Edith Cowan University

Research Online

2021

A game-based online tool to measure cognitive functions in students

Valeska Berg
Edith Cowan University

Mark McMahon
Edith Cowan University

Shane L. Rogers
Edith Cowan University

Michael Garrett
Edith Cowan University

Dominic Manley

See next page for additional authors

Authors

Valeska Berg, Mark McMahon, Shane L. Rogers, Michael Garrett, Dominic Manley, and Gabbie Miller

A Game-Based Online Tool to Measure Cognitive Functions in Students

Valeska Berg[1], Mark McMahon[1], Shane L. Rogers[1], Michael Garrett[1,2], Dominic Manley[2], Gabbie Miller[2]

[1]*Edith Cowan University, Perth, Australia, v.berg, m.mcmahon, shane.rogers}@ecu.edu.au*

[2]*Cinglevue International, Perth, Australia, michael.garrett, dominic.manley}@cinglevue.com, gabbiemmiller@gmail.com*

Abstract

Cognitive assessments can be expensive, lengthy and fatiguing for students and are often conducted in an artificial clinical context. In an effort to make the assessments more fun, researchers have started to introduce game elements to traditional cognitive tasks and training. This comes with a number of challenges. The main challenge is to develop an engaging tool that at the same time reliably assesses cognitive constructs in students. To address these challenges, this research aims to improve cognitive assessment with a new game-based assessment app that has been designed and developed in collaboration with researchers, teachers, students, and software engineers based on established cognitive theories, and subsequently validated through iterative testing in real world settings. The iterative development process is based on design-based research and includes cycles of design explorations, testing, analyses, redesign, and evaluation with students in authentic educational settings. The knowledge gained from the iterative process of designing a valid cognitive function app can inform other researchers who are aiming to develop cognitive assessment tools in an educational context.

Keywords*: Assessment, Design-based Research, Cognitive Assessment, Education, Gamification, Cognitive Functions*

1 Introduction

While there is an increasing demand for cognitive assessment tools for children, designing child-friendly assessment tools that are enjoyable and valid is challenging. Tests that assess cognitive skills can be extensive, tiring and stressful. Cognitive assessments have a long history and were initially designed for clinical purposes over seven decades ago (e.g. [1, 2]). These *traditional assessments* can be lengthy and repetitive, often require the assistance of trained professionals, and are typically applied in a context that does not reflect the real life of a student [3]. Thus, there is a need to shift away from traditional cognitive testing to a more enjoyable and accessible way of testing in a context that reflects the environment of the student [4, 5]. The challenge is to design a tool for students that meets these criteria while maintaining the validity of the assessment.

Traditional cognitive assessments typically involve children being assessed in a private room with one or more assessors. This environment can be anxiety-inducing and stressful, especially for children who are unfamiliar with these types of situations [6]. Furthermore, lengthy cognitive assessments can be boring for children to the point where they disengage with the tasks [3]. Therefore, research is shifting away from traditional cognitive testing to a more fun, accessible, and easy approach [4, 7]. To make the tasks more enjoyable, assessors have started to add gamified elements to traditional cognitive task designs [8, 9]. Improving the task design of traditional cognitive tasks is thought to increase students' engagement with the tasks, thereby raising the probability of measuring children's full cognitive capabilities [7, 10]. However, this is an emerging field that is still lacking valid research (tools) [8, 11]. This is partly due to the fact that a large amount of research in the gamification literature focuses on cognitive *training* instead of *assessment* [10, 12, 13]. While testing environments are slowly evolving, most still require one-on-one support from a trained assessor.

This article describes the iterative process of developing a new cognitive assessment app called eFun, which was designed in an effort to make cognitive assessments more accessible, valid, and fun for students. eFun is a self-administered series of games that measure a subset of cognitive functions called 'executive functions. Executive functions have been shown to predict academic and career success and are related to quality of life [14-16]. eFun was developed with a theory-driven design through iterative cycles of design, it was tested in real world settings, and is based on established executive function theories and user feedback.

1.1 Measuring executive functions

Executive functions (EFs) are a set of interrelated cognitive skills that enable us to successfully navigate daily challenges, and have repeatedly been shown to predict success across the life span [2, 14, 17-19]. These regulatory skills are necessary to remember and manipulate the information we receive (working memory), inhibit irrelevant thoughts and behavior (inhibition control), and flexibly apply rules to the right context (cognitive flexibility). Collectively they enable us to solve tasks and reach goals and as such are duly recognized as an important factor in educational success [14, 15, 17, 20, 21]. Thus, improving executive function assessment tools in educational contexts is a crucial first step to support students' cognitive development and create a foundation for learning.

Many researchers face the challenge of finding or developing cognitive assessment tools for students that are valid yet engaging [10]. We will discuss the problems underlying this challenge with an example of the iterative design process of eFun. The main problem is that existing validated executive function assessment tools are not always suited for use with students. For example, if the student perceives the task as too effortful, frustrating, and/or repetitive, it can result in participant disengagement which in turn, may negatively impact data quality [10]. It has been shown that data quality can be negatively affected if the student puts in low effort on cognitive tasks including executive function tasks [22]. Furthermore, it has been shown that if the student enjoys the task and finds it interesting, performance is higher [23]. Task enjoyment has also been found to be positively associated with attention and task persistence [24, 25], which can lead to better performance [26, 27].

To increase task engagement and effort, some researchers have added game elements to traditional tasks [10, 28]. However, introducing game-like elements can diminish their potential motivational benefits if they distract the participant to the point that the construct of interest is no longer being measured reliably [3, 28]. For example, from a gaming perspective it would be appealing to have an attractive background, but this could risk players' ability to distinguish elements on the screen that they need to interact with. Similarly, game designers value dynamic game elements, however, from a cognitive perspective it can interfere with the assessment of cognitive skills [29]. This means that game elements need to be introduced carefully without distracting the player from the core task. In line with this, Lumsden et al. [10] suggest that gamification can provide a way to develop engaging and yet scientifically valid cognitive assessments if it is applied carefully. Thus, the aim is to successfully import game design elements into EF tasks without undermining their validity. This is expected to improve the quality of the outcome data and enhance the experience for participants (Lumsden et al., 2016).

Another challenge is to make the tasks enjoyable while maintaining the underlying theoretical construct of the original task. There is a need to design measurement tools for students that use game elements to increase engagement, while simultaneously keeping the task focused on the assessment of psychological capacities consistent with theoretical guidelines [7]. This is in line with design-based research (DBR) which emphasizes theory driven-design [30, 31]. This means that the development team needs to have a common understanding of the theoretical underpinnings that are needed to develop the tasks [32]. On the other hand, DBR is conducted in order to generate, advance and refine theory [30, 31]. As hypotheses are rejected or confirmed in DBR, theoretical models get refined/retheorized [33, 34]. Thus, DBR is based on established theories while allowing for theory generation and modification [30].

When a tool has been designed with all the aforementioned aspects in mind it needs to go through cycles of testing. Often task design issues become apparent only after it has been tested with the target group. The data gives a good indication of the appropriateness of the task design and difficulty level.

For example, if the task layout and difficulty levels are not appropriate, the risk of ceiling effects in the data increases [8]. Ceiling effects are a consistent issue in the EF literature, especially when conducting research with children [35, 36]. For example, Willoughby et al. [36] found floor and ceiling effects on several executive function tasks with children. One possible reason for this was limited task variation within the task (for more information see [36]). Similarly, Petersen et al. [37] mention variability issues with regards to executive function tasks. Petersen et al. [37] explain that ceiling and floor effects are associated with lower variability in the measured construct, which increases Type II error and reduces power to detect associations with other variables.

A few recent attempts have been made to adapt EF tasks to children by using strategies such as shortening task length and modifying the design and delivery method [9, 35, 36, 38-41]. Howard and Melhuish [42], for example, modified their EF task to suit the reaction times of young children, to test whether children's performance is affected by animated vs static stimuli and whether using an iPad versus a laptop has an effect on children's performance. The researchers found better task performance on the iPad compared to the laptop version. They attribute this finding to the additional cognitive demands that come with reorienting attention between the laptop screen and the keyboard. This cognitive demand is reduced when using an iPad because the response location and the visual display of the task are on one screen. Thus, using an iPad provided a better foundation for accurately measuring the cognitive construct (inhibition) and resulted in better reliability measures. Furthermore, the stimulus presentation time was tested to suit children's reaction times and adapted accordingly. Lastly, the inclusion of animations did not affect performance, however, it increased reliability and resulted in stronger correlations with other EF measures. This shows the importance of testing cycles with the target group in design-based research.

In line with these findings, several researchers have decided to use tablets instead of computers for cognitive assessment tasks [9, 35, 36, 40, 41]. Using a tablet instead of a computer has several advantages [43]. Firstly, the immediacy of a touch can help reduce the additional time and effort which has been identified as a result of reorienting between keyboard and screen, especially for children [42, 44, 45]. Furthermore, tablets are mobile and can therefore be applied to different contexts, which supports testing opportunities in non-clinical, real life scenarios. Additionally, tablets allow for verbal (standardized) instructions given via headphones alongside visual interactive instructions on the screen. This allows for self-assessment, which eliminates instructor bias and costly instructor training and limits the cognitive demands associated with social interactions. Finally, using an online-connected tool like a tablet enables fast data collection that can be uploaded and analyzed in a more efficient way than the traditional pen and paper recordings [36].

2   eFun

To address the aforementioned challenges and problems with existing cognitive assessments for children, this research presents the iterative design process of eFun; an app to measure executive functions in primary school children. eFun is a measurement tool for students that uses game elements to measure cognitive functions while following the guidelines of psychological theories. Miyake et al. [46] proposed a theory that states that the three core executive functions (working memory, inhibition, and cognitive flexibility) can be measured separately, despite sharing underlying cognitive mechanisms. Diamond [47] and Zelazo et al. [48] further suggest that the three core executive functions work together to engage in more complex cognitive processes such as problem-solving. Thus, when solving a task or problem, several cognitive functions are activated, enabling the solver to plan and set/reach goals. These theories form the foundation for the development of the three eFun executive functions tasks. The three games are based on validated EF tasks (for more information see [8]).

A team of developers, designers, and researchers worked together to ensure that the tasks are enjoyable and valid for children. eFun was initially tested at a primary school with students in Years 1 and 2 and was refined and enhanced accordingly [8]. These design decisions were based on the data and results [8], observations, and the students' and teachers' feedback on the tasks. The design and interface were adjusted to make it easier to use and more appealing for children. For example, hand

icons that demonstrate how the games work were added. The challenge levels were adjusted based on the students' developmental stage considering performance indicators from the first study, yet the tested constructs remained the same (i.e., working memory, inhibition, and cognitive flexibility). The modified eFun app was then tested again with the same students, 9 months later (Years 2 and 3). Results of this comparison are reported in the *Testing results* section 5 of this paper. The iterative cycles of testing, design and development are part of the design-based research approach.

3   Methodology

3.1   Design-based research

We adopt the design-based research (DBR) definition by Wang and Hannafin [30, pp. 5 and 7]. The authors define DBR as, 'a systematic but flexible methodology aimed to improve educational practices through iterative analysis, design, development, and implementation, based on collaboration among researchers and practitioners in real-world settings, and leading to contextually-sensitive design principles and theories'. Design-based research merges research and practice and requires collaborative teamwork between researchers, practitioners and developers. The methodology of the eFun project is based on the DBR approach with an iterative design process that underwent cycles of analysis and design in order to produce educational software to support teachers in their development of individualized learning plans. eFun assesses cognitive skills in students that are linked to educational outcomes (e.g. Maths and English).

3.1.1   Design-based research overview
Design-based research is:

- **Pragmatic**: Design-based research informs and improves practice [30]. New tasks are developed, new practices emerge, and new technology is employed [32]. Innovative digital applications are produced and adapted to teachers' expectations [32].
- **Theoretical:** Design is theory-driven and grounded in relevant research, theory and practice [30]. The hypotheses embedded into the learning situations are confirmed or rejected, and theoretical models are refined [32].
- **Iterative, and flexible**: Processes are iterative cycles of analysis, design, implementation, and redesign, [30] allowing flexible revisions of the design [32].
- **Contributive**: Research conducted in this way provides the opportunity to develop something new (innovative learning approach, innovative learning artefact, etc.) [32].
- **Interactive/Collaborative**: Designers are involved in the design processes and work together with participants [30]. Stakeholders develop a common view of the educational objectives addressed by the project. They build an agreement for the design of innovative tools and a common understanding of the theoretical background needed for the development of these tools [32].
- **Tested in realistic contexts**: Experimental testing of the learning artefacts by practitioners in real contexts (schools) allow the designer to take into consideration the complexity of these contexts and to collect data for the analysis phase [47]. Design is conducted in real-world settings and the design process is embedded in and studied through design-based research [30].

All of these principles were included in the DBR methodology incorporated into the development of eFun.

4   The development of eFun

As a first step, research gaps were identified based on literature reviews, and applicable theories were selected to inform the game structure, design and layout of the games. Design decisions were made based on theories of cognitive science and psychology, and user/child friendliness was identified as a key element in the design process. This informed the development of eFun and the first testing phase.

The eFun app was tested for the first time with primary school students and the initial findings have been analyzed and published [8]. The eFun games were then redesigned based on the feedback and results from the initial testing phase. This modified second version of eFun was tested again with the same students to assess whether the changes made yielded the desired results.

The eFun app consists of three child-friendly cognitive games that measure the three core executive functions working memory, inhibition and cognitive flexibility. The next section describes the games in more detail, and how the cycles of testing, redesign, and analyses enabled flexible revisions to the app [31, 32].

4.1   Overview of the core eFun games

This section will briefly explain the core eFun games. The structure and story of the core eFun games remained the same throughout the cycles of design. An explanation of the changes that were made to the game elements can be found in the following sections.

The *Ice Steps* game is based on the backward Corsi Block test [49]. The students are asked to remember a sequence of ice floats in backwards order as a measure of working memory. In this game, *Pongo* the penguin crosses a river on ice floats and the student is asked to bring Pongo back to the other side. The story behind this game is that Pongo must collect fish from one side of the river and get back to the other side of the river to feed hungry chicks. If the child gets three trials in one level wrong, the game is discontinued, with a rewarding screen showing the chicks being fed with the collected fish. This rewarding screen reduces the chance that the child gets frustrated or bored because the game is too difficult or too easy [29].

The *Log Chop* game is based on established Go/No-Go tasks [42, 50-52]. The students are asked to chop (swipe) vertically falling logs but avoid chopping icicles. The control required to withhold the response in the presence of an icicle allows for an assessment of inhibition. Difficulty increases as the speed of the logs and icicles falling increases along with the variety of the items' location. The story behind this game is to chop logs in order to make a fire to keep the eFun villagers warm.

The *Ice Cube Sorting* game is based on adapted card sorting tasks [38, 53] and aims to measure cognitive flexibility by switching between sorting rules. The students are asked to sort ice cubes into four tubes according to three changing rules (color, shape, and number), each displaying a variation in type of fruit, color and quantity (1-4) (e.g. one red apple, two green pears etc.). The story behind this game is that *Eski* the husky must store (and sort) ice cube fruits in order to prepare for an upcoming cold storm, during which he cannot leave the house to get food.

A questionnaire at the end of each game was used to gain students' feedback on the games. After each game, the students were asked to fill out a brief questionnaire with seven questions evaluating how enjoyable, fun, exciting, easy, hard, boring and frustrating they found the games. The questionnaire is embedded within the app after each of the other games on the iPad. The questions were simplified and presented one at a time to better accommodate the target audience. All questions and answers are read out to the child along with the text on screen. The questionnaire is based on the Intrinsic Motivation Inventory [54, IMI], which is a multidimensional measurement device intended to assess participants' subjective experience on a target activity. The IMI has been used in the context of intrinsic motivation and self-regulation assessment, and includes questions assessing interest and enjoyment. For the purpose of this project with students, questions from the interest/enjoyment scale were adapted and a 4-point response scale was used: 'no, not at all', 'a little bit', 'quite a bit' and 'yes, a lot'. A very similar type of response scale has previously been shown to be clear and useful for studies involving young children [55].

4.2   Version 1 of eFun

For a detailed description of the eFun Version 1 games please see Berg, et al. [8]. The first version of eFun was developed with a team of game developers, 3D designers and researchers using the game engine Unity. It was decided that all games should be contained in one app to allow for a quick and smooth assessment. In order to assist children's limited reading abilities, it was decided that all instructions would be verbalized and visual cues would be added for additional support, allowing for self-administration without assistance. A guiding character was developed called *Owly*. Owly appears in the left corner of the screen to give instructions and tell the story of each game. The student hears the instruction that is simultaneously displayed in writing in a speech bubble on the screen. Instructions and stories cannot be skipped. Stories/storyboards and goals for each game were developed to enhance motivation and engagement with the app. The assessment duration was kept to a minimum to reduce cognitive load. A procedural plan for the application of eFun through testing cycles with multiple primary schools was developed.

Interface decisions were based around user-friendliness and appeal to children. For example, it was decided to display a top-down overview map screen that includes locations for each game. The map appears between games to indicate the student's progress. The husky character Eski walks from one game location to the next during transitions between games. The characters and environment were developed with a 'winter wonderland' theme (see fig. 1 for an example).

In order to minimize frustration, some games end early based on the performance of the student. However, other games require a full set of data and therefore were decided to have standardized lengths. It was decided that a displayed score count would increase competition and might distract the child from the game itself [28], therefore verbal and visual forms of encouragement and praise were agreed on. Overall, all games were designed to be as simple and straight-forward as possible, while including subtle elements to keep the students engaged. For example, the background in the inhibition game was kept neutral to avoid it distracting the student from the stimuli. Yet, to make it more engaging, sounds and verbal feedback were added for assistance and encouragement. A meta-analysis by Deci et al, 2001 showed that verbal rewards enhanced intrinsic motivation. For example, audio feedback was given for swiping the correct stimuli in the inhibition game, and a chopping sound occurred when swiping the wood (correct stimuli). Different audio feedback was given for swiping the incorrect stimuli, such as a crystal sound when swiping the icicles. Characters were animated to make them more appealing to children. Practice phases were included to ensure that students understood how the game worked before they started playing. Furthermore, it was decided to end all games with an achievement screen showing how the child has reached the goal (e.g. a fire with the swiped logs).

To get feedback from the students during the studies, we employed a questionnaire assessing how fun, enjoyable, exciting, boring, frustrating, easy, and hard the child found their experience with the games. This questionnaire appears after every game on the iPad. The questions and answers are read out to the children and are repeated when the child clicks on them. Additionally, a focus group was conducted with each participating class in order to give the children the chance to express their opinion on the games and ask questions where needed. Finally, educators, developers and researchers were asked to play the games to provide a further source of feedback.

During internal pilot testing we observed that adults got bored/frustrated with some of the games. This was in line with their feedback that some of the games were too slow. We therefore increased the speed slightly to maintain an appropriate pace, but not more than would be suitable for children. Furthermore, we noticed that children and adults in the internal pilot testing found it hard to understand some of the games, therefore instructions were improved to be more child-friendly. Spelling mistakes, bugs, and mechanics (e.g. visual feedback to swiping logs and icicles in line with touch) were fixed before version 1 was released.

4.3   Testing eFun Version 1 with students

The first version of eFun was tested with primary school students in Years 1 and 2. For a detailed description of the results please see Berg et al. [8]. The feedback questionnaire results showed that overall, the students enjoyed playing the eFun games. The Log Chop game was the most liked, the Ice Steps game was the second most liked game and the Ice Cube Sorting game was the least liked game. The focus group was in line with the questionnaire feedback, showing that the majority of children were

frustrated with the Ice Cube Sorting game because they found it hard or didn't understand it. The students found Owly's instructions during the Ice Cube Sorting game to be repetitive, annoying and distracting. To reduce the number of verbal instructions, it was decided to add visual cues to the next version of eFun.

When the students were asked how they knew that they were doing well in the game, they noted praise such as 'good job' and 'well done' or the ice steps turning green to indicate that they were doing well. This showed us that we were able to communicate success without a score, so we kept this approach for the next development cycle.

Most students understood what they needed to do in order to get to the goal in the game (e.g. chop logs to make a fire). The focus group also revealed that Eski the huski was the most liked character, therefore it was decided to make him the main character of the next game that was to be developed (a problem-solving game). Furthermore, the students reported that they wanted to help the characters in the game and that they felt in control in most of the games. However, some students felt less in control during the Ice Cube Sorting game.

4.4 eFun redesign (Version 2)

As a result of feedback, data analysis, and observations undertaken during study 1, changes were made to make eFun more child and user friendly. Overall, instructions were made clearer and easier to understand. Owly's voice was changed, visual cues were added, and the difficulty levels of the games were adjusted where needed. A demonstration phase and hand icon were added to visually show students how to play the game and to explain how the mechanics work (e.g. tapping ice floats, swiping logs). Additionally, further sound effects and animations were added to enhance appeal. For example, correct responses were accompanied by a high-pitched ascending tone, whereas incorrect responses were accompanied by an error sound. Character animations were also enhanced, with Owly*'s* explanations being accompanied by an animated beak when talking and *Pongo* the penguin sliding to some locations instead of walking.

4.4.1 eFun Ice Steps game (Working Memory)

The main problem that was observed during this game was the difficulty level of the practice phase. Students found the practice phase in this game very difficult. Some students needed assistance to get through the last trials of the practice phase. The practice phase required the student to remember up to 5 floats in reversed order correctly to progress to the game. Therefore, the practice phase was adapted from the initial/first version of this game described in Berg et al. [8]. The updated version requires the participant to only remember three ice floats in the practice phase. Furthermore, an animated hand was added to the practice phase to demonstrate tapping the ice floats to move *Pongo* the penguin (see fig. 1).

The speed of the game was slightly increased to make it more engaging/dynamic (Pongo moves faster). Speeding up the game also decreased the time the sequence of ice floats had to be held in mind/working memory, accounting for a student's attention span. However, the speed was adapted only slightly to avoid making it too easy and to account for students' reaction time abilities.

Sound and visual effects were added to capture children's attention and to serve as a reward. For example, particle effects and sounds were added when Pongo is jumping from float to float. Furthermore, question marks appear above Pongo's head to indicate that it is the student's turn to decide which ice steps to tap on, and a popping sound was added to sustain the student's attention to the game. If the student taps on a *wrong* float, Pongo falls into the water (ice float breaks), whereas before Pongo would just reappear on land blinking. Additionally, error sounds (Pongo call of dismay) and a water sound were added when Pongo falls into the water.

When the student crosses the river on the correct ice floats an enlarged fish appears as a visual reward for successfully bringing fish for the chicks to the other side of the river. Additionally, a happy sound appears when the fish is dropped off to the little chicks. This serves as an encouragement to motivate the student to continue with the game and collect more fish for the hungry chicks. The student also receives a rising pitch feedback sound (jingle sound) when tapping on the correct ice floats that

aims to serve as an auditory reward. The student cannot see the amount of fish that has been collected. This was a careful consideration of the researchers to avoid competition between classmates which could interfere with the assessment process and result in frustration.

Additionally, Pongo was reanimated with improved design and new animations such as a sliding animation to make it more interesting to watch. Besides, the little chicks were further animated and shadows were added to the characters and items in the game.

Figure 1. *eFun Ice Steps Game measuring working memory, showing the demo phase in which the student learns how to bring Pongo to the other side of the river to feed the hungry chicks.*

4.4.2 *eFun Log Chop game (inhibition)*

The main problem with the inhibition game became apparent during the data analysis. The results of the questionnaire indicated that the students enjoyed this game the most, however, the performance data showed that the game was too easy for the students which resulted in ceiling effects (a large number of high scores). Therefore, this game was adjusted to be more difficult. The stimuli (icicles and logs) were changed to look more alike, with more similar colours, widths, and shapes. Additionally, the size of the stimuli was decreased slightly to make it more difficult.

Furthermore, a fourth level with varying inter-stimulus intervals (ISI) was added, and the number of stimuli per level was increased from 25 to 30 to have more outcome data. Having more stimuli was thought to improve the variability of the data. Additionally, a demonstration hand was added to show the student how to swipe the logs (see fig. 2), and how to inhibit swiping the icicles.
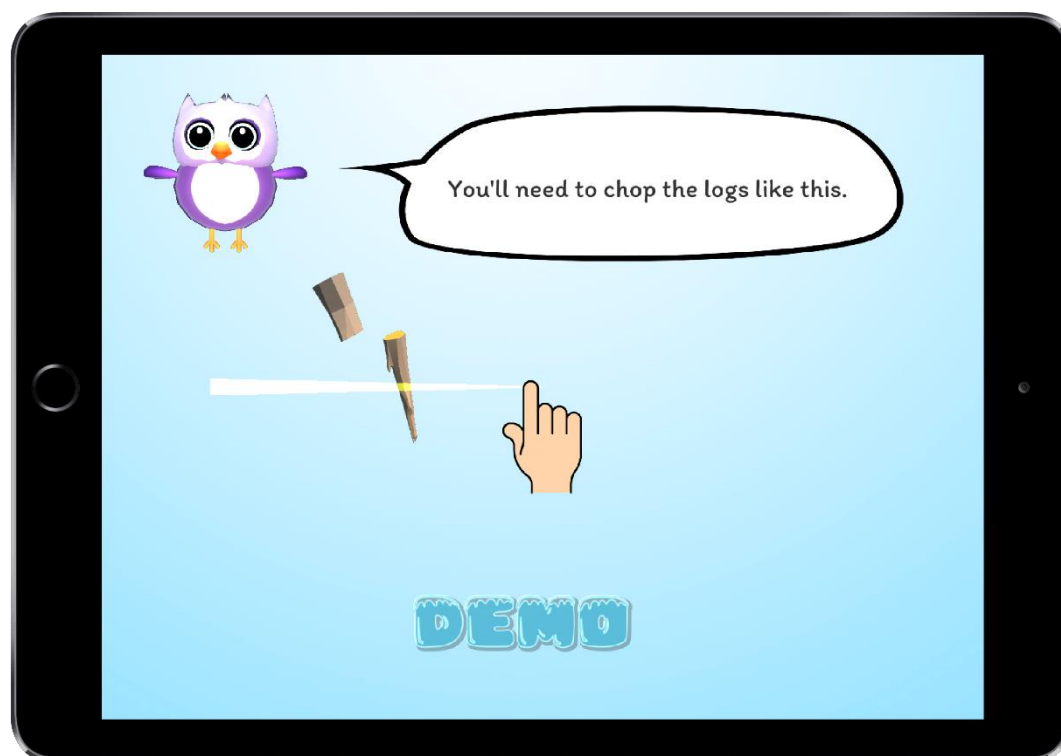
Figure 2. *eFun Log Chop game measuring inhibition showing the demo phase in which the student sees how the logs (go stimuli) get chopped.*

Additionally, the background colour was changed from grey to a light blue gradient to look more appealing while not being too distracting. Dockterman et al. [29] recommend minimizing distracting graphics to avoid drawing student's attention away from the assessed skill (inhibition). Lastly, increasing pitch sounds were added to the Log Chop game as an auditory reward when swiping the correct stimuli (logs), and the crystal sound for swiping the icicles remained.

4.4.3   eFun Ice Cube Sorting game (cognitive flexibility)

The main problem with this game was that the instructions were difficult to understand and confusing for some students. Students reported that the repetition of instructions and feedback were frustrating. Therefore, instructions were adjusted to be easier to understand, and visual cues and demonstrations were added to reduce the amount of verbal feedback and instructions. In the initial study [8] we observed that some participants kept forgetting about the three rules. Therefore, feedback on the sorting rules was added to reduce memory demands. For example, the student was reminded of the rules when sorting according to a wrong rule: when *colour* was chosen incorrectly, the explanatory character would say: 'That's not right! We're not sorting by *colour* this time.' In the initial version of eFun the student was only told that the move was incorrect but not why ('That's wrong, try another rule.').

Additionally, visual cues appear on the sorting tubes after sorting, whereas in Version 1 only verbal feedback was given. For a correct sort, the sorting rule is displayed as a symbol (e.g. fruit) with a green tick (see fig.4), and for a wrong sort, a red circle with a line appears on the tube. For example, when correctly sorting according to the rule *type/shape* of fruit, fruit images with a green tick are displayed on the tube that matches the shape of the fruit in the ice cube. However, if the student incorrectly sorts according to the rule *colour*, a red cross symbol is displayed on the tube. Along with this, audio cues were added for correct and incorrect sorts. An error sound appears for wrong sorts and a rewarding

jingle with increasing pitch (same sound as in the other two games) appears for correct sorts. These cues were added to avoid verbal repetitions. For example, the cues for a correct sort replace the repetitive verbal feedback 'that's right, well done!'.

In Version 2, Owly and the speech bubble disappear after instructions/feedback are given, rather than staying on screen (as in Version 1). They only come back for inactivity reminders if the student does not respond. This eliminates distractors and is supposed to help the student focus on the task itself. Additionally, terminologies were adjusted to suit the student's developmental level. For example, the term *rule* was changed to *how we sort*: 'The rule has changed' was replaced with 'we've changed how we sort the fruit'. These and other script changes make it easier for the student to understand the game. Lastly, a demonstration phase with a visual hand was added to show the student how to drag the ice cubes into the tubes.

Figure 3. *eFun Ice Cube Sorting game measuring cognitive flexibility showing the practice phase. The ice cube needs to be sorted into one of the tubes. If the correct rule was 'color', the cube would need to be sorted into the second tube, for the rule 'shape' it needs to be sorted into the very left tube (as shown) and for the rule 'number' it would need to be sorted into the very right tube.*

4.4.4 Feedback questionnaire

To avoid students skipping through the questionnaire without listening to the questions and answers, the option to select an answer while Owly is talking was disabled. In Version 1 the student was able to interrupt the verbal instructions by clicking on an answer, which moved them onto the next question. Additionally, wording was changed from *task* to *game* to make it sound more appealing to the student.

5 eFun (Version 2) testing results

After the adjustments were made, the second version of eFun was tested again 9 months later with the same primary students as in the initial study. Before the second study commenced, approval from both the Edith Cowan University Ethics Committee and the participating school was sought. Information and consent forms were again sent to the school principal, the teachers and parents. The teachers distributed information and consent forms for the children to parents. The information letters outlined the procedure, possible risks, and purpose of the study. Results between the two testing sessions were analyzed to assess differences between the two data sets, including feedback questionnaires, performance data and focus groups. Overall, students equally enjoyed playing the games and reported no difference between the two eFun versions. However, the performance data indicated that the changes to the eFun app resulted in different scores. For the raw data files see https://figshare.com/projects/A_Game-Based_Online_Tool_to_Measure_Cognitive_Functions_in_Students/97636

5.1   Student evaluation

After each game, the students were asked to fill out a brief questionnaire with seven questions evaluating how enjoyable, fun, exciting, easy, hard, boring and frustrating they found the games. Answers were given on a 4-point response scale ranging from 'no, not at all', 'a little bit', 'quite a bit' to 'yes, a lot'. The ratings for the adjectives enjoy, fun, exciting, and boring (reverse scored) were combined into an enjoyment score to compare the overall enjoyment for each task. A reliability analysis was carried out on the task rating adjectives comprising 4 items. The Cronbach alpha reliability coefficient for the adjectives enjoy, fun, exciting, and boring ranged between 0.86 and 0.69. Therefore, it was deemed appropriate to combine these evaluations into an enjoyment score.

The results showed that students still enjoyed playing the eFun games, with the Log Chop game being the most enjoyed game and the Ice Cube sorting game the least enjoyed game. In order to compare students' enjoyment ratings of the eFun games from the initial to the second study, a paired sample t-test was conducted. The t-test showed that there was no significant difference between the scoring of the three eFun games from the initial study to the second study (Ice Steps: t (66) = -1.17, p = .247, d = .166, M_{s1} = 3.26, SD_{s1} = 0.77, M_{s2} = 3.38 SD_{s2} = 0.66; Log Chop: t (66) = -1.78, p = .080, d = .208, M_{s1} = 3.73 , SD_{s1} = 0.47, M_{s2} = 3.73, SD_{s2} = 0.47; Ice Cube Sorting: t (64) = -.95, p = .346, d = .145, M_{s1} = 3.08, SD_{s1} = 1.0, M_{s2} = 3.21, SD_{s2} = .85). This means that students' enjoyment ratings of the eFun games remained similar across the two testing phases.

When looking at the individual questions of the evaluation (enjoyable, fun, exciting, easy, hard, boring, and frustrating) from the initial to the second study, a paired sample t-test showed that there was only a significant difference in two of the Log Chop questions and in one of the Ice Steps questions. More students rated the Log Chop game as *too easy* in the initial study as compared to the second study (t_{66} = -4.10, p < .001). Furthermore, more students rated the Log Chop game as *boring* in the initial study as compared to the second study (t_{66} = -2.95, p = .004). Lastly, for the Ice Steps game, more students answered that they enjoyed the game in the second study than in the initial study (t_{66} = 2.08, p = .041).

However, it is important to also consider the meaning of the mean scores for the answers. For the Log Chop question 'Do you think the game was too easy?', the mean was 3 (SD = 1.28) in the initial study and 2.24 (SD = 1.17) in the follow-up study, meaning most students answered that the Log Chop task was *too easy* with *quite a bit* in the initial study and *a little bit* in the second study.

The Log Chop question 'Do you think the game was boring? had a mean score of 1.50 (SD = 1.06) in the initial study and 1.15 (SD = .47) in the follow-up study. This means that most students responded *No, not at all* (scored as 1) to this question after playing the Log chop game.

Similarly, it is important to consider the means for the Ice Steps game. The Ice Steps evaluations for 'Did you enjoy the game?' were high with a mean of 3.3 (SD = .95) for the initial study and a mean of 3.5 (SD = .70) for the second study. This means that most students answered *quite a bit* to *Yes, a lot* to this question after playing the Ice Steps game.

5.2   *Difference between students' game performances from the first to the second study*

To test the difference between the game performance scores from the initial study to the second study, a paired sample t-test was applied. It was expected for the Ice Steps game to have higher scores in the second study because the practice phase was made easier, thus potentially reducing frustration and increasing motivation with the game. Furthermore, the Log Chop game was expected to have lower scores than in the initial study because this game was made more difficult to eliminate ceiling effects and improve the distribution of scores. For the Ice Cube sorting game, higher scores were expected due to clearer instructions and more visual cues.

Paired sample t-test results showed that there was a significant average difference between the eFun game performance of the Ice Steps game (M_{s1} = 16.68, SD_{s1} = 10.13, M_{s2} = 31, SD_{s2} = 17.30) and the Log Chop game (M_{s1} = 0.70, SD_{s1} = 0.11, M_{s2} = 0.32, SD_{s2} = 0.13) from the initial study to the second study (Ice Steps: t_{61} = -7.48, p < .001; Log Chop: t_{60} = 20.16, p < .001). On average, Ice Step scores were 14.37 points higher in the second study than in the initial study (95% CI [18.21, 10.53]). For the Log Chop task, the scores in the follow-up study were on average .47 points lower (95% CI [-.43, -.52]) than in the initial study. No significant difference in performance scores were found for the Ice Cube Sorting game from the initial to the second study.

6   Discussion

In order to improve the eFun games, the first version and the modified version of eFun (Version 2) were tested in two studies with the same students. Based on initial data and student/teacher feedback, the app was modified to be more child-friendly, with an aim of increasing the validity of the tasks. The second study reinvestigated students' ratings of the games and the performance scores using the modified version (Version 2). The results were in line with expectations; overall students still enjoyed playing the eFun games in the second study, meaning that students rated the eFun games as equally enjoyable as in the first study. Thus, enhancing some of the game features (e.g. adding sound and visual effects) did not seem to have significantly changed the students' overall enjoyment ratings of the games. The time in-between testing (9 months) was expected to be too long for the students to remember design details, and the core mechanics and background stories of the games remained the same. Furthermore, it is possible that our game-like features were not salient enough to make a difference. Only subtle changes were made to the design and the main task structure remained the same.

Furthermore, based on the initial study with eFun [8] it was expected that not all games would be rated as equally enjoyable (composite enjoyment score) in the second study. As with the initial study, the eFun Log Chop game was rated as more enjoyable than the other two eFun games, with the eFun Ice Cube Sorting game being the least enjoyable game. The Log Chop game was the most fast-paced and simple of the three eFun games, whereas the Ice cube sorting game was the most complex game and had fewer dynamic elements than the Log Chop game.

Thus, we suggest that complexity, difficulty, and pace influenced the perceived overall enjoyment of the eFun games. A simple game can often be successful in drawing people's attention in; for example, *Tetris* is simple, yet engaging. Tetris has a very simple layout and the rules are easy to understand, similar to the eFun Log Chop game. Thus, it is important to avoid distracting graphics and make instructions as simple and straightforward as possible to reduce complexity, and focus the students' attention on the core task [29]. The Log Chop task achieves this by implementation of simple rules and minimalistic design elements in a 3D environment (only 2 stimuli and the background). As explained earlier, the background in the Log Chop game was kept simple to avoid distraction. The above rating results are related to the composite enjoyment score (combining answers for enjoy, fun, exciting, and boring (reverse scored)).

However, when comparing the difference between the individual answers in the two studies we found that fewer students rated the Log Chop as *too easy* in the second study compared to the initial study. Commensurately, fewer students rated the Log Chop task as *boring* in the second study as

compared to the initial study. The Log Chop game was made more difficult by adjusting the two stimuli so that they looked more alike, which made them harder to differentiate. Furthermore, the number of stimuli in each level was increased and a level with increased difficulty was added (varied inter-stimulus intervals and increase in number of stimuli). Therefore, fewer students experienced the game as too easy in the second study. The Log Chop game was also rated as less boring in the second study. This can be explained with the flow theory which states that tasks should neither be too easy (boring) nor too difficult (frustrating) when aiming to increase enjoyment and immersion [56, 57]. This is in line with literature stating that participants were more engaged when the perceived challenge of the task was high, along with other factors such as relevant instructions, controlled learning environment, and a high skill set [58].

With regards to the Ice Steps game, more students answered that they enjoy the game in the second study than in the initial study. We expect this can be attributed to the facilitation of the practice phase in the second study compared to the initial version. In summary, students' overall enjoyment rating of the tasks was high. Design modifications did not seem to affect students' evaluations of the games, however, modifications to the difficulty level were noticeable to the students.

Modifications to the difficulty levels also affected the students' performance scores. For the Ice Steps game with an easier practice phase in the second study, the students' overall performance increased from study 1 to study 2. This is in line with research suggesting that if the student finds the task interesting and enjoys it, the performance score on the task is higher [23]. Furthermore, researchers suggest that when a child is able to solve a task, or at least work on it without facing major difficulties, motivation increases and anxiety decreases [59, 60]. Consequently, this may improve overall performance on the task [61] [62]. Thus, facilitating the practice phase in the Ice Steps game is likely to have contributed to an increase in performance scores and enjoyment, possibly due to better concentration and motivation in the subsequent game. However, for the Log chop game, which was made and experienced as more difficult in study 2, the overall performance scores were lower in study 2 compared to study 1. The change in the overall performance is likely to be a result of the adjustments that were made to the game structure (increasing stimuli, facilitating practice phase by reduction of items to be remembered and so on), as the structure of a task is seen as one of the most important components of a game [4, 63]. Thus, the researchers' intention of increasing/decreasing difficulty levels via adjustments made to the game structure was successful.

The Ice Cube Sorting game did not show any significant differences in students' feedback ratings or performance data. Yet, a significant amount of work went into trying to improve instructions for this game (simplified language and added visual cues). This is in line with research that found no difference in data between tasks with added game elements vs task without added game elements [4]. Similar to the rating of the Ice Cube Sorting game, participants in Hawkins' study rated both versions of the tasks as equally boring and repetitive [4]. However, the game-like versions of the tasks in Hawkins' study were rated as more interesting and enjoyable, whereas in our study no difference in enjoyment ratings for the Ice Cube sorting game was found. Overall, this indicates that including game mechanics does not necessarily invalidate the data of cognitive tasks and has the potential to increase engagement. However, future research is needed to clearly separate the effects of task modifications and the addition of game-like elements.

7   Conclusion

Traditional cognitive assessments are often lengthy, repetitive, and tiring [3]. In response, eFun has been developed as a novel and engaging measurement tool to assess primary school-aged children's executive functions. The app was developed through a design-based research methodology to develop three minigames that each aim to assess a different executive function (i.e., working memory, inhibition, and cognitive flexibility). The iterative development approach provided a basis for constant evaluation and modification of the game and led to a range of findings that improved the game but also provided some recommendations that can be generalized beyond this particular product. Specifically:

- Sustain student's attention: distracting graphics were avoided [64, 65]. The tasks are short with subtle game elements and a narrative that runs through all games.
- Create an appropriate level of challenge: difficulty levels were adjusted to suit students' developmental levels.
- Give positive feedback: all games include rewarding end screens and positive feedback. Sounds, narration, and visuals focus on emphasizing positive feedback and achievements rather than punishing incorrect moves [66].
- Reduce test-anxiety: there is no judgment of the test-taker, which aims to reduce anxiety and stress. The games are self-administered in a familiar context.
- Indicate progress without scores: a map appears between games showing movement completion by displaying the games that have been complete and that are to come. The map screen highlights progress toward an endpoint, which can enhance sustained engagement and a feeling of achievement [67]. Scores are not displayed in order to avoid frustration, competition, and speeding through the game to gain extrinsic rewards [29].
- Maximize the potential of a game environment for self-assessment: no instructor or assessor is required to be present during the assessment.
- Utilize mobile devices' potential for implementation in real life contexts: the games are played in contexts that are familiar to the student (i.e. schools).

The results indicate that the eFun games offer an enjoyable experience for students, and that adjustments to the structure of the game/underlying task rather than the incorporation of gamified elements are more likely to affect the performance data. Future studies with eFun will examine the difference between traditional cognitive tasks and the eFun games.

References

[1] E. A. Berg, "A Simple Objective Technique for Measuring Flexibility in Thinking," *The Journal of General Psychology,* vol. 39, no. 1, pp. 15-22, 1948, doi: 10.1080/00221309.1948.9918159.

[2] A. Diamond and D. S. Ling, "Conclusions about interventions, programs, and approaches for improving executive functions that appear justified and those that, despite much hype, do not," *Dev Cogn Neurosci,* vol. 18, pp. 34-48, Apr 2016, doi: 10.1016/j.dcn.2015.11.005.

[3] J. McPherson and N. R. Burns, "Assessing the validity of computer-game-like tests of processing speed and working memory," *Behavior Research Methods,* vol. 40, no. 4, pp. 969-981, 2008, https://doi.org/10.3758/BRM.40.4.969

[4] G. E. Hawkins, B. Rae, K. V. Nesbitt, and S. D. Brown, "Gamelike features might not improve data," *Behavior research methods,* vol. 45, no. 2, pp. 301-318, 2013, https://doi.org/10.3758/s13428-012-0264-3

[5] A. Lueckenhoff, C. Wessels, M. Kyrarini, and F. Makedon, *Towards Game-based Assessment of Executive Functions in Children.* 2020, doi: 10.1371/journal.pone.0230498

[6] V. J. Shute, L. Wang, S. Greiff, W. Zhao, and G. Moore, "Measuring problem solving skills via stealth assessment in an engaging video game," *Computers in Human Behavior,* vol. 63, pp. 106-117, 2016, https://doi.org/10.1016/j.chb.2016.05.047

[7] J. F. Vermeir, M. J. White, D. Johnson, G. Crombez, and D. M. L. Van Ryckeghem, "The effects of gamification on computerized cognitive training: Systematic review and meta-Analysis," *JMIR serious games,* vol. 8, no. 3, p. e18644, 2020, https://doi.org/10.2196/18644

[8] V. Berg, S. L. Rogers, M. McMahon, M. Garrett, and D. Manley, "A Novel Approach to Measure Executive Functions in Students: An Evaluation of Two Child-Friendly Apps," *Frontiers in Psychology,* vol. 11, p. 1702, 2020, https://doi.org/10.3389/fpsyg.2020.01702

[9] S. J. Howard and E. Melhuish, "An Early Years Toolbox for Assessing Early Executive Function, Language, Self-Regulation, and Social Development: Validity, Reliability, and Preliminary Norms," *Journal of Psychoeducational Assessment,* vol. 35, no. 3, pp. 255-275, Jun 2017, doi: 10.1177/0734282916633009.

[10] J. Lumsden, E. A. Edwards, N. S. Lawrence, D. Coyle, and M. R. Munafò, "Gamification of cognitive assessment and cognitive training: a systematic review of applications and efficacy," *JMIR serious games,* vol. 4, no. 2, p. e11, 2016, https://doi.org/10.2196/games.5888

[11] C. Blair, P. D. Zelazo, and M. T. Greenberg, "The measurement of executive function in early childhood," *Dev Neuropsychol,* vol. 28, no. 2, pp. 561-71, 2005, doi: 10.1207/s15326942dn2802_1.

[12]   M. Ninaus *et al.*, "Game elements improve performance in a working memory training task," *International journal of serious games,* vol. 2, no. 1, pp. 3-16, 2015, https://doi.org/10.17083/ijsg.v2i1.60

[13]   S. Dovis, S. Van der Oord, R. W. Wiers, and P. J. Prins, "Improving executive functioning in children with ADHD: training multiple executive functions within the context of a computer game. a randomized double-blind placebo controlled trial," *PLoS One,* vol. 10, no. 4, p. e0121651, 2015, doi: 10.1371/journal.pone.0121651, https://doi.org/10.1371/journal.pone.0121651

[14]   T. E. Moffitt *et al.*, "A gradient of childhood self-control predicts health, wealth, and public safety," *Proc Natl Acad Sci U S A,* vol. 108, no. 7, pp. 2693-8, Feb 15 2011, doi: 10.1073/pnas.1010076108.

[15]   I. Nayfeld, J. Fuccillo, and D. B. Greenfield, "Executive functions in early learning: Extending the relationship between executive functions and school readiness to science," *Learning and Individual Differences,* vol. 26, pp. 81-88, 2013, doi: 10.1016/j.lindif.2013.04.011.

[16]   C. Blair and R. P. Razza, "Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten," *Child development,* vol. 78, no. 2, pp. 647-663, 2007, https://doi.org/10.1111/j.1467-8624.2007.01019.x

[17]   P. D. Zelazo and S. M. Carlson, "The neurodevelopment of executive function skills: Implications for academic achievement gaps," *Psychology & Neuroscience,* vol. 13, no. 3, p. 273, 2020, https://doi.org/10.1037/pne0000208

[18]   E. Kroesbergen, J. Van Luit, E. Van Lieshout, E. Van Loosbroek, and B. Van de Rijt, "Individual differences in early numeracy: The role of executive functions and subitizing," *Journal of Psychoeducational Assessment,* vol. 27, no. 3, pp. 226-236, 2009, https://doi.org/10.1177/0734282908330586

[19]   S. W. Toll, S. H. Van der Ven, E. H. Kroesbergen, and J. E. Van Luit, "Executive functions as predictors of math learning disabilities," *Journal of learning disabilities,* vol. 44, no. 6, pp. 521-532, 2011, https://doi.org/10.1177/0022219410387302

[20]   S. Shaul and M. Schwartz, "The role of the executive functions in school readiness among preschool-age children," *Reading and Writing,* vol. 27, no. 4, pp. 749-768, 2014, https://doi.org/10.1007/s11145-013-9470-3

[21]   J. A. Welsh, R. L. Nix, C. Blair, K. L. Bierman, and K. E. Nelson, "The development of cognitive skills and gains in academic school readiness for children from low-income families," *Journal of educational psychology,* vol. 102, no. 1, p. 43, 2010, https://doi.org/10.1037/a0016738

[22]   J. DeRight and R. S. Jorgensen, "I just want my research credit: Frequency of suboptimal effort in a non-clinical healthy undergraduate sample," *The Clinical Neuropsychologist,* vol. 29, no. 1, pp. 101-117, 2015, https://doi.org/10.1080/13854046.2014.989267

[23]   S. Schukajlow and A. Krug, "Are interest and enjoyment important for students' performance," in *Proceedings of the Joint Meeting of PME*, 2014, vol. 38, pp. 129-136.

[24]   J. Reeve, "The interest-enjoyment distinction in intrinsic motivation," *Motivation and emotion,* vol. 13, no. 2, pp. 83-103, 1989, https://doi.org/10.1007/BF00992956

[25]   J. B. Engelmann and L. Pessoa, "Motivation sharpens exogenous spatial attention," 2014, https://doi.org/10.1037/2333-8113.1.S.64

[26]   J. B. Engelmann, E. Damaraju, S. Padmala, and L. Pessoa, "Combined effects of attention and motivation on visual task performance: transient and sustained motivational effects," *Frontiers in human neuroscience,* vol. 3, p. 4, 2009, https://doi.org/10.3389/neuro.09.004.2009

[27]   L. Pessoa and J. B. Engelmann, "Embedding reward signals into perception and cognition," *Frontiers in neuroscience,* vol. 4, p. 17, 2010, https://doi.org/10.3389/fnins.2010.00017

[28]   B. Katz, S. Jaeggi, M. Buschkuehl, A. Stegman, and P. Shah, "Differential effect of motivational features on training improvements in school-based cognitive training," *Frontiers in human neuroscience,* vol. 8, p. 242, 2014, https://doi.org/10.3389/fnhum.2014.00242

[29]   D. Dockterman, Y. Petscher, A. McAfee, E. Klopfer, S. Osterweil, and C. Diefenthaler, "Gaming Considerations for Educational Assessments," 2020, https://doi.org/10.31234/osf.io/en23t

[30]   F. Wang and M. J. Hannafin, "Design-based research and technology-enhanced learning environments," *Educational technology research and development,* vol. 53, no. 4, pp. 5-23, 2005, https://doi.org/10.1007/BF02504682

[31]   C. Design-Based Research, "Design-based research: An emerging paradigm for educational inquiry," *Educational Researcher,* vol. 32, no. 1, pp. 5-8, 2003, https://doi.org/10.3102/0013189X032001005

[32]   L. Oubahssi, C. Piau-Toffolon, G. Loup, and E. Sanchez, "From Design to Management of Digital Epistemic Games," *International Journal of Serious Games,* vol. 7, no. 1, pp. 23-46, 2020, https://doi.org/10.17083/ijsg.v7i1.336

[33]   P. Bell, C. M. Hoadley, and M. C. Linn, "Design-based research in education," *Internet environments for science education,* vol. 2004, pp. 73-85, 2004, https://doi.org/10.4324/9781410610393

[34]   G. Stahl, *Computer Support for Collaborative Learning: Foundations for a Cscl Community (Cscl 2002 Proceedings)*. Routledge, 2002, https://doi.org/10.3115/1658616

[35]   K. Holmboe, C. Larkman, C. de Klerk, A. Simpson, C. Christodoulou, and H. Dvergsdal, "The Early Childhood Inhibitory Touchscreen Task," 2019, https://doi.org/10.31234/osf.io/k7g4a

[36]   M. T. Willoughby, B. Piper, D. Kwayumba, and M. McCune, "Measuring executive function skills in young children in Kenya," *Child Neuropsychology,* vol. 25, no. 4, pp. 425-444, 2019, https://doi.org/10.1080/09297049.2018.1486395

[37]   I. T. Petersen, C. P. Hoyniak, M. E. McQuillan, J. E. Bates, and A. D. Staples, "Measuring the development of inhibitory control: The challenge of heterotypic continuity," *Developmental Review,* vol. 40, pp. 25-71, 2016, https://doi.org/10.1016/j.dr.2016.02.001

[38]   Y. Kado *et al.*, "Executive function in children with pervasive developmental disorder and attention-deficit/hyperactivity disorder assessed by the Keio version of the Wisconsin card sorting test," *Brain and Development,* vol. 34, no. 5, pp. 354-359, 2012/05/01/ 2012, doi: https://doi.org/10.1016/j.braindev.2011.08.008.

[39]   A. Diamond and A. Wright, "An effect of inhibitory load in children while keeping working memory load constant," *Frontiers in psychology,* vol. 5, p. 213, 2014, https://doi.org/10.3389/fpsyg.2014.00213

[40]   P. D. Zelazo, "Executive function: Reflection, iterative reprocessing, complexity, and the developing brain," *Developmental Review,* vol. 38, pp. 55-68, 2015, doi: 10.1016/j.dr.2015.07.001.

[41]   K. Józsa, K. C. Barrett, and G. A. Morgan, "Game-like Tablet Assessment of Approaches to Learning: Assessing Mastery Motivation and Executive Functions," *Electronic Journal of Research in Educational Psychology,* vol. 15, no. 3, pp. 665-695, 2017, doi: 10.14204/ejrep.43.17026.

[42]   S. J. Howard and A. D. Okely, "Catching Fish and Avoiding Sharks: Investigating Factors That Influence Developmentally Appropriate Measurement of Preschoolers' Inhibitory Control," *J Psychoeduc Assess,* vol. 33, no. 6, pp. 585-596, Sep 2015, doi: 10.1177/0734282914562933.

[43]   G. Falloon, "Young students using iPads: App design and content influences on their learning pathways," *Computers & Education,* vol. 68, pp. 505-521, 2013/10/01/ 2013, doi: https://doi.org/10.1016/j.compedu.2013.06.006.

[44]   M. I. Posner and Y. Cohen, "Components of visual orienting," *Attention and performance X: Control of language processes,* vol. 32, pp. 531-556, 1984.

[45]   A. R. Hunt and A. Kingstone, "Covert and overt voluntary attention: linked or independent?," *Cognitive Brain Research,* vol. 18, no. 1, pp. 102-105, 2003/12/01/ 2003, doi: https://doi.org/10.1016/j.cogbrainres.2003.08.006.

[46]   A. Miyake, N. P. Friedman, M. J. Emerson, A. H. Witzki, A. Howerter, and T. D. Wager, "The unity and diversity of executive functions and their contributions to complex 'frontal lobe' tasks: A latent variable analysis," *Cognitive Psychology,* vol. 41, no. 1, pp. 49-100, 2000, doi: 10.1006/cogp.1999.0734.

[47]   A. Diamond, "Executive functions," *Annu Rev Psychol,* vol. 64, pp. 135-68, 2013, doi: 10.1146/annurev-psych-113011-143750, https://doi.org/10.1146/annurev-psych-113011-143750

[48]   P. D. Zelazo, A. Carter, J. S. Reznick, and D. Frye, "Early development of executive function: A problem-solving framework," *Review of General Psychology,* vol. 1, no. 2, pp. 198-226, 1997, doi: 10.1037/1089-2680.1.2.198.

[49]   B. Milner and L. Taylor, "Right-hemisphere superiority in tactile pattern-recognition after cerebral commissurotomy: Evidence for nonverbal memory," *Neuropsychologia,* vol. 10, no. 1, pp. 1-15, 1972/04/01/ 1972, doi: https://doi.org/10.1016/0028-3932(72)90038-3.

[50]   F. C. Donders, "On the speed of mental processes," *Acta psychologica,* vol. 30, pp. 412-431, 1969, https://doi.org/10.1016/0001-6918(69)90065-1

[51]   A. Simpson and K. J. Riggs, "Conditions under which children experience inhibitory difficulty with a "button-press" go/no-go task," *J Exp Child Psychol,* vol. 94, no. 1, pp. 18-26, May 2006, doi: 10.1016/j.jecp.2005.10.003.

[52]   S. A. Wiebe, T. D. Sheffield, and K. Andrews Espy, "Separating the fish from the sharks: a longitudinal study of preschool response inhibition," *Child Dev,* vol. 83, no. 4, pp. 1245-61, Jul-Aug 2012, doi: 10.1111/j.1467-8624.2012.01765.x.

[53]   C. Cianchetti, S. Corona, M. Foscoliano, D. Contu, and G. Sannio-Fancello, "Modified Wisconsin Card Sorting Test (MCST, MWCST): Normative Data in Children 4–13 Years Old, According to Classical and New Types of Scoring," *The Clinical Neuropsychologist,* vol. 21, no. 3, pp. 456-478, 2007/04/02 2007, doi: 10.1080/13854040600629766.

[54]   E. Deci and R. Ryan, "Intrinsic motivation inventory (IMI)," *Retrieved July,* vol. 23, p. 2006, 2005.

[55]   S. L. Rogers, L. Barblett, and K. Robinson, "Investigating the impact of NAPLAN on student, parent and teacher emotional distress in independent schools," *The Australian Educational Researcher,* journal article vol. 43, no. 3, pp. 327-343, July 01 2016, doi: 10.1007/s13384-016-0203-x.

[56] M. Csikszentmihalyi, "Flow and education," *NAMTA journal,* vol. 22, no. 2, pp. 2-35, 1997.

[57] L. S. Vygotsky, "Socio-cultural theory," *Mind in society,* pp. 52-58, 1978.

[58] D. J. Shernoff, M. Csikszentmihalyi, B. Schneider, and E. S. Shernoff, "Student engagement in high school classrooms from the perspective of flow theory," in *Applications of flow in human development and education*: Springer, 2014, pp. 475-494, https://doi.org/10.1007/978-94-017-9094-9_24

[59] M. Tenorio Delgado, P. Arango Uribe, A. Aparicio Alonso, and R. Rosas Diaz, "TENI: A comprehensive battery for cognitive assessment based on games and technology," *Child Neuropsychology,* vol. 22, no. 3, pp. 276-291, 2016, https://doi.org/10.1080/09297049.2014.977241

[60] L. S. Vygotsky, "Thought and language-Revised edition," *Cambridge, MA: Massachusetts Institute of Technology,* 1986, https://doi.org/10.1017/s0272263100008172

[61] M. N. Desrochers, M. J. Pusateri Jr, and H. C. Fink, "Game assessment: Fun as well as effective," *Assessment & Evaluation in Higher Education,* vol. 32, no. 5, pp. 527-539, 2007, https://doi.org/10.1080/02602930601116789

[62] C. L. Fales *et al.*, "Anxiety and cognitive efficiency: differential modulation of transient and sustained neural activity during a working memory task," *Cognitive, affective, & behavioral neuroscience,* vol. 8, no. 3, pp. 239-253, 2008, https://doi.org/10.3758/CABN.8.3.239

[63] R. T. A. Wood, M. D. Griffiths, D. Chappell, and M. N. O. Davies, "The structural characteristics of video games: A psycho-structural analysis," *CyberPsychology & behavior,* vol. 7, no. 1, pp. 1-10, 2004, https://doi.org/10.1089/109493104322820057

[64] J. Sweller, "Cognitive load theory, learning difficulty, and instructional design," *Learning and instruction,* vol. 4, no. 4, pp. 295-312, 1994, https://doi.org/10.1016/0959-4752(94)90003-5

[65] J. Sweller, P. Ayres, and S. Kalyuga, "Measuring cognitive load," in *Cognitive load theory*: Springer, 2011, pp. 71-85, https://doi.org/10.1007/978-1-4419-8126-4_6

[66] E. L. Deci, R. Koestner, and R. M. Ryan, "Extrinsic rewards and intrinsic motivation in education: Reconsidered once again," *Review of educational research,* vol. 71, no. 1, pp. 1-27, 2001, https://doi.org/10.3102/00346543071001001

[67] G. Siemens, D. Gašević, and S. Dawson, "Preparing for the digital university: A review of the history and current state of distance, blended, and online learning," 2015, https://doi.org/10.3102/00346543071001001