2018), 153–169. DOI:http://dx.doi.org/https://doi.org/10.1016/j.humov.2018.04.004

[119] Rodrigo Vicencio-Moreira, Regan L Mandryk, and Carl Gutwin. 2015. Now You Can Compete With Anyone: Balancing Players of Different Skill Levels in a First-Person Shooter Game. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 2255–2264. DOI:http://dx.doi.org/10.1145/2702123.2702242

[120] Regina Vollmeyer and Falko Rheinberg. 2006. Motivational Effects on Self-Regulated Learning with Different Tasks. *Educational Psychology Review* 18, 3 (oct 2006), 239–253. DOI:http://dx.doi.org/10.1007/s10648-006-9017-0

[121] Michael G Wagner. 2006. On the Scientific Relevance of eSports.. In *International conference on internet computing*. 437–442.

[122] Alan T Welford. 1976. Skilled performance: Perceptual and motor skills. (1976).

[123] Daniel B. Willingham and Kelly Goedert-Eschmann. 1999. The Relation Between Implicit and Explicit Learning: Evidence for Parallel Development. *Psychological Science* 10, 6 (1999), 531–534. DOI:http://dx.doi.org/10.1111/1467-9280.00201

[124] A. Wingfield. 1979. *Human learning and memory: an introduction*. Harper & Row.

[125] Brian M. Winn. 2009. The Design, Play, and Experience Framework. In *Handbook of Research on Effective Electronic Gaming in Education*. IGI Global, 1010–1024. DOI:http://dx.doi.org/10.4018/9781599048086.ch058

[126] Carolee J. Winstein and Richard A. Schmidt. 1989. Sensorimotor Feedback. In *Human Skills*. John Wiley & Sons, 17–47.

[127] Melody Wiseheart, Annalise A. D'Souza, and Jacey Chae. 2017. Lack of spacing effects during piano learning. *PLoS ONE* 12, 8 (2017), 1–13. DOI:http://dx.doi.org/10.1371/journal.pone.0182986

[128] Bob G. Witmer and Michael J. Singer. 1998. Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments* 7, 3 (jun 1998), 225–240. DOI:http://dx.doi.org/10.1162/105474698565686