If at First You Don't Succeed: Helping Players Make Progress in Games with Breaks and Checkpoints

COLBY JOHANSON, University of Saskatchewan, Canada
BRANDON PILLER, University of Saskatchewan, Canada
CARL GUTWIN, University of Saskatchewan, Canada
REGAN L. MANDRYK, University of Saskatchewan, Canada

Developing skill and overcoming in-game challenges is of great interest to both players and game designers. Players can improve through repetition, but sometimes practice does not lead to improvement and progress stalls. It would be useful if designers could help players make progress without compromising their long-term skill development. We carried out a study to investigate how two techniques—checkpoints and breaks—affect in-game progress and player skill. Checkpoints allow multiple attempts at a challenge without having to repeat earlier sections; this aids progress, but could potentially hinder skill development. Second, breaks in gameplay have been shown to accelerate skill development, but their effectiveness is unknown when the breaks are integrated into the game's design. Our study evaluated the effects of game-integrated breaks and checkpoints on players' in-game progress (when the techniques were present) as well as two test sessions (with all techniques removed). Our results showed that both checkpoints and breaks aid progress (combining both had the largest effect) and that neither technique reduced performance in the transfer task, suggesting that skill development was not hindered. Our work provides evidence that checkpoints and breaks are valuable techniques that can assist both player progress and skill.

CCS Concepts: • **Human-centered computing → Human computer interaction (HCI)**; **HCI theory, concepts and models**;

Additional Key Words and Phrases: practice, spaced practice, checkpoints, video games, skill development

ACM Reference Format:
Colby Johanson, Brandon Piller, Carl Gutwin, and Regan L. Mandryk. 2023. If at First You Don't Succeed: Helping Players Make Progress in Games with Breaks and Checkpoints. *Proc. ACM Hum.-Comput. Interact.* 7, CHI PLAY, Article 387 (November 2023), 27 pages. https://doi.org/10.1145/3611033

1 INTRODUCTION

Many players want to play games that challenge them [26] — overcoming the challenges within a game is satisfying [28] and motivates players to continue playing [40]. Players and game designers therefore have a strong interest in understanding how players can improve their skills to overcome in-game challenges. Simple repetition (i.e., practice) is the main way this happens: repeatedly attempting a challenge allows players to learn from their mistakes and improve at the game [18, 23, 28, 67]. Although often a viable strategy, there are situations in which simple repetition is ineffective. In particular, players may feel that their rate of improvement is not fast enough, or

Authors' addresses: Colby Johanson, colby.johanson@usask.ca, University of Saskatchewan, Saskatoon, Saskatchewan, Canada; Brandon Piller, brandon.piller@usask.ca, University of Saskatchewan, Saskatoon, Saskatchewan, Canada; Carl Gutwin, gutwin@cs.usask.ca, University of Saskatchewan, Saskatoon, Saskatchewan, Canada; Regan L. Mandryk, regan@cs.usask.ca, University of Saskatchewan, Saskatoon, Saskatchewan, Canada.

2573-0142/2023/11-ART387 $15.00
https://doi.org/10.1145/3611033

they may reach a plateau and be unable to make progress, sometimes causing them to abandon the game [8].

Game designers can try and address this problem in two ways. First, they could focus on skill development, under the assumption that as skill improves, the player will also be able to make progress in the game. There are many potential strategies for improving skill development, many of which have been explored in the domain of perceptual-motor learning (a research area that encompasses many of the types of skills used in digital games) (e.g., [12, 16, 47, 48]).

Second, designers could focus on the player's in-game progress, and add mechanisms or assists that move the player along in the game. This strategy can ensure continued progress — but if the assist is too extreme or too artificial, it could compromise the player's longer-term skill development. For example, a powerup that boosts a struggling player over a challenge could prevent them ever learning the skills needed to overcome the challenge on their own.

In the first approach, there are several techniques that have been shown to improve skill development in previous work [47], including adjusting the variety of what the learner practices (e.g., practicing variations of the task to better handle novel scenarios [46]), introducing part-task practice (e.g., focusing on specific components of a skill by breaking a complex skill into components that can be practiced separately [17]), or adjusting the spacing of the practice sessions. Spaced practice adds short ***breaks*** to the practice session — that is, forced delays that interrupt practice. Taking breaks in games has been shown to improve performance over playing continuously [25, 38], but previous studies presented breaks on a fixed schedule rather than integrating them into the existing mechanics and presentation style of the game; this artificial presentation would not be acceptable in a real-world game. Breaks and pauses do occur in many games — e.g., as spawn timers, cut scenes, or mini-games — so it is possible that breaks intended for skill development could be better integrated into the game design. It is currently unknown, however, whether game-integrated breaks will continue to be effective in improving skill development.

In the second approach, there are many game mechanisms that could be used simply to assist progress, such as powerups, dynamic difficulty adjustment, or ***checkpoints***. Checkpoints create new restart points within a game that allow a player to save their progress — when players fail at a challenge, they restart at the most recent checkpoint [19] and can immediately attempt the challenge again. Checkpoints can aid player progress, but their effect on skill development is unclear. Checkpoints could hinder skill development by allowing players to avoid practicing skills in less challenging contexts, or by artificially moving players forward to a part of the game that is too difficult. However, checkpoints could potentially also help with skill development by allowing players to focus on the part of the game they struggle with (i.e., providing part-task practice), or by exposing players to a wider variety of game scenarios and skill demands (i.e., adjusting the variety of practice).

It is possible that combining the two approaches — game-integrated breaks and checkpoints — will result in even more progress being made in the game or affect skill development more than either modification on its own. Checkpoints on their own might prompt players to take risks, but the break could reintroduce a consequence by enforcing a pause after a failure. This might give players time to reconsider what they are doing.

To investigate these questions, we carried out a between-participants study in which people played a side-scrolling platform game. Participants were divided into groups who saw different versions of the game with different combinations of breaks and checkpoints, in a 2x2 design. Game-integrated breaks were implemented as variable pauses of up to 10 seconds after dying, and checkpoints were implemented as an automatic save mechanism using several fixed points in each game level. Participants played their version of the game for 20 minutes and then completed an additional 10-minute transfer session in which the game had neither checkpoints nor breaks. We additionally invited participants back after seven days to complete a retention session that had players play the game for an additional 10 minutes, using the same levels as in training but without breaks or checkpoints. The study assessed in-game progress by measuring how many levels the player completed (with the technique present); and because skill development can only be inferred by seeing changes to performance over time [47, 50], we assessed development through the player's performance in the transfer and retention sessions.

For this study we had five research questions:

- **RQ1:** For breaks, would pauses that are tightly integrated into the game design still improve a player's skill (i.e., translate into improved progress within the game)?
- **RQ2:** For checkpoints, would saving a player's progress help them make progress within the game?
- **RQ3 and RQ4:** Would having played with breaks (RQ3) or checkpoints (RQ4) have any effect on a player's skill development?
- **RQ5:** Would using both techniques have a larger effect on progress or skill development than using each individually?

For both techniques, the study also considered whether the technique detracted from or changed the play experience: for example, it is possible that players dislike forced breaks that prevent them from immediately re-attempting a challenge, and that players will view checkpoints as "false progress" that takes away from their sense of accomplishment.

Our study showed positive results for both game-integrated breaks and checkpoints as techniques for supporting player progress without hindering skill development. First, participants who had breaks in their game completed about three more levels than the baseline condition (which had no breaks and no checkpoints), and there were no reductions in performance in the transfer and retention tasks. This result provides evidence that breaks are still effective even when integrated into the game design. Second, checkpoints also allowed players to complete about three more levels than the baseline group, and performance in retention and transfer tests was again unaffected. This result indicates that a checkpoint's positive effects on progress do not come at the cost of reduced skill development. Third, players who had both checkpoints and breaks completed nearly eight more levels than the baseline, and were also no worse during transfer and retention, suggesting that combining the two techniques may provide added value.

In addition to our primary measure of progress, we also checked the effects of checkpoints and breaks on player deaths, and analysed the game logs to look for two patterns of inhibited progress: stalls (in which players get stuck at one part of the game), and regressions (in which players do worse than in past attempts). We found that player deaths were affected by the presence of breaks (likely due to players being more cautious because of the after-death pause). We also found that both checkpoints and breaks helped players overcome both types of inhibited progress: regressions occurred fewer times, and players were stalled for less time.

Finally, our player experience measures found mixed results: there were no differences in player experience for people who had checkpoints, but we did find differences in subjective flow, curiosity, and meaning for groups who had breaks. Players who started with breaks reported improvements in these measures once they moved to the transfer and retention sessions (which had no breaks). Subjective preference questions showed that none of the participants liked the breaks — but several people noted that they used the time to reconsider what they were doing in the game, suggesting that the breaks may have been having the intended effect.

Overall, our findings provide strong evidence that both techniques — breaks and checkpoints — can be successfully used to support players' progress in digital games without hindering skill development. Even when integrated into the game mechanics and presentation, breaks successfully improved progress; and even though players stated that they disliked the breaks, there were no negative effects on game-experience measures. Checkpoints allowed players to progress further without any negative effect on skill development, and were liked by participants. Our results give game designers a new understanding of how they can support a player's progress and help them improve at the game using techniques that can be woven into existing game mechanics and presentation without substantially compromising the play experience.

2 RELATED WORK

2.1 Spaced Practice in Games

Spaced (or distributed) practice is the concept of scheduling periods of rest to break up periods of work within a training session [47]. Compared to continuous (or massed practice), adding breaks generally results in strong gains to short-term performance and slightly weaker gains to long-term performance (i.e., learning), as measured via transfer tests (testing participants on a very similar but different task) or retention tests (re-testing participants after some period of time, such as a day or a week) [12, 30, 47].

Spaced practice has been found to benefit player performance in games over four experiments [25, 32, 38, 49] and two analyses of data-sets [51, 52]. This work demonstrates that benefits of spaced practice can be had within experiments lasting less than an hour [25, 32, 38], to ten hours [49], or over periods of weeks or longer [51, 52]. It has been observed in games as simple as *Breakout* [32] or as complex as *Destiny* [51].

Because breaks improve performance, introducing breaks could potentially help players make progress and overcome a game's challenges. However, no past work has provided guidance on how to integrate spaced practice into a game. Past work introduces breaks on a fixed schedule and involves rests of two minutes or longer (e.g., [25, 38]). However, spaced practice is defined in relative terms rather than absolute terms [47] — there is no one accepted schedule for the timing of the rest periods relative to the work periods, the times for the rest and work period need not be fixed, and trials might be used to determine spacing instead of fixed time periods [47, 63]. Additionally, the rest period does not need to be as long as two minutes; breaks as short as 15 seconds have been found to be beneficial [22, 58]

2.2 Checkpoints

Very little prior work in digital games has studied the effects of checkpoints or save states on performance, learning, or progress. One paper examined the relationship between the frequency of saving progress (with one implementation using checkpoints) and several player experience measures, and found that the frequency of deaths was related to perceived challenge [10]. Because of the relative lack of prior work, however, the remainder of this subsection considers ways that checkpoints could affect a player's experience of the game and how this changed experience might affect performance and learning.

2.2.1 Skill-Challenge Balance. Practice is often an effective method of skill development because a game's difficulty increases in line with their progress [18, 23, 28, 56, 67]. To facilitate continued learning, challenges should be challenging but within or at the edge [65] of a player's abilities. Players encountering challenges of this difficulty feel that they can overcome them [27] and are motivated to do so [18]. If checkpoints help players progress through the game they will be more quickly taking on tough challenges — challenges they might not be prepared for. This could negatively affect learning.

2.2.2 Practice Conditions. Checkpoints may alter practice conditions in two ways. First, if checkpoints help players make progress then they may affect the *variability of practice* — they become exposed to more levels and more variations of the game's mechanics. Second, checkpoints may introduce *part-task practice* by allowing players to attempt a challenging obstacle without needing to consider the ones that come before — they can focus on overcoming an obstacle they struggle with.

Increased variability in practice is thought to promote motor learning by strengthening a learner's ability to cope with novel conditions [47, 48]. These benefits are subject to factors such as the nature of the task (e.g., benefits of variable practice are found most commonly with simple tasks) and the expertise of the learner (e.g., increased variability may be most beneficial for early learning) [68]. Increased variety typically increases errors [31], leading to a short-term drop in performance which would be undesirable in games. In games, players would experience this loss of performance alongside making progress through the game, so they may struggle more when taking on harder obstacles. The previously mentioned data analysis [51] of the online game, *Destiny*, looked at variability in terms of the propensity for "social play" and playing style, but did not find that increased variability enhanced skill acquisition.

Part-task practice is thought to be beneficial for complex skills [16, 31], as it allows a learner to focus on aspects that need improvement [34]. When players are repeatedly going up against a challenging obstacle, checkpoints may let them focus on and refine a subset of skills. Past work has found part-task practice to be effective for skill learning in digital gaming contexts [17, 21].

3 MODIFYING GAME PRACTICE WITH BREAKS AND CHECKPOINTS

Modifying in-game practice by introducing breaks and checkpoints is something that some games already do. Sometimes this is done explicitly, while other times it is an implicit consequence of the design of the game or its systems. In this section, we discuss how these systems are presented within commercial games.

3.1 Breaks

In digital games, spaced practice (i.e., breaks) is often included incidentally. For example, breaks may exist in games due to technical limitations — loading screens, in particular, are commonly found in many games because of the delays encountered when loading content [43] (although these breaks may become shorter or less common as hardware improves). Other examples include players waiting for others to connect in a multiplayer game, breaks from the game's core mechanics in the form of mini-games [38], or menu systems for activities such as inventory management. Breaks also may be implicitly introduced in games simply by having enemies spaced apart from one another, with the travel time between battles acting as a break.

Although previous studies provide evidence for the benefits of breaks, there is little guidance from past work in terms of integrating them into the game. Past research presented breaks to players on fixed schedules, which could interrupt a player in the middle of play. This could be difficult to integrate into the design of many games. Less of an interruption would be preferred and could be accomplished if the break was integrated into the events of the game. Death is a common game mechanic and one where, incidentally, breaks can already be found (e.g., when waiting to respawn in an online game). Considering that breaks have been found to improve performance, this may be an ideal time to provide a break — players already doing well will receive fewer breaks whereas players who are struggling receive more breaks, giving them more opportunities to benefit from spaced practice.

If breaks *are* integrated into the game explicitly, then it is usually as a reminder to take a break. Many games prompt players to take a break after some time has passed. While this is not done for the benefits of spaced practice (it is to reduce the likelihood of seizures and repetitive strain injuries [36]), it is a common way that players are presented with the idea of taking a break. A glance at online discussions quickly reveals players who mock the suggestion or go out of their way to keep playing in response [61]. In general, the idea of taking breaks to get better goes against many players' desires or intuitions — those who might believe that if they just keep trying, they will eventually succeed. However, players do accept breaks in certain contexts, for example, the need to wait is common within multiplayer games as players wait for a new round to start, or wait to be connected to other players.

3.2 Checkpoints

In contrast to breaks, checkpoints support a player's desire to get back to a challenge right away. Checkpoints save a player's progress so that players can start playing again from the checkpoint when they fail [19]. These saves are typically strategically timed, such as before difficult segments or only when specific conditions are met. Checkpoints are generally accepted by players, although they occasionally aggravate players when saving occurs at an inconvenient time [29].

An alternative approach to checkpoints is allowing players to choose when to save the game via save states. When taken to the extreme, some players consider it "cheating" (e.g., [60, 62]), as players can continually save the game and prevent any amount of progress from being lost. Players opposed to this argue that players doing this will not learn the game and will rely on the aid [60, 66]. Other players point out that strategic use of saving can help with learning because it can allow for quick trials of different approaches [62] and can help one learn difficult segments of the game [60], or specific enemy patterns [66].

Mechanisms that allow players to save their progress have become more prevalent over time. In the original *Super Mario Bros.* (released in 1985), players must start over from the beginning of the level if they die. With *Super Mario World* (released in 1990), each level had a checkpoint at the midpoint. Many modern titles save progress very frequently or design their levels to be small enough to not require checkpoints. In *Super Meat Boy* (released in 2010) the levels are designed to be short enough to not require checkpoints and in *Celeste* (released in 2018) each level fits on-screen without scrolling and the game is saved at each screen transition.

4 METHODS

4.1 The Game

We created a bespoke 2D side-scrolling platform game for our study (Figure 1), similar to a game used in past work on spaced practice [38]. The game was inspired by games such as *Super Mario World* [37], *SpeedRunners* [13], and *Super Meat Boy* [53]. Our aim was to create a set of mechanics that were easy to understand, but with enough nuance that they would take time to master. While the platforming genre would be familiar to players, our specific implementation of the mechanics and the design of our game levels would not.

The player controls a lumberjack character, moving horizontally with the arrow keys, jumping with the space bar, and pressing and holding "E" to swing with a grapple hook. Horizontal movement includes acceleration and deceleration, and players must anticipate this to avoid overshooting. Pressing the jump button allowed variable upward acceleration depending on the duration of the press, so players must time both the start of a jump as well as the duration. The jump button also allows wall-jumping. Pressing and holding "E" when a grapple location is in range initiates a swing action; the player holds the button down until they want to let go of the rope, and the speed and direction of the swing can be adjusted mid-swing with the arrow keys. Players can also exit the swing with a jump action.

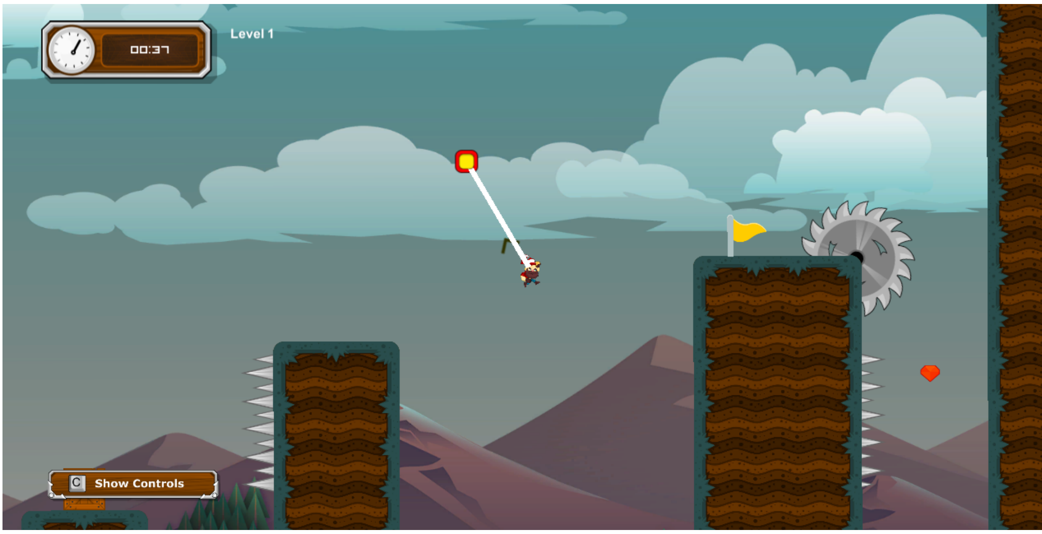


Fig. 1. The game used in the experiment. Pictured is the player using the grapple hook to swing toward the checkpoint (the yellow flag).

The game mechanics were introduced to players through a 47-second video demonstration and through three in-game tutorial levels; these included in-game sign posts that indicated what keys needed to be pressed and were designed so that players were required to perform each skill multiple times to continue. After the tutorials, the game levels increased in difficulty as the player made progress by gradually presenting the player with more frequent or more complex obstacles. If players ever became stuck and were unable to finish a level after 3 minutes, they were given a button that they could click to skip the level.

The game was developed in Unity and presented to participants on a website via WebGL.

4.1.1 Experimental Factors. Participants of our experiment were assigned to one of four treatment groups, formed by crossing two factors (checkpoints and breaks) in a 2x2 design.

Checkpoints: If participants experienced the game with checkpoints, flags were placed throughout the level that saved the player's progress whenever the flag was passed. Flags were positioned between groups of obstacles within the levels. Then, if they died within the level, they would start over at the checkpoint rather than at the beginning of the level. A screenshot showing a checkpoint flag is shown in Figure 2.

Game-Integrated Breaks: If participants experienced the game with game-integrated breaks, then upon death they were made to wait up to ten seconds before another attempt, with an onscreen countdown timer showing them the time remaining. The length of the break was equal to the time they were alive before their death (up to a maximum break of ten seconds) — more time alive meaning a longer break. Ten seconds was chosen as a maximum based on feedback from pilot testing (and to avoid situations where a player spends as much time pausing as actually playing). If players did not receive breaks, they still were made to wait for one second before their next attempt, to prevent accidental deaths from unintentional inputs. A screenshot showing the countdown timer during a break is shown in Figure 3.

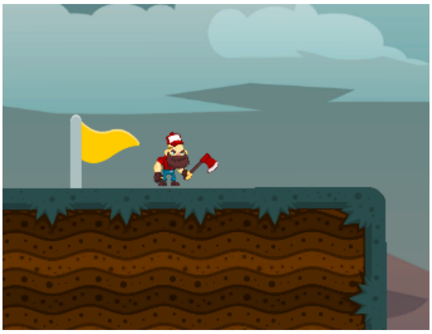


Fig. 2. Screenshot of the game showing a checkpoint flag.

Fig. 3. Screenshot of the game showing the countdown timer for a game-integrated break.

4.2 Procedure

After giving consent, participants were assigned to one of the four treatment groups. Participants then watched a video demonstrating how to play the game before responding to a questionnaire relating to their current motivation toward the task and how they viewed their ability to complete it. Participants then played the game, starting with the tutorial levels. They played the game for a 20-minute *training* session, in which they played the game with the techniques that were specified by their group in the 2x2 design.

After this training session, participants responded to several questionnaires relating to their subjective experience of the task, and a questionnaire relating to demographics. They then played the game again for a 10-minute *transfer* session, in which they played a new set of levels without checkpoints or game-integrated breaks. After, they responded to further questionnaires relating to their subjective experience and two more questionnaires relating to individual differences.

Participants who completed the training session fully were then invited back one week later to complete a *retention* task. Each participant renewed consent, and played the game for 10 minutes, with the same levels and no checkpoints or game-integrated breaks. They then responded to questionnaires assessing their experience and could provide feedback once again.

The game and all questionnaires were presented to participants via a website built using an existing web framework designed to aid the creation of online studies [24].

4.3 Measures

We used a combination of questionnaires and data generated from in-game actions to measure differences between participants and their experience with the game.

4.3.1 Individual Differences. To account for individual differences that could alter a participant's in-game performance or subjective experience of the game, we used the following measures:

Gaming and Platforming Familiarity: We expected that prior experience playing platforming games might affect performance and subjective experience of the game. Therefore, we asked six questions, each presented as a slider from 1 ("Not at all") to 100 ("Very familiar", or "Gamer"). These questions were: *"How much do you self-identify as a gamer on the following scale?"*, *"How familiar are you with side-scrolling platform games?"*, *"How familiar are you with Super Mario games?"*, *"How familiar are you with the game 'Super Meat Boy'?"*, *"How familiar are you with the game 'Speedrunners'?"*, and *"How familiar are you with the game 'Celeste'?"*. The measure was the mean of responses to all six questions. These games were chosen because they are popular and because they feature one or more of the mechanics used within our game. In particular, the grapple mechanic in our game was directly inspired by and modelled on *Speedrunner*'s grappling hook, and during development, our game's movement controls were compared against *Super Meat Boy* in an attempt to replicate the feel of that game's controls for greater external validity.

Attentional Control: We expected that a participant's ability to give our game their complete attention might affect their performance and subjective experience of the game. We therefore used Derryberry and Reed's [11] Attentional Control Scale (ACS), which provides a self-report measure of *attentional control* — which relates to an individual's ability to focus on a specific task and shift their attention away from potential threats.

Current Motivation: We thought that a participant's initial motivation upon beginning the game could affect their ability to learn and improve at the game. Therefore, we used Rheinberg et al.'s [42] Questionnaire on Current Motivation (QCM) as a self-report measure of *task-related anxiety*, *probability of success* at the task, *interest* in the task, and *perceived challenge* of the task.

Achievement Orientation: We thought participants who were more competitive or interested in winning might put more effort into learning the game. Therefore, we used Gill and Deeter's [20] Sport Orientation Questionnaire (SOQ) as a self-report measure of *competitiveness* (i.e., the overall desire to meet a standard of excellence or compare favourably to competitors), *win orientation* (i.e., the importance of outperforming the competition), and *goal orientation* (i.e., the importance of achieving specific performance goals).

Initial Performance: Because participants who played our game would have different levels of prior experience playing platform games, we measured their initial performance in the three tutorial levels (where no Checkpoints or Game-Integrated Breaks were included) to further account for individual differences in prior platform game experience.

Total Training Time: Participants would end up actively engaged with the game for different lengths of time in training due to how long they ended up waiting. Therefore, we calculated the total time spent training by subtracting the wait time.

4.3.2 Outcome Measures. Our dependent outcome measures were chosen based on the goal of measuring in-game progress and exploring how players' subjective experience of the game changed due to training with checkpoints or game-integrated breaks.

Progress: As an objective measure of player progress within the game, we logged the *number of levels completed* in each session (training, transfer, and retention). To understand how players' in-game behaviour changed in the presence of checkpoints or game-integrated breaks, we logged and used the *number of deaths* in each session as a dependent measure.

We used two published questionnaires to measure players' subjective experience of the game.

Flow: We used Vollmeyer and Rheinberg's [64] Flow Scale Short (FKS), which measures all aspects of *flow* (i.e., challenge-skill balance, merging of action and awareness, unambiguous feedback, concentration on the task, time transformation, and fluency of action), combined into a single measure. Situations that lead to the joyful experience of flow can also lead to *worry*, and so the questionnaire measures this as well.

Player Experience: We used Vanden Abeele et al.'s [1] Player Experience Inventory (PXI) to assess several different aspects of player experience. The PXI measures player experience in terms of *functional consequences* (i.e., arising directly due to the game's design), as well as *psychosocial consequences* (i.e., second-order emotional experiences). Functional consequences consist of *ease of control*, *progress feedback*, *audiovisual appeal*, *goals and rules*, and *challenge*. Psychosocial consequences consist of *mastery*, *curiosity*, *immersion*, *autonomy*, and *meaning*. We measured the psychosocial consequences as well as the functional consequence of *challenge* after each session.

Written Responses: We also asked participants to respond to several open-ended questions, once after training and again after the test session. The questions presented after training changed depending on which version of the game they played.

If the participant played the version of the game in which checkpoints were present, then we asked them the following questions:

- "*How did you feel about the checkpoints in the game? Were there too many or too few?*"
- "*Do you think that the checkpoints made the game easier?*"
- "*Did knowing that you would re-start at a checkpoint affect how you played the game?*"

If the participant played the version of the game without checkpoints, then they were asked:

- "*When you died you had to walk back to that point in the level. Did you find this process easier than the part of the level where you died?*"
- "*Was there anything in particular you focused on or thought about while you were walking back?*"
- "*Did knowing you would need to walk back if you died affect how you approached the game?*"
- "*How did you generally feel about starting each level from the beginning after each death? Do you have any further comments on this?*"

If they played the version of the game where they had to wait to respawn, we asked them:

- "*How did you feel about waiting to play the game after each death? Was the time too short or too long?*"
- "*When waiting to play, what did you do with your time?*"
- "*Did the possibility that you would need to wait to try again change how you approached the game?*"

If they played the version of the game where they restarted instantly, then we asked:

- "*Did you take any breaks while playing? (Either intentional or unintentional.)*"
- "*Do you think you would have benefited from taking a break while you played?*"
- "*Do you think the ability to attempt the level again immediately affected how you approached it?*"

After the transfer session, in which checkpoints and game-integrated breaks were not present, we asked participants to compare the two versions of the game: "*Of the two versions of the game you played, which did you prefer?*", and "*Of the two versions of the game you played, did you find one of them to be more difficult?*". Finally, we gave them one last opportunity to provide general comments.

4.4 Recruitment

Because our game was designed to be somewhat challenging and we wanted participants to be able to make progress in the game, we were interested in recruiting participants who were already experienced gamers. All participants were recruited on Amazon's Mechanical Turk, so we first recruited participants to complete a task that simply had them respond to the question "How much experience do you have with playing side-scrolling platform games?" by dragging a slider to a value between 1 ("None") and 100 ("A great deal"). Participants who specified a value of 60 or higher were then assigned a "qualification" that would allow them to see the invite to the experiment.

Because we expected that self-report measures of experience might not be the best predictor of performance in our game, we advertised our task as involving as little as 5 minutes of gameplay or as much as 30. Participants would then be given a maximum of 5 minutes to complete the first three levels which served as an in-game tutorial to introduce the game's mechanics. If they could not complete the levels in that time, they were redirected to the end of the experiment and paid $1.75

USD. If they completed the levels within 5 minutes, the game would continue uninterrupted for the full 20 minutes, and would further go on to complete the full study, at which point participants were compensated with an additional $8.50 USD.

Participants were assigned to whichever group had the fewest participants at the time they gave consent. All participants were at least 18 years of age and had an average approval rating of at least 97% and more than 500 HITs (human intelligence tasks) completed. The design of this study received ethical approval from the behavioural research ethics board of the first author's university.

4.5 Participants

We estimated the participant count by using an a-priori power analysis in G*Power [14] with the following parameters: .25 effect size, alpha of .05, power of .80, numerator df of 1, 4 groups, and 6 covariates. The power analysis estimated that 128 participants would be required. A total of 190 participants completed our study. Of these, we filtered out a total of 40 participants: 8 due to attempting the tutorial levels more than once, 15 due to having too low of a framerate to properly play the game (<30 frames per second), 4 due to the age they entered (<18 or >90 years), 5 due to not completing any levels beyond the three tutorial levels, and 8 due to not completing any of the levels on the transfer test. This left us with 150 participants, with an average age of 33.5 years (min=18, max=64, SD=7.79). 101 identified as men, 48 identified as women, and 1 identified as non-binary. 38 played without checkpoints or breaks, 37 played with checkpoints but no breaks, 38 played with checkpoints and breaks, and 37 played with breaks but no checkpoints.

After one week, we invited back the 150 participants who were not filtered out. Of these, 117 completed the retention task: 28 had played without checkpoints or breaks, 29 had played with checkpoints but no breaks, 30 had played with checkpoints and breaks, and 30 had played with breaks but no checkpoints.

Our participants were experienced gamers and identified as such (81.1 out of 100; SD=21.1). They had quite a bit of familiarity with side-scrolling platform games (89.2 out of 100; SD=13.7), as well as with *Super Mario* games (93.1 out of 100; SD=11.2). However, they were less familiar with *Super Meat Boy* (47.5 out of 100; SD=38.6), *Speedrunners* (37.1 out of 100; SD=35.8), and *Celeste* (36.3 out of 100; SD=36.4).

4.6 Data Analyses

The data and associated results include confirmatory analyses and three different kinds of separate additional analyses.

4.6.1 Confirmatory Analyses. To verify that spaced practice positively affected performance in our platforming game, we computed individual ANCOVAs (analysis of covariance) for each outcome measure. Because of our 2x2 design, we had two between-subject factors: Breaks and Checkpoints.

Attentional Control was used as a covariate for every ANCOVA due to there being a significant difference between the groups ($F_{3,146} = 3.39$, $p = .020$), as evaluated by a one-way ANOVA with the trait measure as the dependent variable (all other trait measures were evaluated in the same way, but were not significant). For the progress measures (levels completed and total deaths), Total Training Time was used as a covariate to account for the differing amount of time actively playing the game, and Tutorial Completion Time was used as a covariate to account for individual differences in initial performance at the game. Further covariate selection was made based on whether each correlated with the specific outcome measure being tested and are reported alongside the results of the ANCOVAs. These covariates are presented in Section 5.4.

Jamovi [55] was used for all quantitative analyses. Alpha was set at .05. All pairwise comparisons used the estimated marginal means and Bonferroni corrections. An ANCOVA was chosen because when group sizes are equal, it is robust to violations of normality and homogeneity of variance [15]. Our group sizes were similar (see Section 4.5); however, we also inspected the data to ensure normal distributions and homogeneity of variance before proceeding. An ANCOVA is not robust to violations of independence [15]; however, our between-subjects experiment design ensures that observations across groups are independent.

4.6.2 Additional Analyses. Our second set of results are are not based on specific hypotheses. In addition to the quantitative results for our subjective measures, we also analyzed written responses to open-ended questions, and visualizations generated by processing game logs.

For the qualitative measures of subjective experience, we asked participants to complete the Flow Scale Short (FSS) [64] and the Player Experience Inventory (PXI) [1] after each session (Training, Transfer, and Retention). Because the PXI measures functional consequences (i.e., consequences that result from the design of the game) as well as psychosocial consequences (i.e., emotional experiences that result from playing the game), we chose to measure the psychosocial consequences after every session while also including the functional consequence of challenge, as this could conceivably change due to Checkpoints, Breaks, or the slightly different levels played during the transfer session. For all repeated measures (Challenge, Mastery, Curiosity, Immersion, Autonomy, Meaning, Flow, and Worry), we used separate repeated-measure ANCOVAs, with Session as a within-subjects factor, and Breaks and Checkpoints as between-subjects factors. Attentional Control was included as a covariate (for the same reason as the confirmatory analyses). All post-hoc tests used Bonferroni corrections. Because our analyses are exploratory, we focus on significant effects only.

With the open-ended responses, we explored players' perceptions of game-integrated breaks and checkpoints. The responses were analyzed by a thematic analysis [6] performed by two of the authors. The two authors worked together to generate initial codes and group them into themes and then each author worked independently, coding all of the responses. After, the two authors met up to discuss and resolve all discrepancies. For brevity, instead of reporting the results of a full qualitative analysis, we highlight key findings that help to explain our results as well as highlight how players feel about these two mechanics. Any reported percentages are the percentage of responses that reflects the given theme unless otherwise stated.

Participants were always asked for written responses after training and again after testing. It is possible that biases related to recall or peak-end experiences may have affected their responses; however, these biases would have influenced each experimental condition similarly, and thus we do not expect to see systematic differences in qualitative responses based on which version of the game was played.

To better understand the ways that players made progress as they played the game, we generated visualizations from logs of their progress. Throughout the levels, we placed invisible triggers that we used to track progress throughout the level (shown in Figure 6). These were placed before and after obstacles where players would likely die. We also used the checkpoints to trigger a log of the player's progress; if the player was playing the version of the game without checkpoints, this trigger would still occur, but the checkpoint itself would be invisible and deactivated.

5 CONFIRMATORY RESULTS

5.1 Were Game-Integrated Breaks Beneficial? (RQ1 and RQ3)

We hypothesized that the short game-integrated breaks would positively affect the progress made in training and that skill development — progress in retention and transfer — would be unaffected. During training, we found a significant main effect of Breaks on the adjusted number of levels that players completed (see Table 1), after accounting for individual differences — players were able to

Measure	Session	df	Checkpoints				Breaks				Checkpoints*Breaks			
			F	p	η_p^2	Δ	F	p	η_p^2	Δ	F	p	η_p^2	Δ
Levels Completed	Training	1, 139	**38.8**	**<.001**	**.218**	**3.98 ± 0.64**	**5.5**	**.020**	**.038**	**3.74 ± 1.59**	3.0	.084	.021	7.72 ± 2.09
	Transfer Test	1, 141	0.4	.533	.003	0.36 ± 0.57	3.6	.060	.025	2.69 ± 1.42	2.9	.092	.020	3.05 ± 1.84
	Retention Test	1, 110	0.4	.535	.004	-0.39 ± 0.63	1.5	.224	.014	1.88 ± 1.54	1.07	.303	.010	1.49 ± 2.00
Death Count	Training	1, 142	1.1	.293	.008	2.46 ± 2.33	**78.3**	**<.001**	**.209**	**-51.4 ± 5.81**	**84.2**	**<.001**	**.372**	**-48.9 ± 7.55**
	Transfer Test	1, 143	**4.9**	**.028**	**.033**	**-3.57 ± 1.61**	**11.8**	**<.001**	**.076**	**-13.7 ± 3.98**	2.7	.105	.018	-17.3 ± 5.19
	Retention Test	1, 109	**6.2**	**.014**	**.053**	**-4.54 ± 1.82**	2.0	.157	.018	-6.52 ± 4.50	0.4	.542	.003	-11.0 ± 5.82

Table 1. Results of statistical analyses for our measures of in-game progress and deaths, showing the main effects of Checkpoints and Breaks, as well as interaction effects between Checkpoints and Breaks. Each row contains the results of a separate ANCOVA. The Δ columns show the mean difference between having Checkpoints or Breaks (or the combination) and not having them, with ± indicating standard error.

Fig. 4. Estimated marginal means of levels completed and death count, from the ANCOVAs. Error bars are standard error.

complete more levels with Breaks. We similarly found a significant main effect of Breaks on the adjusted number of deaths.

These breaks did provide players with an opportunity to rest during training. Participants who were assigned to train with the version of the game that included breaks waited an average of 274 seconds throughout training (SD=68.4, Min=124, Max=451), compared to 61.2 seconds for those who played the game with a fixed 1-second wait after each death (SD=18.8, Min=29, Max=108).

During the immediate transfer test, in which we had all participants play the game without Breaks, with a new set of levels, we did not find a main effect of Breaks on the adjusted levels completed (see Table 1). We did, however, find a main effect of Breaks on the adjusted number of deaths — training with breaks resulted in fewer deaths.

During the delayed retention test, we again had all participants play the game without Breaks, but this time with the same levels as in training (except the early tutorial levels). We found no main effect of Breaks on adjusted levels completed (see Table 1). We also did not find a main effect of Breaks on the adjusted number of deaths (see Table 1).

5.2 Were There Any Drawbacks to Adding Checkpoints? (RQ2 and RQ4)

We hypothesized that checkpoints would help players make progress in training, and not affect skill development — progress in transfer and retention. In training, we found a significant main effect of Checkpoints on the adjusted number of levels completed, accounting for individual differences (see Table 1) — players completed significantly more levels with Checkpoints. There was no main effect of Checkpoints on the adjusted number of times players died.

During the transfer test, in which all participants played the game without Checkpoints on a new set of levels, we did not find a main effect of Checkpoints on levels completed, but we did

	Training						Transfer						Retention					
	Levels Completed			Death Count			Levels Completed			Death Count			Levels Completed			Death Count		
Covariate	F	p	η_p^2	F	p	η_p^2	F	p	η_p^2	F	p	η_p^2	F	p	η_p^2	F	p	η_p^2
Time spent training	10.6	.001	.071	37.5	<.001	.209	4.5	.092	.031	12.9	<.001	.083	2.3	.135	.020	4.9	.029	.042
Tutorial levels time	69.7	<.001	.334	1.3	.254	.009		n/a		<0.01	.972	<.001	196	<.001	.231	17.9	<.001	.138
Platforming familiarity	6.5	.012	.045		n/a		8.0	.005	.054		n/a		13.9	.001	.113		n/a	
Attentional control	0.15	.697	.001	2.3	.133	.016	0.3	.590	.002	0.01	.939	<.001	0.5	.495	.004	1.0	.329	.009
Probability of success	0.03	.952	<.001		n/a			n/a			n/a		0.01	.926	<.001		n/a	
Perceived task challenge	0.2	.865	<.001		n/a			n/a			n/a			n/a			n/a	
Task-related anxiety	4.9	.030	.033	5.9	.017	.040	4.2	.043	.029		n/a		0.5	.485	.004		n/a	

Table 2. The effects of the covariates, from the separate ANCOVAs. A "n/a" indicates that the individual difference measure was not used as a covariate for that outcome measure.

find a main effect of Checkpoints on death count — they died fewer times if they trained with Checkpoints. For the retention test, we found the same; there was no significant main effect of Checkpoints on levels completed, but there was a significant main effect on death count.

5.3 Were There Any Interactions Between Breaks and Checkpoints? (RQ5)

We hypothesized that the combination of checkpoints and breaks would positively affect progress beyond what either checkpoints or breaks would on their own in training, but not affect transfer or retention. This was explored by looking for significant interactions between Checkpoints and Breaks. There was just one significant interaction effect between Checkpoints and Breaks, affecting the number of deaths during training (see Table 1). All post-hoc pairwise comparisons for the interaction were significant ($p < .001$). Examining the estimated marginal means (Figure 4), the interaction indicates that the introduction of Checkpoints affects the death count differently depending on whether Breaks were also present. When Breaks were present, including Checkpoints resulted in an *increase* to the death count. However, without Breaks, the inclusion of Checkpoints results in fewer deaths.

No other interactions were significant.

5.4 Individual Differences and Covariates

We corrected for individual differences between participants using covariates. The effects of these covariates are reported in Table 2.

Our measures of individual differences of attentional control, perceived probability of success, and perceived task challenge were not significantly related to levels completed or death count, however, task-related anxiety was significantly related to levels completed in Training and Transfer, as well as death count in Training. Platforming game familiarity was significantly related to all measures for which it was included.

We found that total time spent training was significantly related to levels completed during Training only, as well as death count for all sessions. Time spent on the tutorial levels was significantly related to levels completed in Training as well as both levels completed and death count in Retention.

6 ADDITIONAL RESULTS

6.1 Quantitative Subjective Experience

We found a significant interaction between Session and Breaks for Flow ($F_{2,228} = 7.5$, $p < .001$, $\eta_p^2 = .062$). From post-hoc tests, we found that Flow increased between the Training session and Transfer session ($p < .001$) as well as between the Training session and Retention session ($p < .001$) for participants who trained with Breaks.

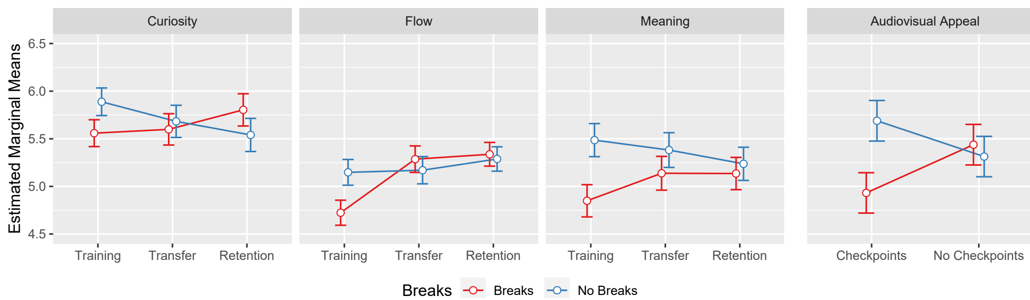


Fig. 5. Estimated marginal means for curiosity, flow, meaning, and audiovisual appeal questions to help understand the significant interactions.

For Worry, we found a significant main effect of Session ($F_{2,228} = 3.5$, $p = .033$, $\eta_p^2 = .030$). Post-hoc tests show that there was a significant increase in Worry between the Training session and the Transfer session ($p < .001$), as well as between the Training session and Retention session ($p = .039$), but no significant difference between the Transfer and Retention sessions ($p = .125$).

For Curiosity, we found a significant interaction between Session and Breaks ($F_{2,228} = 5.5$, $p = .004$, $\eta_p^2 = .046$). Post-hoc tests do not show any significant differences between any pairs; however, when examining the means we see that when players were made to wait during training, their Curiosity about the game increased as they played the transfer and retention versions of the game that did not have players wait (see Figure 5). The opposite occurred for players who trained without Breaks — their Curiosity about the game decreased over the Sessions.

For Meaning, we found a significant interaction between Session and Breaks ($F_{2,228} = 4.9$, $p = .008$, $\eta_p^2 = .041$). Post-hoc tests do not show any significant differences between the groups, so we again look to the means to understand the interaction (see Figure 5). Participants who trained with Breaks found increased Meaning in the Transfer and Retention sessions compared to training; however, if participants trained without Breaks, then they perceived slightly less Meaning as they played the game for longer.

For the measures of Mastery, Immersion, Challenge, and Autonomy, we found no significant within-subject effects or interactions for Session, Breaks, or Checkpoints. We also found no significant between-subject effects or interactions for Breaks or Checkpoints.

6.2 Open-ended Responses

6.2.1 Game-Integrated Breaks. We asked participants who were made to wait to respawn after they died how they felt about it, what they did during their short breaks, and whether the need to wait changed how they approached the game. In terms of how they felt, the majority of participants (89%) gave negative comments, suggesting that the wait was pointless (i.e., there was no in-game context for it), frustrating, punishing (i.e., not only would they have to repeat part of the level, they had to wait to do so), or that it should be shorter. One noteworthy response stated that although the wait time was objectively short, it felt subjectively long. This could be because, as other participants stated, they were simply eager to get back to playing the game.

Most participants' (77%) focus remained on the game during the break. Either by simply watching the timer tick down (37%) or by considering what they were doing within the game (40%). Most participants (56%) who waited thought that the breaks did not change their approach to playing the game, although others stated that they did. In particular, many suggested that the breaks dissuaded them from taking risks in the game, especially as they got further through the level.

For participants who did not have to wait when they died (their respawn time was fixed to 1 second), we asked them to speculate about whether a break might have helped if they happened to take a break at any point, and if the ability to attempt the game again immediately affected how they approached the game. Only some participants thought that a break would be useful (21%), and over half (56%) thought a break would provide no benefit. Participants thought that a break could hinder their concentration or performance, but might help them deal with fatigue or frustration, particularly if the game had gone on longer. However, about a quarter of these players (24%) did in fact take a short break of some kind. Usually, it was to attend to a distraction, but some took a purposeful break (e.g., to stretch, refocus themselves, or hydrate).

When asked if they thought that being able to attempt the level again right away affected how they played, the majority (79%) said that it did, particularly by allowing them to approach the game more recklessly and by intentionally making many attempts to learn the game.

Participants who trained with game-integrated breaks had an opportunity to play the game without breaks in transfer. When asked which version of the game they preferred, about 43% said they preferred the training version (with breaks), compared to 52% who said that they preferred the transfer version (with no breaks). Of these, 15% stated explicitly that they liked not having to wait as long. When asked which version of the game was more difficult, 67% said that the second version was more difficult, compared to 25% who said that the first version was more difficult (the remainder did not answer the question).

Participants for whom the only change in the game was the new levels (i.e., they trained without checkpoints or breaks) preferred the first version of the game (60.5%), but they were split on which version of the game was more difficult (42.1% said version 1 as more difficult and 44.7% said version 2).

6.2.2 Checkpoints. For participants who trained with checkpoints enabled, we asked them how they felt about checkpoints, whether checkpoints made the game easier, and if they approached the game differently because of the checkpoints. The majority of participants seemed to like the checkpoints. No participant stated that they disliked the checkpoints, and the majority said that they were spaced about right (75%). Most participants (85%) said that checkpoints made the game easier, although two participants pointed out that the game wasn't necessarily easier, just that their progress was saved.

When asked if and how checkpoints altered their approach, most (63%) said that it did. The primary reason being that it allowed them to take risks that they otherwise would have avoided; another was that they felt less pressure knowing that they would not lose all of their progress.

For participants who trained with checkpoints disabled, we asked whether the walk back was easier than the section of the game where they died if they focused on anything while they walked back, if they approached the game differently due to needing to walk back, and if they had any general comments regarding this design choice. More than half (64%) said that this process was easier than the section of the game where they died. Most (73%) participants focused on some aspect of the game while they worked their way back to where they died. This included how to move through the level more efficiently, avoiding mistakes they made previously, thinking about how to overcome obstacles later in the level, and reflecting on the controls. On whether the walk back changed their approach to the game, many (75%) said that it did, by approaching the game more cautiously, putting more effort into playing the game, or by causing them to focus on learning aspects of the game or a specific level.

General comments on the need to walk back prompted more negative comments than positive comments (44% vs. 21%). Players said they were frustrated, annoyed, anxious about dying, or that they simply hated it. Some positive comments said that this process motivated them to put in more

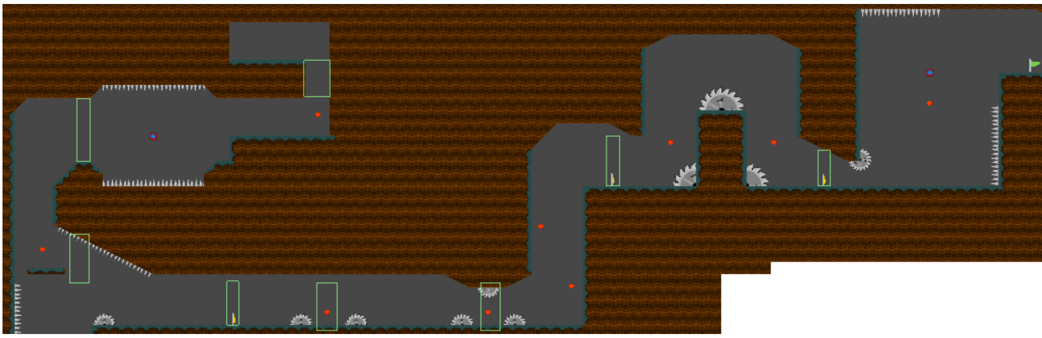


Fig. 6. Example of one of the game's levels. The spots on the map where progress is logged are represented by the green rectangles. The checkpoints are the yellow flags, and the end of the level is the green flag.

effort, or that it made the completion of a level feel more rewarding. About half of the comments (52%) discussed the game's design, for example, that the walk back was simply part of the game's challenge, or that this design was okay because it was done in other games. Other responses touch on the absence of checkpoints, with many stating that the game would benefit from checkpoints, although some stated that the levels were short enough to not need them. Related to this, several participants commented on the length of the level being a factor; they stated that the walk back was not a big deal as long as the level was relatively short. A few participants felt that repeating parts of the level allowed them to learn and improve at the game.

Participants who trained with checkpoints were able to play the game without checkpoints during transfer. These participants had a strong preference towards the training version of the game, with checkpoints enabled (73%). Of these participants, a third said they liked the checkpoints of the training version. In terms of difficulty, nearly all participants (87%) said that the second version of the game was harder, and 38% of these participants said that the lack of checkpoints was a reason why the transfer version of the game was more difficult.

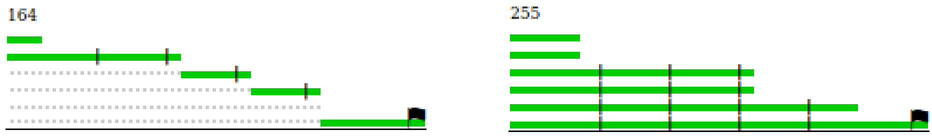
6.3 Visualizations

We generated visualizations from the game logs to represent player progress. These show each attempt made by each player as a green line (Figure 7); vertical bars indicate checkpoints. When a player dies, the green line restarts from the beginning (if no checkpoints), or from the last checkpoint reached; dashed lines represent the part of the level that was skipped due to the checkpoint. Completed levels are marked with a flag.

The power law of practice [35] suggests that players should continuously improve at the game and continue to make progress. Figure 7 shows two examples that approach the "ideal" learning situation: the player makes steady progress in the level without stalling. In contrast to this ideal, we observed two types of inhibited progress (see Figure 8):

- **Stalled progress**. The player is unable to progress beyond past attempts; they are "stuck" and unable to proceed, at least for a time.
- **Regressed progress**. The player fails to reach a part of the level that they had reached in a previous attempt.

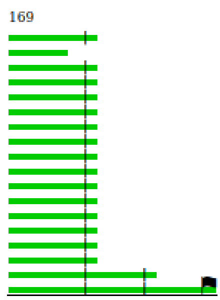
These observations prompted us to quantify challenges to making progress within the game to better understand how checkpoints and breaks were affecting player progression.



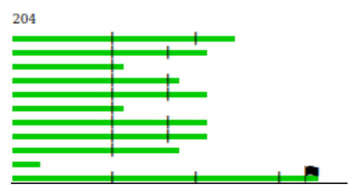
(a) Example with checkpoints enabled.

(b) Example with checkpoints disabled.

Fig. 7. Examples of near-to-ideal scenarios in which players make consistent progress during training and do not require many attempts to overcome a challenge. The first attempt is the top green line and the last attempt is the bottom green line. The examples are from the same level; the number at the top is the participant's ID.



(a) Stalled progress. Lines of similar length indicate that the player is stuck and unable to get further in the level.



(b) Regressed progress. A line that is shorter than the line above it indicates that the player was unable to make as much progress as they had in previous attempts.

Fig. 8. Examples visualizations of the two scenarios relating to lack of progress during training.

6.3.1 Quantifying Stalled and Regressed Progress. Based on these observations, we used the game logs to work out two measures of ways that progress deviated from the ideal: the time for which players' progress was stalled, and the number of times that regressed progress occurred.

For *stalled progress time*, the logs of attempts were processed in the order in which they were collected. For each attempt, the total progress made was compared to past attempts. If the progress made was less than the best past attempt, then that attempt would be considered to be an attempt in which the player was stuck, and that attempt's time would be added to the total time spent stuck.

For *regressed progress count*, we calculated how their current attempt's progress compared to their best attempt. If the player failed to make as much progress, then the attempt was considered to be one where progress regressed. All regressed progress attempts are also stalled progress attempts as they are just a more severe form of being stalled.

The descriptive results of these analyses are presented in Figure 9. To check for significant differences between the groups, we used separate two-way ANOVAs for Stalled Progress Time and Regressed Progress Count, with Breaks and Checkpoints as between-subject factors. Any post-hoc tests used Bonferroni corrections.

In Training, we found significant main effects of Breaks ($F_{1,146} = 90.1$, $p < .001$) and Checkpoints ($F_{1,146} = 22.8$, $p < .001$) on Stalled Progress Time, with both decreasing Stalled Progress Time. There was also a significant interaction between Checkpoints and Breaks ($F_{1,146} = 4.9$, $p < .029$) — post-hoc tests show that the combination of Breaks and Checkpoints greatly reduce stuck time than

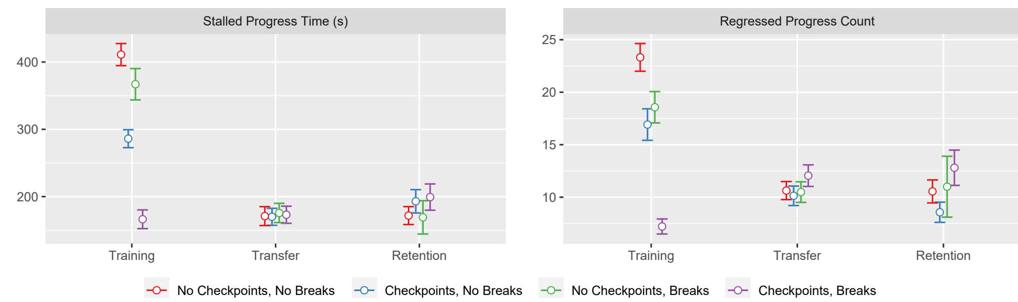


Fig. 9. Results of our measures of players' progress being inhibited, in terms of stalled progress time and regressed progress count. Error bars are standard error.

either on their own. We also found significant main effects of Breaks ($F_{1,146} = 47.2$, $p < .001$) and Checkpoints ($F_{1,146} = 31.3$, $p < .001$) on Regressed Progress Count, with both reducing Regressed Progress Count. The interaction was not significant ($F_{1,146} = 31.3$, $p = .057$).

In the Transfer test, there were no significant main effects, nor any significant interaction effects for Stalled Progress Time or Regressed Progress Count (all $p \geq .280$). Similarly, in the retention test, there were no significant main effects, nor any significant interaction effects for Stalled Progress Time or Regressed Progress Count (all $p \geq .148$).

Therefore, Checkpoints and Breaks both helped players make progress while they were present, and this different training did not affect their progress in the later test sessions. There was also an interaction between Checkpoints and Breaks for Stalled Progress Time — Breaks were much more effective when Checkpoints were also present.

7 DISCUSSION

The study provided several new findings about the effects of breaks and checkpoints on progress, skill development, and player experience:

- Players completed significantly more levels with both breaks and checkpoints (three additional levels in both cases compared to the baseline) during the training session (when the techniques were present), and the combination of breaks and checkpoints showed the largest increase in progress (eight levels better, although the interaction was not statistically significant);
- There were no significant differences in levels completed on either the transfer or retention tasks, suggesting that neither technique reduced (or improved) skill development beyond the training session;
- Both checkpoints and breaks showed improvements in our secondary measures of progress — breaks reduced player death rate (likely due to the forced pause after dying), and the combination of the two techniques significantly reduced the number of stalls and regressions;
- For measures of player experience, no differences were found between training and transfer sessions for checkpoints; for breaks, we found improvements in three measures (flow, curiosity, and meaning) when participants moved to the tasks where breaks were removed;
- No participants stated that they disliked checkpoints, and no participants stated that they liked the breaks.

7.1 Explanation of Results

7.1.1 Why Did Game-Integrated Breaks Help Progress? Past work that has applied theories of spaced practice to games has done so on a fixed schedule, for example, a two-minute break after five minutes of play [25, 38], whereas this work presented breaks to players dynamically, after each death. We found benefits to presenting breaks in this way to players that were similar to the benefits found within other work that also applied spaced practice to games [25, 38]. Participants completed more levels, stalled fewer times, and spent less time not making progress when breaks were present. This strongly suggests that spaced practice can also be effective with a dynamic schedule and with dynamic break lengths. Past work not looking at games or even perceptual-motor skill have also found that shorter breaks can be effective [22, 58] and also that a variety of schedules for the breaks can be effective [47]. We find that the same is true for games.

More generally, spaced practice aids performance in a variety of ways. It aids skill development by helping learners develop memories that allow them to carry out a task [58]. This might be possible due to encoding variability, where events spaced by time can be encoded into memory in different ways [5, 58]. Spacing may also assist in the *consolidation* process, in which memories become more stable and resistant to decay [9, 52], as well as force a retrieval of the relevant memory traces when returning to the task, which reinforces them [44, 63]. Short break intervals in particular might be effective because of the concept of deficient processing, in which less attention is given to the second or subsequent attempts [22]. A short time delay between attempts could prompt a learner to direct more of their voluntary attention toward a task [22]. Similar to this is the thought that a short break could give one time to recover from physical [47] or cognitive [2, 63] fatigue.

Spaced practice helps learners transition from early stages of learning to later stages of learning [54], in which learners can execute a skill with more fluency and better attend to the relevant stimuli [57]. It also aids the process of *knowledge compilation* [3], in which declarative knowledge (verbal information about a skill) becomes encoded as procedural knowledge, which can be more directly applied to executing a skill.

Finally, we must acknowledge that game-integrated breaks did come at a cost — players overall spent less time actively playing the game when these types of breaks were used. When breaks were included, players spent an average of 274 seconds waiting (min=124, max=451, SD=68.4), compared to 61.2 seconds (min=29, max=124, SD=18.8). Therefore, unlike in past experiments, the total training time was not fixed; players who were given more breaks simply had less time to complete the levels. We corrected for this by including time spent training as a covariate in our analyses, and this variable was significantly related to the levels completed in training and death count in training, indicating that this reduction in training time did have a cost.

7.1.2 Why Did Checkpoints Help Progress? We observed that checkpoints helped players make progress when they were present and did not get in the way of learning the game. Increased progress was made in part due to players making more consistent progress, stalling fewer times and for less time.

In terms of theory, in Section 2.2.2, we proposed that checkpoints could affect the variability of practice, as well as facilitate the process of part-task practice. Considering part-task practice, checkpoints implicitly allow this to occur by allowing a player to focus only on overcoming a single challenging obstacle or section of the game. Players can try to correct errors they are making by attempting different strategies. In the written feedback, we found that many participants took advantage of this learning opportunity. In cases where this worked, players overcame obstacles, which means that they progress in the game and can then be also exposed to a greater variety of practice.

Practice would have more variety if players were able to make progress in the game, and they did. In non-game contexts, variability in practice generally comes with increased errors in early learning [31], which we did observe somewhat, as players died more times (though only when game-integrated breaks were not present). However, due to the task being a game in which players make progress, the increased errors did not mean that performance was worse, as players were clearly able to make more progress when checkpoints were present.

We compared checkpoints to the absence of checkpoints, and so part of the reason why checkpoints were beneficial is that playing without checkpoints — making players work their way back through the level — was likely not helpful. We found no evidence that working through the potentially easier, early parts of the level was beneficial training.

We did not find any long-term benefits to progress due to checkpoints that would have indicated greater skill development. This does not mean there were no learning benefits at all, only that there were no additional benefits over the other variations tested. One important consideration, however, is that unlike traditional perceptual-motor skills, immediate performance can be more important than long-term performance in games, as a player's primary goal is often to make progress [28].

7.2 Do Players Hate Breaks?

None of our participants stated that they liked taking a break, and an overwhelming majority were critical of the wait. It was viewed as too long, punishing, and pointless. This makes sense considering that games are played for leisure, often as a change or a distraction from other aspects of a person's life. So why take a break? We show that there are good reasons to do so. In addition to any benefits of reducing fatigue or repetitive strain issues [36], we find clear and immediate performance benefits of taking a break, even one as short as ten seconds. And yet, the suggestion that players take a break to aid their progress might be met with even more opposition. When a player is engaged with a game, intently focusing on overcoming a challenge, all a player might want to do is keep trying. Despite our participants disliking the breaks, we did have participants use the breaks productively, as an opportunity to refocus or collect themselves. A player's dislike of breaks does not mean that they are ineffective, and without a break, they may have never been prompted to reflect on what they were doing. In fact, the players who want to get back to playing the game might be more inclined to use their break to consider what they are doing in the game.

Disliking the breaks within a game does not necessarily translate to disliking the game as a whole. In our game, the presence of game-integrated breaks did not significantly affect how they experienced the game in terms of ease of control, progress feedback, goals and rules, mastery, immersion, challenge, and autonomy. A similar lack of change in subjective experience was also found in past work [25, 38]. We only found that flow and meaning were slightly reduced when breaks were present. Flow might have been reduced due to the interruptions that our breaks created, and a reduction in meaning means that participants thought the task was less important. Even with this, when players had the opportunity to play the game again without breaks, there wasn't a strong preference for this version of the game, and players tended to think that the version of the game with breaks was easier, even though it really was not; they were just performing better.

If breaks do bother the player, this raises the question of whether it is more important that a player gets better at the game or enjoys it. This question is difficult to answer, as competence within the game is very much linked to a player's motivation to keep playing [40, 45]. However, not every game has an emphasis on performing procedural skills well. It may not make much sense, for example, to include breaks in games such as *The Sims* or *Animal Crossing*, which are simulation games that place fewer demands on perceptual-motor skill development. In a game like *Super Meat Boy*, in which the ability to execute a series of precisely timed inputs is needed to succeed, the benefits of breaks may greatly increase the overall enjoyment of the game.

7.2.1 Improving Acceptance of the Breaks. A criticism of breaks mentioned by some participants was that there was no context for the break. For example, one participant pointed out that if the game had a visible time limit, then the wait would make more sense (possibly because it would more explicitly be a form of punishing the player for failure). Other participants pointed out that a break would make more sense for different types of games such as multiplayer games, where players already accept breaks in the form of respawn timers or waiting for the matchmaking system to pair them with other players. If there seems to be a reason for the breaks, players may be more accepting of them.

Players may also be more accepting of the breaks in our game if their presentation to players was improved. When players died, we simply showed a timer. Death in other games involves animations, kill cams (replays where you can see how you die), spectating other players, or amusing cut-scenes such as how characters taunt Batman when he dies in the *Arkham* series of games [4]. Additionally, our breaks were predictable. Players learned that when they died they would need to take a break. It might be better to wait for indications that the player is struggling before trying to provide aid, such as after a certain length of time with stalled progress or after a certain number of progress regressions. Finally, it might be possible for players to continue to engage with the game in some way while they wait. For example, past work explored what players could do during breaks and found that simply switching to a separate but related task is beneficial [38].

7.3 Implications for Players and Game Designers

We learned that many players' intuition regarding breaks and checkpoints is wrong. Breaks, something that many players are opposed to, can be beneficial and enhance player performance. Checkpoints, unlike what many believe, do not necessarily make the game easier or result in players becoming reliant on them. In our game, checkpoints did not affect skill development and were beneficial because they helped players make progress. Game designers should consider whether there is value in including breaks or checkpoints in their games.

For checkpoints, we modelled our checkpoints on those already found within many commercial games. However, current biases about checkpoints and the notion that they may hinder learning could mean that there is room for improvement. Designers should carefully consider the frequency of checkpoints. We found no benefits to replaying earlier parts of the game compared to making use of checkpoints, and checkpoints can help players more rapidly apply a trial-and-error approach to overcome challenges. Checkpoints should not be considered to be a crutch, but instead a valuable tool.

In contrast to checkpoints, breaks are not typically included in games with the intention of helping players make progress and improve at the game. We extend past work by demonstrating that breaks can be presented with a dynamic schedule and the breaks can even be quite short — as short as ten seconds — while still improving performance. We gave players a break after each time they died, an implementation already found in some games (such as in many online first-person shooters) and one that is relatively easy to implement. While players disliked this, we note that it did not meaningfully affect their enjoyment of the game as a whole. Additionally, it should be possible to design breaks that are less apparent but still serve to improve performance. Finally, breaks prompt players to approach the game with more caution, taking fewer risks and dying fewer times. If players are repeatedly throwing themselves at a challenge without taking time to consider what they are doing or why they are failing, there is a good chance that a break will prompt them to do so.

Combined, checkpoints and breaks together provide game designers with a lot of control over how players make progress in a game. Being able to predict and dictate a player's progress can allow game designers to craft enjoyable game experiences by ensuring players are consistently and reliably progressing at the appropriate pace. For example, in the scenario presented in the introduction, what can be done if the player is repeatedly trying to overcome a challenge in a game but is unable to do so? If they are playing with checkpoints enabled, but no break, then it is likely they are attempting various strategies and observing the results. But they might also be simply throwing themselves at the problem without actively considering what they are doing. In this case, a short break may be just what the player needs to allow them to reconsider what they are doing and redirect their attention towards making the most of their attempts, rather than trying the same approach repeatedly.

7.4 Limitations and Future Work

The main limitation of this study is that we only tested one game, and one implementation of breaks and checkpoints. Furthermore, our game was specifically built for the study — it was not a commercial game. Further work is needed to determine if our results generalize to other side-scrolling platform games, and to other game genres. Future work could reproduce the results using other games, as well as testing different variations of our implementation of checkpoints and breaks. For breaks, given the variety of prior games in which spaced practice has been tested, we are confident that the benefits found with our specific implication of spaced practice will also be present in other games and game genres. In terms of checkpoints, this concept is also found in genres other than platformers and the way in which it is implemented is similar across genres. There is nothing specific about checkpoints found in side-scrolling platformers, compared to, for example, single-player first-person shooters; however, we must acknowledge that differences are a possibility. More generally, our platformer may not have been as polished as a commercial game. A professionally made and play-tested game designed to be played for many hours may result in a different experience.

Future work could investigate how breaks could be designed and implemented into commercial games. In particular, additional studies could explore how to improve the presentation of breaks and whether different frequencies or lengths would also be effective. The design of our breaks made them very obvious and uninteresting. This gives several opportunities for future work. Our breaks were highly predictable — players knew that any time they died they would need to wait. Future work could explore variations on dynamically presenting the breaks. Considering that spaced practice can enhance performance, our motivation was to provide aid to a player at a time when it might be effective, but it may be possible to time the breaks differently or present them only under specific circumstances. For example, by considering indicators of inhibited progress. Another opportunity to improve our breaks is to make them more interesting. Future work should investigate whether presenting the break differently can make the break less apparent or increase acceptance of the break. For example, instead of a simple timer, animations or short activities (like those in [38]) could be included to entertain the player and potentially improve the experience of breaks or waiting.

Testing different frequencies or lengths of breaks is also something future work should explore. Our breaks were only up to ten seconds long, but breaks as long as two minutes [38] or even one day [25] have also been tested. Based on past work, breaks of a variety of different lengths and frequencies could be useful. For short breaks, it is thought that a short timer interval between rehearsals prompts a learner to direct more of their voluntary attention towards the task [22] while also aiding with processing [58]. Longer breaks are thought to be beneficial due to strengthening the memory of how to execute the relevant skills [33, 41] or due to the second event being encoded differently in memory [5]. However, there may be upper or lower limits to explore. At some point, the breaks may be too long or too short, or presented too often or too infrequently to be of any benefit. It may also be that the most effective break length and frequency changes as players improve [7].

Additionally, future work could test whether breaks could be turned into short training scenarios. Prior work found that simply switching to a new task for an implicit break had benefits similar to a more explicit break [38] and so it may be possible to do this without losing the benefits of spaced practice. Further, it may be that some of the breaks found in commercial games are already doing this. Consider, for example, spectating your teammates in an online first-person shooter game while you wait to respawn. Past work has found that there are performance benefits to watching demonstrations of a game [39]. Simply spectating other players may also act as a demonstration that could benefit a player's performance.

There are opportunities for future work to test varying implementations of checkpoints. Checkpoints — or more generally, saving progress — can be presented in many different ways. Ours were presented as visible goal markers for players to reach, but they may also be invisible (i.e., auto-saves [59]). We placed checkpoints after what we felt were difficult obstacles in the level, but we could have spaced them more or less frequently than we did. It is possible that a system that prevents even the smallest loss of progress will affect performance and learning differently than our implementation.

We logged the number of deaths as a way of understanding how checkpoints and breaks were affecting player behaviour in the game. The death count was affected by breaks and checkpoints, but we do not know precisely why. With breaks, players may have avoided dying within the game to avoid the delay (some participants brought this up in their written responses). Those who played with checkpoints stated that the checkpoints allowed them to take additional risks (although we found no main effect of Checkpoints on death count). However, if breaks were given then the death count was low regardless of whether checkpoints were active. There is an interaction between the two that is worth further investigation — are players making an intentional decision to take more risks when checkpoints are present, but only if breaks are not also present? Furthermore, past work has found a relationship between death and a player's perception of challenge [26] — players view a game as more challenging if they die more frequently [10], yet we found no differences in challenge even though we found differences in death counts due to checkpoints and breaks. In our study, death was not necessarily an indicator of a lack of progress within the game in terms of levels completed — players might have died more often yet made more progress, and players may have even intentionally chosen to die in order to attempt different strategies. The relationship between deaths and perception of challenge is likely more complex than past work suggests, and future work should examine the relationship further.

A final limitation of our study is that our players were paid to play this game. This means that players might have been willing to put up with frustrations more than they typically would in a game. This also could have influenced them to rate their experience more positively than they may have otherwise, considering how interesting a game is when compared to other tasks on Amazon's Mechanical Turk.

8 CONCLUSION

Ensuring that players continue to make progress in a game is of great importance to both game designers and players. We carried out a study to test two techniques of supporting this, breaks and checkpoints. We found that both were effective at supporting players' progress without hindering skill development. Even with relatively short breaks (no longer than ten seconds) that were fully integrated into the game, players were able better able to overcome challenges and make progress, although participants tended to dislike the breaks. Checkpoints were also effective when integrated into the game before difficult obstacles, and were liked by our participants. Our work provides several contributions that can change the way players and game designers think about practice within games:

- We show that spaced practice can be integrated into games using a dynamic schedule rather than a fixed schedule and that the breaks can be as short as ten seconds while still effectively improving immediate performance and aiding progress.
- We show that checkpoints are an effective method of improving immediate performance and helping players make progress and that they do not come with any apparent drawbacks in terms of skill development.
- We show that benefits occur despite player beliefs about checkpoints and breaks, and the subjective experience of the game is largely unaffected.

Our results provide useful information for players who want to improve their skills, practical suggestions for designers who are interested in ways of helping their players make progress through their games and add to our overall understanding of how skill development occurs within games.

REFERENCES

[1] Vero Vanden Abeele, Katta Spiel, Lennart Nacke, Daniel Johnson, and Kathrin Gerling. 2020. Development and validation of the player experience inventory: A scale to measure player experiences at the level of functional and psychosocial consequences. *International Journal of Human-Computer Studies* 135 (3 2020), 102370. https://doi.org/10.1016/j.ijhcs.2019.102370

[2] Phillip L Ackerman, Ruth Kanfer, Stacey W Shapiro, Sunni Newton, and Margaret E Beier. 2010. Cognitive Fatigue During Testing: An Examination of Trait, Time-on-Task, and Strategy Influences. *Human Performance* 23, 5 (10 2010), 381–402. https://doi.org/10.1080/08959285.2010.517720

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

[4] Batman: Arkham Wiki. 2021. Game over scenes. Retrieved July 22, 2021 from https://batmanarkham.fandom.com/wiki/Game_over_scenes

[5] Aaron S. Benjamin and Jonathan Tullis. 2010. What makes distributed practice effective? *Cognitive Psychology* 61, 3 (nov 2010), 228–247. https://doi.org/10.1016/j.cogpsych.2010.05.004

[6] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (jan 2006), 77–101. https://doi.org/10.1191/1478088706qp063oa

[7] Nicholas J. Cepeda, Harold Pashler, Edward Vul, John T. Wixted, and Doug Rohrer. 2006. Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin* 132, 3 (2006), 354–380. https://doi.org/10.1037/0033-2909.132.3.354

[8] Gifford K. Cheung, Thomas Zimmermann, and Nachiappan Nagappan. 2014. The first hour experience. In *Proceedings of the first ACM SIGCHI annual symposium on Computer-human interaction in play*. ACM. https://doi.org/10.1145/2658537.2658540

[9] S E Criscimagna-Hemminger and R Shadmehr. 2008. Consolidation Patterns of Human Motor Memory. *Journal of Neuroscience* 28, 39 (9 2008), 9610–9618. https://doi.org/10.1523/jneurosci.3071-08.2008

[10] Marjorie Ann M. Cuerdo, Anika Mahajan, and Edward F. Melcer. 2021. Die-r Consequences: Player Experience and the Design of Failure through Respawning Mechanics. In *2021 IEEE Conference on Games (CoG)*. IEEE. https://doi.org/10.1109/cog52621.2021.9618894

[11] D Derryberry and M Reed. 2002. Anxiety Related Attentional Biases and Their Regulation by Attentional Control. *Journal of Abnormal Psychology* 111, 2 (2002), 225–236. https://doi.org/10.1037//0021-843X.111.2.225

[12] 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. https://doi.org/10.1037/0021-9010.84.5.795 arXiv:00028282

[13] DoubleDutch Games. 2016. SpeedRunners. Game [PC].

[14] Franz Faul, Edgar Erdfelder, Axel Buchner, and Albert-Georg Lang. 2009. Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods* 41, 4 (nov 2009), 1149–1160. https://doi.org/10.3758/brm.41.4.1149

[15] A. Field. 2018. *Discovering Statistics Using IBM SPSS Statistics*. SAGE Publications. https://books.google.ca/books?id=0QTyswEACAAJ

[16] Fabio E. Fontana, Ovande Furtado, Oldemar Mazzardo, and Jere D. Gallagher. 2009. Whole and Part Practice: A Meta-Analysis. *Perceptual and Motor Skills* 109, 2 (oct 2009), 517–530. https://doi.org/10.2466/pms.109.2.517-530

[17] John R. Frederiksen and Barbara Y. White. 1989. An approach to training based upon principled task decomposition. *Acta Psychologica* 71, 1-3 (1989), 89–146. https://doi.org/10.1016/0001-6918(89)90006-1

[18] James Paul Gee. 2005. Learning by Design: Good Video Games as Learning Machines. *E-Learning* 2, 1 (2005), 5–16. https://doi.org/10.2304/elea.2005.2.1.5

[19] Giant Bomb. 2019. Checkpoints (Concept). https://www.giantbomb.com/checkpoints/3015-292/. Accessed February 5, 2021.

[20] Diane L. Gill and Thomas E. Deeter. 1988. Development of the sport orientation questionnaire. *Research Quarterly for Exercise and Sport* 59, 3 (1988), 191–202. https://doi.org/10.1080/02701367.1988.10605504

[21] Daniel Gopher, Maya Weil, and Daniel Siegel. 1989. Practice under changing priorities: An approach to the training of complex skills. *Acta Psychologica* 71, 1 (1989), 147–177. https://doi.org/10.1016/0001-6918(89)90007-3

[22] Douglas L Hintzman. 1974. *Theoretical Implications of the Spacing Effect.* Lawrence Erlbaum Associates, 77–99.

[23] 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. https://doi.org/10.1145/2658537.2658697

[24] Colby Johanson. 2020. *colbyj/bride-of-frankensystem 1.1.* https://doi.org/10.5281/zenodo.3738761

[25] Colby Johanson, Carl Gutwin, Jason T. Bowey, and Regan L. Mandryk. 2019. Press Pause when you Play. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*. ACM. https://doi.org/10.1145/3311350.3347195

[26] Jesper Juul. 2009. Fear of failing? the many meanings of difficulty in video games. *The video game theory reader* 2, 237-252 (2009).

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

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

[29] Ben Kuchera. 2014. The passion of the checkpoint: Why gaming's most frustrating failure is so hard to fix. *Polygon* (2014). Retrieved February 5, 2021 from https://www.polygon.com/2014/2/25/5422328/the-passion-of-the-checkpoint-why-gamings-most-frustrating-failure-is

[30] 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. https://doi.org/10.1080/02701367.1988.10609373

[31] R.A. Magill. 2007. *Motor Learning and Control: Concepts and Applications.* McGraw-Hill. https://books.google.ca/books?id=rCPdAAAACAAJ

[32] Spero A Metalis. 1985. Effects of massed versus distributed practice on acquisition of video game skill. *Perceptual and Motor Skills. Vol 61(2)* (1985), 457–458. https://doi.org/10.2466/pms.1985.61.2.457

[33] Everett Mettler, Christine M. Massey, and Philip J. Kellman. 2016. A comparison of adaptive and fixed schedules of practice. *Journal of Experimental Psychology: General* 145, 7 (jul 2016), 897–917. https://doi.org/10.1037/xge0000170

[34] James C. Naylor and George E. Briggs. 1963. Effects of task complexity and task organization on the relative efficiency of part and whole training methods. *Journal of Experimental Psychology* 65, 3 (1963), 217–224. https://doi.org/10.1037/h0041060

[35] 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.

[36] Nintendo. 2021. Wii - Health and Safety Precautions. *Nintendo* (2021). Retrieved July 12, 2021 from https://www.nintendo.com/consumer/systems/wii/en_na/health_safety.jsp

[37] Nintendo EAD. 1990. Super Mario World. Game [SNES].

[38] Brandon Piller, Colby Johanson, Cody Phillips, Carl Gutwin, and Regan L. Mandryk. 2020. Is a Change as Good as a Rest? Comparing Break Types for Spaced Practice in a Platformer Game. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*. ACM. https://doi.org/10.1145/3410404.3414225

[39] Barbara J. Pollock and Timothy D. Lee. 1992. Effects of the Model's Skill Level on Observational Motor Learning. *Research Quarterly for Exercise and Sport* 63, 1 (1992), 25–29. https://doi.org/10.1080/02701367.1992.10607553 PMID: 1574658.

[40] 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. https://doi.org/10.1037/a0019440

[41] Mary A. Pyc and Katherine A. Rawson. 2009. Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory? *Journal of Memory and Language* 60, 4 (may 2009), 437–447. https://doi.org/10.1016/j.jml.2009.01.004

[42] 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.

[43] Scott Rogers. 2014. *Level Up! The guide to great video game design.* John Wiley & Sons.

[44] Kelly Braun David C Rubin. 1998. The Spacing Effect Depends on an Encoding Deficit, Retrieval, and Time in Working Memory: Evidence. *Memory* 6, 1 (1 1998), 37–66. https://doi.org/10.1080/741941599

[45] 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 (2006), 347–363. https://doi.org/10.1007/s11031-006-9051-8

[46] Richard A. Schmidt. 1975. A schema theory of discrete motor skill learning. *Psychological Review* 82, 4 (jul 1975), 225–260. https://doi.org/10.1037/h0076770

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

[48] Charles H. Shea and Robert M. Kohl. 1990. Specificity and Variability of Practice. *Research Quarterly for Exercise and Sport* 61, 2 (jun 1990), 169–177. https://doi.org/10.1080/02701367.1990.10608671

[49] Wayne L Shebilske, Barry P Goettl, Kip Corrington, and Eric Anthony Day. 1999. Interlesson spacing and task-related processing during complex skill acquisition. , 413–437 pages. https://doi.org/10.1037/1076-898X.5.4.413

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

[51] Tom Stafford, Sam Delvin, Rafet Sifa, and Anders Drachen. 2017. Exploration and Skill Acquisition in a Major Online Game. In *Proceedings of the 39th Annual Meeting of the Cognitive Science Society.*

[52] Tom Stafford and Erwin Haasnoot. 2016. Testing Sleep Consolidation in Skill Learning: A Field Study Using an Online Game. *Topics in Cognitive Science* 9, 2 (nov 2016), 485–496. https://doi.org/10.1111/tops.12232

[53] Team Meat. 2010. Super Meat Boy. Game [PC].

[54] 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.

[55] The Jamovi project. 2022. jamovi (Version 3.2) [Computer Software]. Retrieved from https://www.jamovi.org.

[56] L Thomsen, F Petersen, A Drachen, and P Babaei. 2016. Identifying Onboarding Heuristics for Free-to- Play Mobile Games : A Mixed Methods Approach Identifying Onboarding Heuristics for Free-to-Play. October (2016), 1–8.

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

[58] Thomas C. Toppino and Emilie Gerbier. 2014. About Practice. In *Psychology of Learning and Motivation.* Elsevier, 113–189. https://doi.org/10.1016/b978-0-12-800090-8.00004-4

[59] TV Tropes. 2021. Auto-Save - TV Tropes. Retrieved July 26, 2021 from https://tvtropes.org/pmwiki/pmwiki.php/Main/Autosave

[60] u/Cosmic_Colin. 2014. r/truegaming - Is using save states really cheating? *reddit* (2014). Retrieved February 5, 2021 from https://www.reddit.com/r/truegaming/comments/24miq7/is_using_save_states_really_cheating/

[61] u/imaunitard. 2012. r/gaming - Why not take a break? *reddit* (2012). Retrieved July 12, 2021 from https://www.reddit.com/r/gaming/comments/x6exy/why_not_take_a_break/

[62] u/Kizuxtheo. 2018. r/Megaman - Thoughts on save states? *reddit* (2018). Retrieved August 31, 2020 from https://www.reddit.com/r/Megaman/comments/7ij4so/thoughts_on_save_states/

[63] 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. https://doi.org/10.1016/j.humov.2018.04.004

[64] Regina Vollmeyer and Falko Rheinberg. 2006. Motivational effects on self-regulated learning with different tasks. *Educational Psychology Review* 18, 3 (2006), 239–253. https://doi.org/10.1007/s10648-006-9017-0

[65] L.S. Vygotsky. 1978. Interaction between learning and development. , 79–91 pages. https://doi.org/10.1016/S0006-3495(96)79572-3

[66] Waluigilicious. 2016. Save states on old games. Let's end this argument for good. *NeoGAF* (2016). Retrieved August 31, 2020 from https://www.neogaf.com/threads/save-states-on-old-games-lets-end-this-argument-for-good.1209168/

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

[68] Gabriele Wulf and Charles H. Shea. 2002. Principles derived from the study of simple skills do not generalize to complex skill learning. *Psychonomic Bulletin & Review* 9, 2 (jun 2002), 185–211. https://doi.org/10.3758/bf03196276

Received 2023-02-21; accepted 2023-07-07