DEPARTMENT OF INFORMATICS

TECHNISCHE UNIVERSITÄT MÜNCHEN

Master's Thesis in Informatics: Games Engineering

Dynamic Difficulty Adaption for Serious Games

Florian Herber

DEPARTMENT OF INFORMATICS

TECHNISCHE UNIVERSITÄT MÜNCHEN

Master's Thesis in Informatics: Games Engineering

Dynamic Difficulty Adaption for Serious Games

Dynamische Schwierigkeitsanpassung für Serious Games

Author:	Florian Herber
Supervisor:	Prof. Gudrun Klinker, Ph.D.
Advisor:	Dipl.-Inf. David A. Plecher, M.A.
Submission Date:	14.10.2020

I confirm that this master's thesis in informatics: games engineering is my own work and I have documented all sources and material used.

Munich, 14.10.2020 Florian Herber

Acknowledgments

I want to use this opportunity to express my sincere thanks to everyone who supported me throughout the development of this thesis and my entire course of study. Most importantly, i want to thank Prof. Gudrun Klinker and David Plecher, for giving me the possibility to work on this thesis and all previous projects that led to it, for their continuous support throughout my course of study, their commitment and time investment.

Abstract

In this thesis a new viewpoint on Serious Game balancing is developed by shedding light to the differentiation between learning and gaming domain. Since the skills of a player can differ significantly in the two domains, it is necessary to treat them individually. Different levels of starting experiences and different skill acquiring rates require dynamic solutions in order to make Serious Games applicable to a broad range of users. Serious Games have the possibility to enhance learning processes around the world. Especially but not only, in situations where formal teaching is not available due to time, money or locations constraints, they can become powerful tools. For them to become personalized learning environments dynamic difficulty adjustment based on the individual treatment of the learning and gaming domain are indispensable. To accomplish this a theoretical model for the *Componentwise Serious Game Balance (CSGB)* was conceptualized. Additionally, *Componentwise Dynamic Difficulty Adjustment (CDDA)*, a dynamic difficulty adjustment based on this model, was developed to make it applicable to Serious Game design. The CDDA was then implemented in the already existing Serious Game *HieroQuest*, which is dedicated towards teaching the Middle Egyptian language. In the course of this, another game mode, which makes use of the dynamic properties to create the game world, was created. Possible affects on the learning success and player experience are investigated in a short- and long-term user study.

Kurzfassung

In dieser Arbeit wird eine neue Sichtweise auf Serious Game Balancierung entwickelt, indem die Differenzierung zwischen Lern- und Spieldomäne herausgestellt wird. Da die Fähigkeiten eines Spielers in den beiden Domänen sich signifikant unterscheiden können ist es notwendig beide unabhängig voneinander zu behandeln. Unterschiedliche Level der Anfangserfahrungen und unterschiedliche Raten des Wissensaufbaus benötigen dynamische Lösungen, um Serious Games für eine große Gruppe von Nutzern zugänglich zu machen. Serious Games haben die Möglichkeit Lernprozesse in der ganzen Welt zu verbessern. Speziell aber nicht nur an Orten die aufgrund von zeitlichen, örtlichen oder monetären Einschränkungen keine formelle Lehre erlauben können sie zu mächtigen Werkzeugen werden. Damit sie zu personalisierten Lernumgebungen werden, ist dynamische Schwierigkeitsanpassung basierend auf individueller Behandlung der Lern- und Spieldomäne unabdinglich. Um dies zu erreichen wurde ein theoretisches Model zur *Komponentenweisen Serious Game Balancierung* entwickelt. Zusätzlich, wurde eine *Komponentenweise Dynamische Schwierigkeitsanpassung* entwickelt, eine auf diesem Model basierende Dynamische Schwierigkeitsanpassung, um es auf Serious Game Design anwendbar zu machen. Diese wurde dann in ein bereits existierendes Serious Game *HieroQuest* eingebaut, welches die Mittelägyptische Sprache lehrt. Im Zuge dessen wurde auch ein zusätzlicher Spielmodus, welcher die dynamischen Gegebenheiten nutzt, um die Spielwelt dynamisch zu erstellen, entwickelt. Mögliche Auswirkungen auf den Lernerfolg und die Spielerfahrung wurden in einer Kurz- und Langzeitstudie untersucht.

Contents

1. Introduction

This chapter gives an overview of the presented work. First current challenges of Serious Game design will be explained (Section 1.1). After that the motivation to work on this field of research is presented (Section 1.2). Followed by the overall goals of this thesis (Section 1.3) and the outline (Section 1.4).

1.1. Problem Description

Relying on sophisticated learning approaches can lead to a decrease of motivation. Not only but especially in informal teaching. When learning for own further education without the help of a teacher e.g. by solely relying on books the initial enthusiasm can drastically deplete over time. Serious Games can counteract this effect, by combining immersing gameplay with learning to create an effective and engaging learning environment. Over the years many Serious Games in different genres have been developed to enhance learning content in various disciplines.

The basic intention for their use is to overcome tedious or boring phases, by introducing new motivators. But at the same time they are driven by a digital device, which removes an important aspect of traditional learning approaches, which is self-control during the learning process, mainly in terms learning speed. While reading a book, a learner can choose at anytime to step back and reread the last paragraph or even the complete chapter. In Serious Games this is most of the time not possible and therefore the difficulty of the learning content has to be treated with caution. But this is not only true for the difficulty of the learning content. Since a Serious Game combines learning and gaming, also the difficulty of the gaming content has to account for players with different levels of experience.

But the digital property of a Serious Game can also be an opportunity. To use their dynamic aspect to the fullest a personalized experience based on the players previous knowledge, their goals and their capabilities has to be created. This personalized experience has to account for knowledge, goals and capabilities regarding the learning and gaming domain, at the same time. Since both domains require very different skills, both have to be treated individually. Without adjusting the difficulty of the learning and the gaming content on a personalized basis the Serious Game can at most be balanced towards the average learner/player, but this will cause the outer edges of the spectrum to either be overwhelmed or underwhelmed by the learning and or the gaming content. Because previous experience in one of the two domains does not imply experience in the other, a Serious Game has to account for different levels of starting experience as well as different skill growing rates. Therefore this thesis conceptualizes a new viewpoint on Serious Games and the balance between challenges and skills, which treats both domains independently. To make the model applicable to dynamic Serious Game design a dynamic difficulty adjustment based on this model is developed.

1.2. Motivation

Serious Games and their implications on learning and gaming are still a young yet promising field of research. They can enhance the learning process by introducing new viewpoints and possibilities to interact with the learning content. Since they are digital games, they also allow for dynamic adjustments. Serious Games that are able to adjust themselves towards an individual's needs are a good example for the enhancement of traditional learning with digital applications. Serious Games are not yet a fully understood construct. To allow for a personalized learning experience, which is tailored towards an individual's needs in every aspect, new powerful models have to be developed. Before Serious Games can find their way into learning in informal and formal teaching solutions, that make them applicable to a broad group of target audience, are needed. When done right Serious Games cannot only give new viewpoints and enhance the learning process in formal teaching, but can also teach where no teaching is available due to time, location or money constraints. The possibilities for Serious Games to enhance informal teaching across the world in various disciplines are endless. But at the current point in time there is still yet a lot of research on Serious Games to be performed.

1.3. Goal

The overall goal of this thesis is the development and research of new models for Serious Games, which allow for personalization. This is divided into five incremental sub goals. The first goal of this thesis is to conceptualize a theoretical balance model for Serious Games, which focuses on the separation between the learning and the gaming domain: *Componentwise Serious Game Balance (CSGB)*. The second goal is the conceptualization of a dynamic difficulty adjustment, which makes use of the newly developed CSGB: *Componentwise Dynamic Difficulty Adjustment (CDDA)*. The third goal is the implementation of the CDDA in the already existing Serious Game *HieroQuest*. The fourth goal is the development of a new dynamic mode, which allows the players to train the acquired knowledge. The mode should be based on CDDA and furthermore make long-term investigations possible. The fifth goal is the conduction of a user study to evaluate the affects of the CDDA on the short-term and long-term learning outcome and the player experience.

1.4. Outline

The rest of the thesis is structured as follows. Chapter 2 focuses on Serious Games, the theory behind learning, motivation and flow. First a definition (2.1.1) and a classification (2.1.2) of Serious Games are presented in Section 2.1. After that different learning theories (2.2.1), learning environments (2.2.2) and the concept of learning difficulty (2.2.3) are presented in Section 2.2. Followed by the theory of motivation (2.3.1), the history of flow (2.3.2) and flow in digital games (2.3.3) in Section 2.3.

Chapter 3 describes the theory and concept behind the *Componentwise Serious Game Balance (CSGB)* model. First the necessity of the model is discussed in Section 3.1. Followed by the theory behind the model in Section 3.2. After that the model itself is described in detail in Section 3.3. To conclude the chapter an enhancement of the model with time as an additional component is explained in Section 3.4.

Chapter 4 focuses on the balancing of digital games. First simple static approaches to account for different types of players are presented in Section 4.1. After that more advanced dynamic approaches for entertainment games (4.2.1) and Serious Games (4.2.2) are explained in Section 4.2.

Chapter 5 focuses on *Componentwise Dynamic Difficulty Adjustment (CDDA)*, the enhancement of the dynamic difficulty adjustment based on the CSGB. First the kind of Serious Games that allow for CDDA are classified in Section 5.1. After that the five necessary steps to implement the CDDA in an already existing Serious Game are explained in detail in Section 5.2: 1) Dynamically adjustable elements (5.2.1); 2) Measurements (5.2.2); 3) Difficulty levels (5.2.3); 4) Intervals (5.2.4); 5) Update function (5.2.5). Followed by a description of personalized gaming (5.3.1) and learning content (5.3.2) in Section 5.3.

Chapter 6 explains the different refinement circles the Serious Game *HieroQuest*, which is dedicated towards teaching the Middle Egyptian language, went through. First an introduction to the Middle Egyptian language (6.1.1), an introduction to the *Story of the Shipwrecked Sailor* (6.1.2), an ancient story used in the game, and an explanation why Hieroglyphs and Serious Games go hand in hand (6.1.3) are presented in Section 6.1. After that the gameplay (6.2.1) and limitations (6.2.2) of the first iteration of *HieroQuest* are explained in Section 6.2. Followed by the gameplay (6.3.1) and limitations (6.3.2) of the second iteration in Section 6.3. Finally the gameplay (6.4.1) and limitations and the results of the first user study (6.4.2) of the third iteration of *HieroQuest* are presented in Section 6.4.

Chapter 7 shows the implementation of CDDA in *HieroQuest*. First the implementation of the CDDA in *HieroQuest* is explained theoretically in Section 7.1. Afterwards the concrete implementation is shown on one of the rooms within *HieroQuest* in Section 7.2. Followed by an implementation of a new game mode, in which the level structure (7.3.1), the room structure (7.3.2) and the riddles (7.3.3) are dynamically created, in Section 7.3.

Chapter 8 focuses on the evaluation of the newly implemented CDDA. First the expected results of the user study are stated in several hypotheses in Section 8.1. After that the methodology of the first short-term user study is presented in Section 8.2. Followed by the methodology of the second long-term user study in Section 8.3.

Chapter 9 shows the limitations of the presented work and gives an outlook by showing some future directions. Finally Chapter 10 summarizes the presented work and gives a detailed conclusion.

2. Serious Games and Learning

The following chapter gives an overview on the concept of Serious Games (Section 2.1), including a definition (Section 2.1.1) and a classification (Section 2.1.2). Afterwards an important aspect of Serious Games, learning, is discussed (Section 2.2). In doing so several learning theories (Section 2.2.1) and environments along with several viewpoints on learning (Section 2.2.2) are presented, including a concept to rate the difficulty of a learning task (Section 2.2.3). After that the theory behind motivation and flow (Section 2.3) are presented, as they play an important role in the gaming and learning domain of a Serious Game.

2.1. Serious Games

Serious Games are digital learning tools with the possibility to enhance ones learning process with immersing gameplay to possibly create long-term motivation. They can also give new insides on a topic by introducing new technologies e.g. Augmented Reality (AR) [1]. Serious Games have to entertain and educate at the same time, which can be seen as two contrary domains. Combining both into one immersive experience, where both can profit from each other, is very difficult. Therefore they are often depreciated by the pedagogic community and the gaming community. But when done right the player/learner can profit from the dynamic possibilities a digital application can offer. Taking a look at both domains it becomes evident that the core group of target audience are players/learners, which are interested in both the gaming and the learning content. But to strengthen the position of Serious Games and to make them applicable to a broad group of audience, they should be effectively usable by players/learners without any experience to players/learners with a lot of experience in both domains. The dynamic possibilities a digital application offers have to be utilized to tailor a Serious Game towards an individuals needs. This section focuses on a definition and characterization of Serious Games (Section 2.1.1) and the classification of Serious Games into different categories (Section 2.1.2).

2.1.1. A Definition of Serious Games

Dörner et al. [2] define a Serious Game as "a digital game created with the intention to entertain and to achieve at least one additional goal". This definition is relatively generic by giving room to several topics that can function as the additional goal, but in general it states the combination of entertainment and some sort of serious content. The contrariness of the term itself was pointed out by Abt [3] more than 30 years ago. The term builds an oxymoron, trying to achieve the coexistence of serious content and fun. But also Dörner et al. [2] identified this "double mission" a Serious Game has to achieve by "being both effective and attractive". The term serious content is most of the time seen as educational content or learning content [4, 5, 6]. The term learning content will be used throughout this thesis as a generic term for every possible serious content they can have, which for example also includes physical exercises. Also the term gaming content will be used throughout this thesis as a generic term for everything besides the learning content, e.g. walking around in the game world. The state-of-the-art technologies for Serious Games are identical to the technologies available for standard entertainment games [7], which on the one hand opens up a lot of opportunities for possible areas of application, but on the other hand also puts a lot of weight on the developers, by expecting a Serious Game quality, e.g. in terms of visualization, near the gaming industry standard. Since most Serious Games are produced with tight budgets they suffer from poor visualization and overall poor game design and can therefore not fulfill these expectations [4].

With entertainment content and learning content blending into one experience, the question of the general purpose of a Serious game arises [4]. For example, Zyda [8] argues that the entertainment component should be the main focus, since a Serious Game that is not fun would be useless, but one could also argue that without learning content it would be useless. Taking a step back towards the concept of Gamification, learning content that is enhanced with playful elements, but cannot be seen as a digital game, reveals a similar yet not so drastic insight. The purpose of Gamification is to make learning more entertaining, motivating and engaging, which is similar to the purpose of Serious Games [9]. Both domains are immanent components and are equally contributing to the overall goal: entertain and educate. To generate such an enjoyable experience a certain threshold of aesthetic, game design and technology have to be met [4]. But not less important is the smooth, meaningful and homogeneous embedding of mechanisms that allow for knowledge or skill acquisition [5][10].

An important property of digital games is that they are dynamic. To be applicable to a broad group of target audience and therefore be profitable, digital games have to account for the capabilities of an individual [11]. For Serious Games this is a possibility, but at the same time a requirement. The group of target audience of a Serious Game reaches from players with no gaming experience to players with a lot of gaming experience, but also from players with no experience in the learning domain to players with a lot of experience in the learning domain. To account for these very different types of players the game needs to be tailored to an individual, in terms of e.g. prior knowledge, personal preferences, gaming experience, learning style, to maximize the effectiveness in the learning domain and the attractiveness in the gaming domain [11, 2]. The key part is that the game has to adapt to the players, in the best case automatically, by using e.g. predictors [12], rather than the players adapting to the game [13].

The potential areas of application reach from e.g. medical training of surgeons and high-schools to leisure free time activities [14, 15]. Here a distinction between formal and informal scenarios can be made [15]. In formal scenarios two possible use cases exist. Either they are used as a replacement for teachers or lecturers or they are used as a supplement. In the latter case teachers or lecturers not only have to be experts in the field of the learning content, but also at the game, which generates a completely new teaching scenario [15, 16].

2.1.2. A Classification of Serious Games

Serious Games can be classified according to various properties. For example, it would be possible to classify them according to their gaming genre, e.g. shooter or puzzle-game, which can be useful for players that are in general interested in a specific genre and want to try out new games. But a more suitable approach seems to be a classification according to the learning content. This can be done on a lower level, e.g. language or cultural heritage, or on an abstract level by taking a look at the different learning outcomes. Bloom [17] established a taxonomy of three possible learning outcomes for learning in general: motivational learning outcome (affective domain), knowledge learning outcome, (cognitive domain), manual/physical learning outcome (psychomotor domain). Garris et al. [18] translated this three domain taxonomy, which is agreed on by many scientists, to Serious Games: (1) skill-based outcome, (2) cognitive outcome and (3) affective outcome.

Skill-based learning outcomes (1) describe the development of physical or motor skills. Cognitive learning outcomes (2) are separated into three sub domains: Declarative knowledge, which accounts for facts and data used to perform a task; Procedural knowledge, which accounts for the procedure of performing a task; Strategic knowledge, which accounts for applying concepts to another context. Affective learning outcomes (3) describe the development of reactions and feelings, e.g. confidence [18].

2.2. Learning Theories and Environments

In this section the theory behind learning and how learning can be understood from different points of views are presented. Staring with learning theories (Section 2.2.1) and learning environments (Section 2.2.2) and the students motivation in these, towards defining the difficulty of a learning task (Section 2.2.3).

2.2.1. Learning Theories

There exist many different theories on how people learn and how people learn best. In general these theories focus on different aspects of the learning process. Many researchers agree on three sophisticated points of view on learning, each with several sub theories [19]: Behaviorism, Cognitivism and Constructivism.

Behavioral Learning Theories: In the behavioral point of view the learning process consists of the association between a stimulus and the corresponding response. In this theory feelings and thoughts are not part of the learning process. A teacher following the behavioral learning theory introduces students step by step to a process. The students are explicitly told what to do. This theory is also known as learning though observation. The teacher will certainly become a role model, which the student imitates. [19]

Cognitive Learning Theories: In the cognitive point of view mental processing is the main focus. The theory emphasized on the encoding in sensory memory, since perception determines which information is stored. The information in the short-term memory is linked to knowledge in the long-term memory and can be reactivated with the right stimulus. To achieve this the theory focuses on the "importance of environmental information perception". A teacher following the cognitive learning theory presents information through different communication channels. Auditory, visual and kinesthetic explanations are used to convey the information. Students are meant to work with the information to later on ensure long-term memory retrieval. [19]

Constructivist Learning Theories: In the constructive point of view new knowledge is purely constructed out of already existing knowledge. This not only includes own knowledge, but also interactions with others. In the theory information is not an exact copy of reality, but is rather a constructs of an individual, dependent on what the person already knows. A teacher following this theory focuses on the previous experience, culture, personality and background of the students to convey information. [19]

The behavioral learning theory is the only of the three that places the student in a passive role. Here the teacher is the only source of knowledge and teaches in a strictly organized way. In the other two theories the student plays an active role in the process and works with the information[19]. Another contrary learning theory, situated learning, specifies the process of learning as the process of becoming a sophisticated member of society [20]: Legitimate peripheral participation.

Situated Learning: Legitimate peripheral participation: In the situated learning point of view knowledge is gained as a process of becoming a full member of a group. People inevitably participate in activities within a community. The intention to gain knowledge and master skills is anchored in the pursuit of becoming an integral part of society. "Legitimate peripheral participation is proposed as a descriptor of engagement in social practice that entails learning as an integral constituent" [20].

Learning Styles: It was investigated that different people learn with different styles. Kolb and Kolb [21] identified four different learning styles: *Diverging*, *Assimilating*, *Convering* and *Accomodating*. The learner with a *divering* style learns best when analyzing a certain aspect from different viewpoints, which results in them performing better in situations that allow for brainstorming. The *assimilating* learner focuses on abstract concepts and is best at understanding information condensed into a concise form. The *convering* learner is best at finding practical use cases for theories. The *accomdating* learners are best at learning from previous experiences and enjoy involving themselves into new challenges. [21]

2.2.2. Learning Environments

Knowing about different learning theories and different learning styles enables the creation of effective learning environments. Bransford et al. [22] define these environments by the degree they are learner-, knowledge-, assessment- or community-centered. Learning theories do not give a basic recipe of creating such learning environments, similar to physic laws not stating a basic recipe to build a bridge [22]. These environments are micro systems with interpersonal relations, activities and roles with material and physical characteristics [23].

Learner-center Learning Environments pay close attention to the learner itself, by taking the knowledge, skills and beliefs of a learner into account. The state of a learner is, for example, acquired trough observation or conversation. A key factor of this learning environment is to develop the learner's knowledge structures by letting them make predictions about situations and afterwards reason about them. By using critical tasks with known misconceptions it is possible to identify flaws in the students knowledge and to help them develop their thinking. [22]

Knowledge-center Learning Environments pay close attention to teaching knowledge in ways that students can transfer this knowledge to other problems. The environment focuses on activities and knowledge that students can use to generalize about various disciplines. [22]

Assessment-center Learning Environments are an addition to learner- and knowledge-centered environments. They enable opportunities for revision, feedback and discussion. The assessment focuses on the learning goals to enable a deeper understanding and this way improve teaching and learning. This feedback and revision process can be used at intermediate steps to improve the learning and also at the end of the learning process to conclude an revise the complete process. [22]

Community-center Learning Environments are a relatively knew direction of research in terms of learning environments. It is stated that social norms, developed in an environment that allow to learn from each other and ensure a safe space to allow for mistakes, are important. It has to be made sure that these norms improve the learning environment and do not hinder the progress. For example a norm, which states that you should not make mistakes, is deconstructive and should instead state something like, "if we do not make mistakes we cannot learn from them". [22]

Learning environment and student motivation: By analyzing the relation between student motivation and learning environments in a classroom setting Radovan et al. [24] investigated a high correlation between goal oriented motivation, control beliefs and self-efficacy. Autonomy and teacher support along with the perception of usefulness seem to be important factors to make a learning environment and education enjoyable. The more a student perceived the learning task as valuable and relevant in terms of practical experience, the more motivation was perceived. The results further show that teachers could increase this motivation by encouraging autonomous work of students and relating real-life examples that are connected to the studied theory, which will overall result in satisfaction and enjoyment. To maximize students motivation a bottom-up teaching approach, that involves the needs and interests of students and gives them the possibility to actively participate in the process, seems to be appropriate. [24]

2.2.3. Learning Task Difficulty

In the *Cognitive Load Theory* Sweller et al. [25] describe the difficulty of a given learning task. The difficulty of a certain learning task can vary significantly. To further specify the difficulty of the learning task the *level of element interactivity* is conceptualized as the degree to which an element can be "learned without having to learn the relations between any other element" [25]. For example a piece of English vocabulary can be learned without the knowledge of other English vocabulary and therefore has a low level of element interactivity. Additionally, *intrinsic cognitive load* of a learning task is defined as the level of element interactivity of all elements within a task. If the interaction between a lot of elements is high than the *intrinsic cognitive load* is high and low otherwise. *Extraneous cognitive load* is defined as the cognitive load caused by the instructional control. Therefore, per definition, the overall difficulty of a task is defined by the intrinsic cognitive load, which is dependent on the learning elements, and the extraneous cognitive load, which is solely dependent on the given instructions. The overall difficulty can therefore be adjusted by varying the presentation of the information. [25]

2.3. Motivation and Flow

In this section the theory behind motivation and the related concept of flow are presented. First the two kinds of motivation and goals related to them, are presented (Section 2.3.1). Followed by the history of the concept of flow (Section 2.3.2) and flow in digital (Section 2.3.3).

2.3.1. Motivation

Taking a look at a person, whose motivation is self-authored and authentic, and a person that is externally controlled to perform an action reveals a difference in excitement, confidence, persistence, creativity and performance. These two persons are driven by two different kinds of motivation: Intrinsic and extrinsic motivation. The difference in all the previously mentioned observations can be found even when both have the same level of competence [26]. Therefore motivation is a construct of immense importance, since it produces [26]. Motivation is responsible for a persons well-being by making sense of their decisions and behavior and is based on three basic psychological needs of a human: autonomy, competence and relatedness [26].

Intrinsic Motivation describes the motivation perceived from an individual to perform activities that bare an intrinsic interest to them. It is the strongest form of motivation and cannot be externally induced. Activities which appear challenging, new and aesthetic can be the target of intrinsic motivation, while those who are neither of the three cannot. Intrinsic motivation is the satisfaction in performing the action itself without tracing a certain outcome. [26]

<table>
<tr><th>Structure</th><th>Instructions</th><th>Motivation</th></tr>
<tr><td>Task</td><td>Meaningful learning
Novelty, variety and diversity
Reasonable challenge
Support Short-term, self-reflected goals
Support development and effective learning strategies</td><td rowspan="3">Focus on effort and learning
Effort-based strategies
Positive affect on high effort tasks
Active engagement
Feeling of belongingness
Tolerance to failure
Effective learning and self-reflection</td></tr>
<tr><td>Authority</td><td>Students involved in decision making
Real choices based on effort
Support development of responsibility and independence
Support self-management and monitoring skills</td></tr>
<tr><td>Evaluation</td><td>Individual improvement and mastery
Private evaluation
Recognize effort
Provide opportunities for improvement
Mistakes as a part of learning</td></tr>
</table>

Figure 2.1.: Mastery goal supporting learning environment. Adapted from Ames [27].

Extrinsic Motivation describes the motivation of performing actions to achieve a certain outcome. Most of what people do is not intrinsically motivated. Students, who do their homework, because they think it is useful for their career, are extrinsically motivated. Students, who do their homework, because their parents told them to are also extrinsically motivated. Since in both cases the students do not enjoy the action itself neither of them is intrinsically motivated. Additionally, two different kinds of extrinsic motivation can be investigated here, the first type encapsulates a feeling of choice, personal endorsement and in general more autonomy, while the second one only involves external regulation. Autonomous extrinsic motivation in general shows better performance, more engagement and a higher quality of teaching and learning compared to the more controlling extrinsic motivation. [26]

Closely related to motivation are goals set from oneself or from an external structure. There are two types of goals: Performance goals and mastery goals. Both types can have different effects on ones motivation and bare different concepts of success:

Performance Goals are closely related to the performance of an individual compared to others. The purpose of a performance goal is to be better than others. It is focused on ones self-worth and abilities, confirmed by outperforming others or having success with only little effort [27]. Public recognition is a key factor of performance oriented goals [27]. In terms of a performance goal, learning is only an instrument used to achieve a certain goal, where the attention lies on the success [27]. Performance goals have an "undermining effect on intrinsic motivation" and should therefore be avoided in terms of teaching [28].

Mastery Goals are closely related to an individual mastering new skills. A mastery goal relates effort and outcome. An individual tracing a mastery goal is trying to improve their competence, understand the pursuit work and is in general self-reflected [27]. In terms of a mastery goal, learning is seen as a process in which effort will lead to success and in which a change of strategy can correct failure [27]. Mastery goals lead to a higher engagement in learning and effective problem-solving strategies [27]. They are especially related to interaction within a learning environment and therefore promote interactions related to peer modeling and intrinsic motivation [29]. In teaching interactions between students and teachers are critical for motivation and mastery related behavior [29].

When creating learning environments mastery goals should be supported to create a higher engagement and motivation. Figure 2.1 shows the structure and the strategies of such a learning environment [27].

2.3.2. History of Flow

The concept of flow is closely related to the concept of motivation, intrinsic motivation in particular. The sophisticated and broadly accepted "Flow State Model" was developed by Csikszentmihalyi [30] and tries to capture what makes an activity enjoyable. Enjoyment is found in very similar ways across the world in various activities regardless of age, gender or social status [31]. It consists of the following basic components [32]: (1) Tasks with a reasonable chance of completion, (2) Clear goals, (3) Immediate feedback, (4) Deep but effortless involvement, (5) Sense of control over our actions, (6) No concern for the self, (7) Alteration of the concept of time.

The concept of flow builds upon these criteria of an enjoyable experience. Additionally the balance between perceived challenge and personal level of skill is suggested to be the base condition for an optimal experience [30, 33, 34]. To experience flow the following conditions must hold [33]:

- Perceived challenges stretch existing skills (neither over-matching nor under-utilizing)
- One is engaging challenges at a level appropriate to one's capacities
- Clear proximal goals and immediate feedback about the progress

A visual representation of the "Flow State Model" can be seen in Figure 2.2 (left). The state of *Anxiety* and the state of *Boredom* are both extreme states in which either challenges exceed the skills or the skills exceed the challenges. The state of *Flow* is reached as long as an individual's perceived challenges and existing skills are in balance, then the individual remains inside the flow channel and experiences flow. In later studies it became evident that, a fourth state exists, the state of *Apathy*. This state is reached when the perceived challenge and level of skill are below an individuals threshold, the average level of activity. Even though challenge and skill are in balance a state of *Flow* is not reached when both are below this threshold, instead apathy is experienced (Figure 2.2) [35, 36].

2.3.3. Flow Models for Digital Games

The original flow model by Csikszentmihalyi [32] was conceptualized for everyday life, but was later on also used to describe states of optimal experience in play and sports. The model

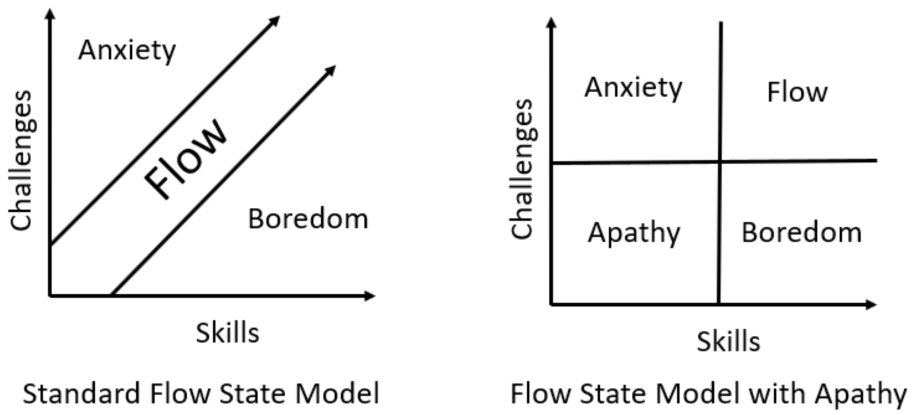


Figure 2.2.: Left Original Flow Channel; Right Flow Model with Apathy. Adapted from Csikszentmihalyi [33].

was used as a base for the "GameFlow" model by Sweetser and Wyeth [31]. Their model consists of eight elements in relation to the major components of enjoyment: concentration, challenge, skills, control, clear goals, feedback, immersion, and social interaction [31, 37, 38]. The criteria of the model allows for reasoning on the success and failure of various kinds of digital games. For entertainment games many of these models exist.

Flow model which pay special respect to Serious Games or which were especially designed for them are few. One recent approach by Nagalingam et al. [39] is the "user experience framework for educational games" (EUDGX), which tries to evaluate an educational game based on six elements: flow, immersion, player context, game usability, game system and learnability. The model includes flow as one of the elements rather than completely building on the original "Flow State Model". Another approach by Hoblitz [40], the educational game motivational model (EduGaM), separates the flow and learning flow to investigate the impact of the learning flow on the learning success. Kiili et al. [41] developed a similar model, which tries to evaluate the flow experience of an educational game by distinguishing between learning and gameplay task. In contrast, Pavlas [42] sees learning as the consequence of an optimal experience, rather than a contribution of the learning to this experience.

Sinclair et al. [43] developed the "Dual Flow" models for exergames. Those are Serious Games, which have a physical activity as the learning domain (Section 2.1.2). The model separates between "attractiveness", based on the original flow model, and "effectiveness", related to the effectiveness of the physical task. In the model both "attractiveness" and "effectiveness" are based on the basic principle of flow, the balance between perceived challenge and skill [43, 30]. The model therefore consists of two independent flow channels, capturing the balance between challenges and skills in the gaming domain and the balance between challenges (intensity) and skills (physical capacity) in the physical domain separately.

All of these models and frameworks try to evaluate the experience a players is going to emphasize based on the design and the mechanics of the game. Another type of approaches tries to evaluate the experience of an individual and therefore build the evaluation based on the players feelings rather than what they are supposed to emphasize. "Several self-report tools have been fashioned [...] to study this inherently unstable, unself-conscious, subjective phenomenon" [33]. The sheer number of different questionnaires [44, 45, 46, 47, 48, 49, 50] shows the general disagreement of the game research community on this topic and the instability of the concept, which is only further encouraged by known flaws in widely used ones, e.g. the Game Experience Questionnaire (GEQ) [51].

3. Componentwise Serious Game Balance

In this chapter the *Componentwise Serious Game Balance* model (CSGB) is explained in detail. The model conceptualizes the two domains of a Serious Game and the balance between the challenges and skills. It is not intended to be used as a criteria or measurements for good Serious Game design, but should rather be used as a basis for reflection on balancing in Serious Games. The model is based on the basic principle of an enjoyable experience: the balance between challenge and skill. First the necessity and the motivation behind the model is described (Section 3.1). Followed by its theoretical background (Section 3.2). After that the concept of the model itself is presented (Section 3.3). Then its dynamic aspect is discussed by including time as an additional parameter (Section 3.4).

3.1. Necessity of the Model

Models or Frameworks that try to evaluate the player experience in Serious Games are few. Most of them are based on a multi-dimensional approach with several categories, with flow being one of them [41, 39, 42]. This is in general in line with approaches for entertainment games [31]. However since flow is already a multi-dimensional construct [30] using it as a component of yet another multi-dimensional approach makes these models very complicated. While existing models see importance in the perceived challenges and skills of the players they do not distinguish between challenges of the learning domain and challenges of the gaming domain. For example, while in the "EduGaM" model [40] the distinction between learning and gaming challenges is only made in form of different questionnaires, but not in relation to actual elements within the game, Sinclair [43] only sees the distinction in the "Dual Flow" model necessary for exergames. The CSGB model includes the distinction between learning and gaming domain while introducing a new perspective on player experience in Serious Games by taking a step back to the most basic principle required for enjoyment, which is the balance between challenge and skill. Therefore the model is based on the original "Flow State Model" (Figure 2.2 left) [33].

3.2. Theory behind the Model

The goal of a Serious Game is to entertain and educate at the same time with equal contribution (Section 2.1.1). While in the learning domain the learning outcome should be maximized, in the gaming domain the fun should be maximized. Taking a look at these two contrary goals already hints towards a separate consideration of the both. This is further supported by the different challenges of both domains and also the different skills required to solve them.

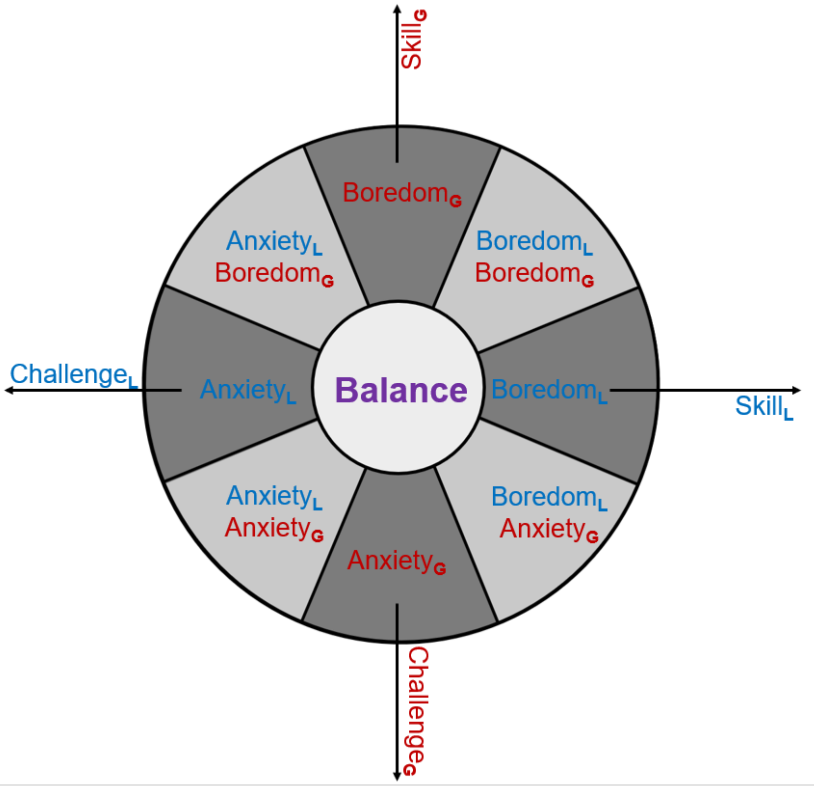


Figure 3.1.: Componentwise Serious Game Balance Model

The tasks the player has to perform in the gaming domain can differ immensely to those in the learning domain. Therefore domain specific knowledge is required to solve the tasks and at the same time domain specific knowledge is gained.

It is investigated that flow in general has a positive impact on the learning outcome of a Serious Game [52, 53], which is inline with the concept of the CSGB model. Even though both domains have to be treated individually they both contribute to the overall experience. In an ideal Serious Game, were learning and gaming perfectly blend into each other, the player will not be able to distinguish between learning and gaming domain and will only experience the sum of both. Therefore the balance in one of the domains does not imply an overall balance, but is rather a requirement for it.

Identically to the "Flow State Model" (Figure 2.2 left) [33], the CSGB model does not provide any values or measurements, since it only builds the theoretical base to later on allow for further implications. Compared to other models, that try to capture the overall experience, the CSGB model is on the one hand simpler, by only taking the balance between challenge and skill into account, but on the other hand also more complex by introducing the separation of the two domains. To act as a starting point for a new perspective on Serious Game design it is required to simplify in one direction to allow for a more complex structure in the other. By accounting for the perceived challenges and skills of an individual the model cannot be seen as a recipe for good Serious Game balance. Per definition it requires the verification

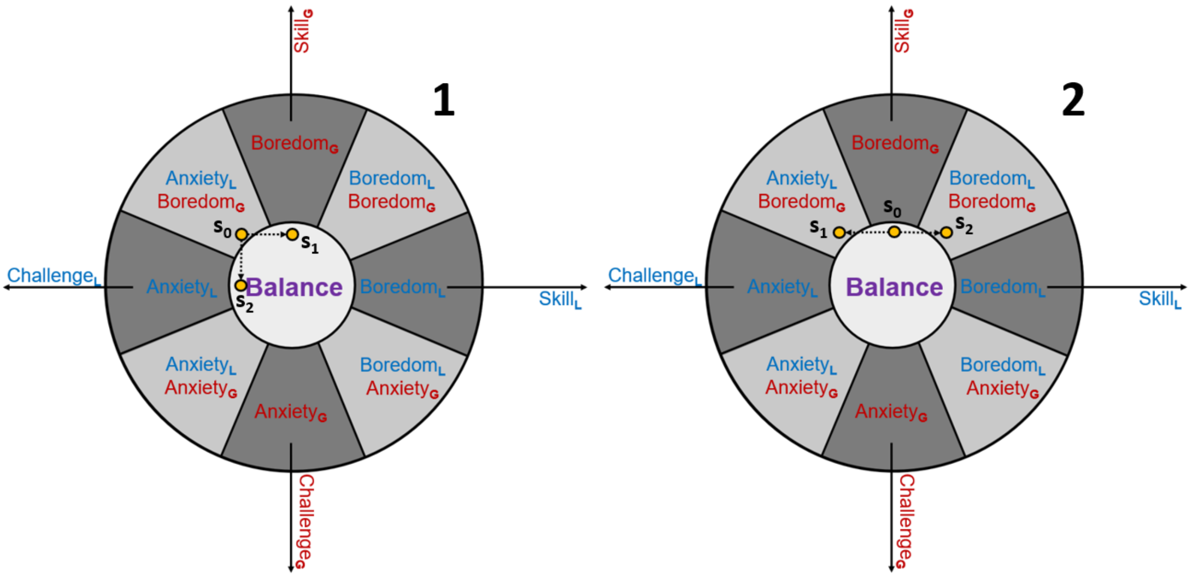


Figure 3.2.: CSGB Model with Different States: 1) From Imbalance to Balance 2) From Balance to Imbalance

of balance in both domains at any point in time for every individual player, which requires dynamic adjustment.

3.3. Componentwise Serious Game Balance Model

A graphical representation of the CSGB model is shown in figure 3.1: The horizontal axis accounts for the perceived challenges and skills of the learning domain. Those are all elements of the game that are directly related to the serious topic. These tasks require some form of knowledge in the serious topic, e.g. repetition or transfer exercises, writing and quizzes. The vertical axis accounts for the perceived challenges and skills of the gaming domain. In general everything that is unrelated to the serious topic e.g. controls, movement, logical riddles, way finding and time critical operations.

If the perceived challenges and skills of both domains are in balance the player will experience an overall state of balance (Figure 3.1 inner circle). This state is only reached when both challenges and skills are above an individuals threshold, otherwise a state of apathy would be reached [33]. The equal contribution of both domains, causes the circular shape of the model (Section 3.2). The overall structure relates to the four state flow model by Csikszentmihalyi [33] (Figure 2.2 right), in which the distance to the base point (inner circle) accounts for the intensity of the experience.

In the case of the CSGB model, the intensity of the balance state is the highest in the center point. When moving away from the center, the intensity of the balance state decreases and the intensity of the outer states increase. When outside the inner circle only one of the outer states is experienced. When the perceived challenges of a domain exceed the player's skills in this domain a state of anxiety is reached. When the player's skills in a domain exceeded the perceived challenges of that domain a state of boredom is reached. The state **s0** in Figure 3.2 annot. 1 shows a possible configuration, where either a relieve of challenge in the learning domain (**s1**) or an increase in challenge in the gaming domain (**s2**) would result in the player reaching the state of balance. This is caused by the equal contribution of both domains (circle shape) (Section 3.2). The state **s0** in Figure 3.2 annot. 2 shows a configuration in which an increase (**s1**) or decrease (**s2**) of the perceived challenges in the learning domain will result in a state of learning anxiety and gaming boredom (**s1**) or learning boredom and gaming boredom (**s2**). Both configurations show that the two domains can counteract each other to a certain degree, so an overall state of balance is still reached.

3.4. The Model over Time

An important part of the model is the dynamic aspect. The representation in Figure 3.1 shows only a snapshot. The different states in Figure 3.2 already hint towards a dynamic aspect of the model. At all points in time the balance between perceived challenge and skills have to be evaluated, because they constantly change. The player constantly develops new skills in both domains, while the challenges within the game constantly change, as well. Figure 3.3 visualizes the time as an additional axis. The purple arrows indicate the history of fictive snapshots of the current player experience. In this example all of the points remain inside the balance state, which can of course change dependent on the current situation.

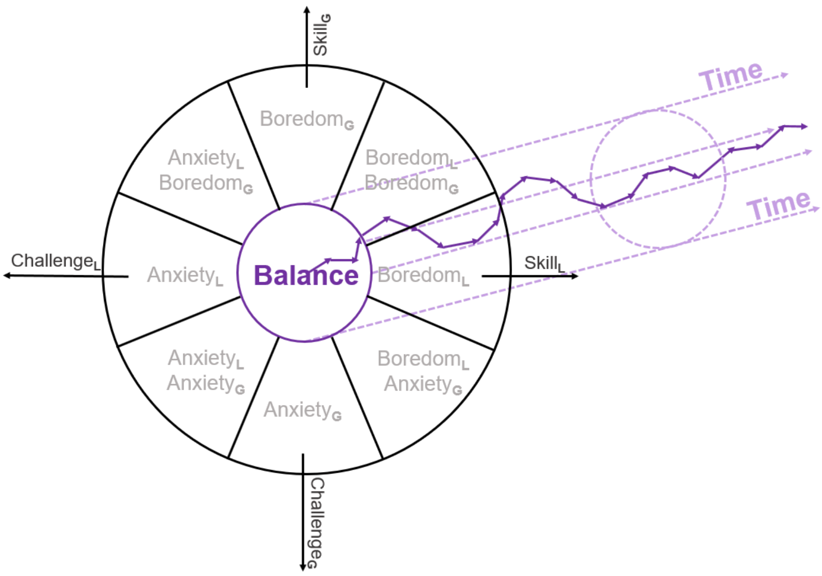


Figure 3.3.: CSGB Model over Time

4. Control Mechanisms for Player Experience in Games

Different approaches, which try to capture the player's experience while playing digital games, were discussed in Chapter 2. These approaches mostly use post-game questionnaires to achieve this. But this methodology only leads to an average value of the experience, but fails to identify in which parts of the game the players' feelings diverged from their average value, which makes refinement based on these values almost impossible. Therefore an average value can only be the basis of the analysis. When following Schell [54] one of the key components to ensure an optimal experience is continuous challenge. But there exists no one-fits-all solution for continuous challenge because challenge can be perceived very differently by different players, since it is depended on different factors. For the most part the previous experience with digital games influences a player's skill level and therefore the perceived challenge in a certain situation. But also other factors can influence the perceived challenge e.g. level of tiredness or current mood. This chapter focuses on the balance between challenge and skill, since it is the basic principle of generating an optimal experience (Section 2.3.2).

The straightforward approaches to achieve this balance are static approaches in which game design experts try to anticipate the player's needs, from a relieve of challenge to an increased challenge, in every situation to balance the game for the average player. Those design choices are than verified through multiple refinement circles that involve feedback from large amounts of testers [55]. In the digital game industry this approach is widely used because it is suitable for common digital games with a narrow group of target audience and high budgets. But for Serious Games the conditions shift. Serious Games are often produced with tight budgets and short amount of development time. Additionally, the learning domain induces a broad heterogeneous group of target audience from different backgrounds, since novice players that are for the most part interested in the learning aspect are now included. This makes this approach difficult to realize in practice [56]. While it is still possible and necessary to make use of these static approaches when first starting to develop a Serious Game, one will recognize that for users from different backgrounds and different previous experiences in the two domains a more dynamic approach is necessary. When including not only different starting experiences, but also different skill growing rates in the equation the limits of static approaches become evident.

In dynamic approaches the needs of an individual are not anticipated beforehand, but are rather tried to be captured during the game by the game itself. This way the game reacts to the players needs by changing its difficulty accordingly. This chapter will at first focus on static approaches (Section 4.1) followed by dynamic approaches (Section 4.2).

4.1. Static Approaches

This section focuses on static approaches that try to generate the best possible player experience, by balancing the difficulty. The most common approach to account for different skill levels is to give players the possibility to change the difficulty setting at the start of the game. A common approach are the three difficulty levels easy, medium and hard. Even though this approach gives the players the opportunity to decide between different difficulty levels, it is not ensured that these predefined difficulty levels are appropriate for every individual. When a player recognizes that the chosen difficulty setting is to hard or to easy there is in many cases no possibility to change it besides restarting the game and losing all of the progress. It therefore leaves the players responsible for their experience rather than the game. Even though in newer games there is the possibility to change the difficulty setting throughout the game, without loosing the progress, it does not change the underlying problem of leaving the player responsible. In addition, not only the previous experience of the players can differ, but also their ability to develop new skills can. To counteract the possible frustration of those, who do not develop the skills as fast as intended by the designers, additional guidance or the possibility to look up and this way relearn certain parts of the game have to be provided.

For example a hint-system or the possibility to click through a tutorial if the players want to do so can be implemented. If the game is large enough to allow for the players to make decisions of which path to choose inside the game, the difficulties of the different paths can also be used to give the players the possibility to account for their current skill level for themselves [57]. After all the designers have to keep in mind that they are dealing with skill levels increasing in different rates from different starting levels.

Since Serious Games introduce the learning as an additional domain, in which the players have different starting skill levels and different skill growing rates, this problematic gets even worse. Additionally, since the group of target audience of a Serious Game is also significantly larger than for an entertainment game, it is almost impossible to balance the game for all players with a static approach, without leaving the players with the decision. Therefore more powerful tools are needed in order to being able to tailor the game towards the individual player needs.

4.2. Dynamic Difficulty Adjustment

Since static approaches can only tailor a game to an individual's needs up to a certain degree, dynamic approaches were conceptualized. The most commonly used dynamic concept to achieve difficulty changes at runtime is Dynamic Difficulty Adjustment (DDA). While basic DDA approaches were already conceptualized at the beginning of the century [58], the field got more attention in the past decade with an ever-growing digital game market [59]. Basic approaches, but also DDA techniques using supervised-learning or clustering, exist. Regardless of their implementation they all try to tailor a game towards an individual's needs. The most basic DDA concept is the rubber-band technique. When the player pulls harder (is more skilled) then the game pulls harder too (gets more difficult).

This basic technique, for example, is used in all *Mario Kart*[1] games. In this multiplayer racing game players drive go-karts on a lapped based racing track. They can collect different random power-ups to give themselves an advantage or sabotage others. The rubber-band technique is used to determine the power of these "random" power-ups: Player that are currently in lower rankings get better power-ups than those in higher rankings to give them the possibility for a comeback. Since *Mario Kart* is a party game, the intention is to not exclude less skilled players from the fun. But this approach, especially in a multiplayer environment, can be very frustrating for the more skilled players, since they have to always fight an uphill battle against the game and are punished for their experience while others are rewarded for their inexperience. In competitive multiplayer scenarios it is very hard to implement a DDA that is not frustrating for more experienced players, therefore the concept of on the Multiplayer Dynamic Difficulty Adjustment (MDDA)[60] was created. Since this scenario induces very different requirements and goals and the focus of this thesis are single player Serious Games, in the rest of the chapter only DDA approaches will be presented.

In single player games the possible frustration of experienced players by seeing that inexperienced players are privileged is reduced simply due to the fact that no other players are present. But players can also get frustrated when they recognize that the difficulty of the game changes and they are now punished for being more experienced. Of course the standard game design choice, regardless of implementing DDA, is to increase the difficulty over the course of the game, since the players' skill levels also increase, but when jumps in the perceived difficulty are noticeable the players can get frustrated. Therefore the DDA should in the best case recognize an individual's needs and dynamically adjust the difficulty accordingly without the individual noticing. A consistent DDA implementation can be arbitrary hard to achieve and is always tied to the game genre and more specifically to the game at hand [61]. For example in games that use a discrete room structure were the players advance in a linear fashion a consistent DDA is easier to achieve than in an game that features an open-world [56].

Before being able to adjust the difficulty according to an individual's needs it is of course necessary to measure them. There exist various techniques that try to achieve that, which vary from game to game and game genre to game genre. A common approach is to use a player's in-game performance to measure them, but also measurements outside the game world can be utilized. When measuring the in-game performance the number of correct actions, the time it takes a player to complete a certain task, the accuracy of clicks and movement, the number of deaths, the amount of damage taken, etc. can be used. When measuring outside of the game world a common approach is to measure a player's heart rate, to determine the current stress level of a player [62, 43].

[1]https://www.nintendo.com/games/detail/mario-kart-8-deluxe-switch, last visited 13.10.2020

4.2.1. Dynamic Difficulty Adjustment in Games

In the following a few examples of DDA implementations in entertainment games are presented:

Liu et al. [63] investigated the usage of an individual's anxiety level to adjust the difficulty of a game. To do so they used "wearable biofeedback sensors" to measure "peripheral physiological signals". These values were then used to control a DDA. A comparison between a performance-based DDA revealed that the players showed lower anxiety during the anxiety-based DDA session and 7 out of 9 participants also showed an improvement in performance. The results suggest that the usage of measurements outside of the game world can possibly enhance the DDA.

Hunicke and Chapman [64] investigated the amount of supplies (Ammunition, Healing Supplies etc.) the strength and accuracy of enemy attacks and the strength of own attacks in a first person shooter, to adjust the difficulty. The DDA is controlled by measuring several in-game variables e.g. the health of a player, the number of times the player has died overall and also in the current level, the time spend in-game and the number of interventions by the DDA. The intention is to keep the arousal rate, measured by the heart rate, equal over time, "while dynamically adjusting to keep them alive longer" [64].

Zook and Riedl [65] created a DDA based on a data-driven player model that uses a temporal component to effectively forecast the players performance on future challenges by taking their skill growth overtime into account. The predicted individual performance in addition to an objective performance on a certain task then guides a DDA to adapt the parameters of the upcoming task.

Shaker et al. [66] implemented an dynamic level generation for *Super Mario Bros*[2] by using a "grammatical evolution-based level generator", which utilizes a fitness function to adjust the game towards the player's needs. The three emotional states engagement, frustration and challenge are used to calculate this function. The results showed that engaging levels were also challenging, while not all challenging levels were also engaging [66]. An interesting aspect of this specific DDA implementation is that it dynamically changes the structure of the overall level, since in a 2D Platformer like *Super Mario Bros* the difficulty of the task is determined by maneuvering through the level.

Denisova [67] measured the in-game performance, enemies killed, of players inside a top down shooting game to control an adaptive timer displayed at the bottom of the screen. When an individual was performing like an average player the timer was ticking down with 1 unit per second, and was sped up or slowed down by a factor of 1.4 without informing the players, when performing better or worse than the average player. In a user study the implemented

[2]https://www.nintendo.com/games/detail/super-mario-bros-35-switch, last visited 13.10.2020

DDA was tested against a control version of the game without the timer manipulation. The results showed that the players of the DDA version felt more immersed than the players of the control version. The immersion score was measured according to the Immersive Experience Questionnaire (IEQ) [48].

4.2.2. Dynamic Difficulty Adjustment in Serious Games

While for entertainment games the usage of DDA is a common approach and many implementations exist only very few Serious Games make us of this concept, even though a more diverse group of target audience urges the need (section 4.1) [56]. It might also be the case, that DDA is not often used in Serious Games because the games are often subject of research and are in most cases relatively short. Nevertheless, it is understood by different researchers that it is a necessary step for Serious Games to generate an optimal experience. For example, Hamari et al. [68] stated that in game-based learning "The game should be able to keep up with the learners", but already Tremblay et at. [56] identified the need for Serious Games "to adapt the difficulty level [. . .] to different sets of skills". For Serious Games identifying an individual's needs is also the first step. For example, in a Serious Game for emergency personnel training Ninaus et al. [62] measure the players' heart rates to determine their stress level. If a player's heart rate increases and reaches an individual threshold additional guidance is provided, to generate a relieve of pressure [62]. But of course also in-game measurements are used. For example, Tremblay et al. [56] used the time it takes the players to solve a task as a measurement in the Serious Game *Number to Number Combat* (Section 5.1.1). In a round based environment the solving speed of the non-player character is changed according to the player's speed in the previous round.

An important difference between Serious Games and entertainment games is the inclusion of an additional domain, e.g. learning. The previous examples do not treat the difficulty of the gaming domain and the learning domain independently. In both examples only the difficulty of the learning domain is adapted even though it is not clear if the inexperience in the learning domain caused the measurement. Sinclair [43] identified this problematic and uses two different types of measurement to change the difficulty of the learning and gaming domain independently. In his exergame (digital game + physical task, Section 2.1.2) the players have to ride an exercise bike to control the height of an in-game character flying in a 2D Side Scroller to avoid enemies and collect coins. While the needs in the gaming domain are measured in-game, amount of avoided enemies and collected coins, the needs in the learning domain are measured outside of the game world, heart rate. Sinclair [43] only identified the need to adjust both difficulties independently for exergames, but the reasoning in Chapter 3 showed the need for this separation for the overall field of Serious Games. The enhancement of DDA towards a concept for Serious Games which accounts for the independence of the two domains is presented in Chapter 5.

5. Componentwise Dynamic Difficulty Adjustment

The CSGB (Chapter 3) conceptualized the existence of the learning and the gaming domain in Serious Game. An individual has independent levels of perceived challenges and independent skill growth rates in the two domains. Both of them contribute to the overall perceived challenges of an individual, while they can also counteract each other at least to a certain degree. For example, it is possible that a slight underload in the gaming domain can be counteracted by a slight overload in the learning domain, which leads to an optimal experience. If the under- or overload in one or both of them is to high, this is no longer the case and the individual feels boredom or anxiety in one or both of them and overall experiences discomfort. Since it is intended to keep the player inside the optimal state at all times, DDA can be used for both domains independently to control the perceived challenges and ultimately generate the best possible experience for an individual. To fulfill this gap an enhancement to the already existing model of DDA was developed, the Componentwise Dynamic Difficulty Adjustment (CDDA). The rest of the chapter focuses on the properties of the CDDA, which Serious Games can profit from CDDA (Section 5.1) and how to implement CDDA in already existing Serious Games or how to design appropriate elements beforehand (Section 5.2).

5.1. Target Games

CDDA can be implemented in different kinds of Serious Game. Yet it shows full potential in Serious Games that focus on declarative knowledge transfer (Section 2.1. Additionally, the game has to include a minimum of gameplay. For example, CDDA cannot be fully implemented in a digital vocabulary trainer, since the difficulty of the gaming domain cannot get dynamically adjusted if there are simply no elements in the gaming domain. The opposite holds for entertainment games in which the difficulty of the learning domain cannot get dynamically adjusted because it does simply not exist. To show that CDDA is a powerful concept and is applicable for various kinds of Serious Games, two games from different genres are used to enhance the explanations of the different steps necessary to implement the CDDA. In the following of this section these games are explained briefly to later on enhance the explanations in Section 5.2.

5.1.1. Number to Number Combat

Number to Number Combat [56] is a Serious Game dedicated towards teaching elementary mathematical equations. The game is a round-based 2D game with a small amount of core gameplay that only allows the players to insert the solution of an equation via a numerical pad. When the player correctly solves the equation an attack animation will be shown and the computer controlled enemy loses a certain amount of health. The enemy also solves equations at a certain speed and hits the player once an equation was solved. If one of the health bars reaches zero the next round will start with a dynamically adjusted solving speed of the enemy. Therefore, in its standard form, the game already implements DDA bound to the player's abilities to solve the equations. In its standard form the equations consist of two numbers and the operations Addition and Subtraction and the solutions of the equations are always single digit numbers, which makes the equations fast to solve for players with elementary school graduation. Therefore basing the dynamic adjustment solely on the solving speed makes it not only dependent on the players skills in the learning domain, but also in the gaming domain because a large amount of the solving time of one equation is dedicated towards finding the number on the numerical pad and clicking on it.

5.1.2. Oppidum

Oppidum [69] is a Serious Game dedicated towards teaching trivia about the Celts. In this Augmented Reality (AR) based board game the player has to build an oppidum, a Celtic fortified town, by placing markers, which represent different buildings, on the board. The spots where markers can be placed are fixed, but the buildings can be arranged to the players liking, which is necessary since buildings next to each other affect their production effectiveness. The player needs these products to build new building and upgrade them. While exploring the buildings with the help of AR, inside and outside, the players learn about the Celts and their culture. The game can be played by two players, which both build a town for themselves and battle each other in trivia quizzes, or by a single player, who will then play against the Computer. Since the game already includes a sufficient amount of gameplay, which allows for the implementation of the CDDA no game-play changes are necessary. The building task is only dependent on the players' skills in the gaming domain and the trivia quizzes are only dependent on the players' skills in the learning domain.

5.2. Implementation Guidelines

In this section the necessary steps to implement the CDDA in an already existing Serious Game are explained in detail. The only difference between the presented steps and the necessary work when deciding to implement CDDA in an early stage of development is that in the early stage one is able to account for the properties of the CDDA beforehand and build the game around it. Therefore the more difficult of the two approaches, implementation in an already existing game, is explained in this section. Implementing the CDDA for an already existing game requires five steps:

- Determine the dynamically adjustable elements and their domain (Section 5.2.1)
- Determine measurements for both domains (Section 5.2.2)

- Assign difficulty levels for all elements for both domains (Section 5.2.3)
- Determine discrete intervals for the CDDA to update the difficulties (Section 5.2.4)
- Assign an update function to adjust the difficulties (Section 5.2.5)

5.2.1. Determine the elements and assign them

The first and most important step of the CDDA implementation is the determination of the respective elements. At first all elements of the game that can be dynamically adapted have to be identified, regardless of their domain. If the game already includes static difficulty settings, which allow the player to select the difficulty at the start, these settings can be utilized when looking for dynamically adjustable elements. If a static difficulty setting is effecting an element, chances are high that the element can also be changed dynamically. When searching for these elements it is important to start with the core gameplay mechanics of the Serious Game and make sure that they are included. If those elements are not identified as dynamically adjustable elements and are therefore not included in the CDDA the whole point of the CDDA gets lost. If the core gameplay elements of a game are not dynamically adjustable they have to be changed accordingly.

For example, in *Number to Number Combat* the core gameplay activity is to solve equations and insert the answers on a numerical pad. The equations only include the two compositions Addition and Subtraction, single digit solution and only contain one composition at a time, which does not allow for an increase or decrease in difficulty because these rules are to restrictive. By loosening up these rules, to allow for less and more difficulty equations, they become dynamically adjustable. By introducing the two operations Multiplication and Division and allowing for equations with multiple operations, while keeping the result at a single digit, the task is now mostly dependent on the players skill in the learning domain. After the equation is solved the player will automatically attack the enemy. Since the attack is automatic and does not require the player to perform any action it also does not require any skills. To make this core activity dynamically adjustable, it has to be changed. In order to increase or decrease the difficulty, the enemy will get the possibility to block the players' attacks, which makes them dynamically adjustable and now require skill.

In *Oppidum*, for example, the core gameplay mechanics are placing buildings on the board and fighting the enemy in a trivia battle. The buildings can only be placed at marked spots, but the player can decide the spot and the kind of the building. Since buildings next to each other affect their productivity this mechanic allows for a dynamic difficulty adjustment. The productivity affects and the costs of the buildings can be dynamically adjusted to force the player to seek for the perfect organization of the oppidum. In addition, the trivia battles can be dynamically adjusted by adjusting the difficulty of the questions and the solving skills of the enemy.

After all dynamically adjustable elements are identified they have to be assign to either the learning or the gaming domain. Elements that requires mostly skills in one of the domains should be assign to this domain. If an element requires skills in both of the domains it can be considered to split the element in smaller parts. Since elements are only affected by the difficulty setting of one domain, they should also only require skills in one. If the split of the element is not possible then the element needs to be redesigned, because the concept of the separation in two domains would get lost.

In *Number to Number Combat* the core gameplay activity to solve the equations solely requires the player's skills in the learning domain therefore this element is assigned to the learning domain. The attack mechanic solely requires the player's skills in the gaming domain, since pressing a button at the right time does not involve any knowledge in the learning domain. Therefore this element is assigned to the gaming domain.

In *Oppidum* the identified elements are the building placements and the trivia battles. Since the building placement does not require knowledge about the Celts and only require the player to logically think about the optimal placement this element is assigned to the gaming domain. The trivia battles require knowledge about the Celts, but do not require the player's skills in the gaming domain, since they are answered by simply clicking on the correct answer without any time restriction. Therefore the trivia battles are assigned to the learning domain.

The next step is to apply this procedure to all elements of the game that can be made dynamically adjustable. If the elements are not part of the core gameplay it can be debated whether these elements should be included in the CDDA, since it is in general only required to adjust the core elements. After all desired elements are adjustable and are assign to one of the domains a measurement for both domains has to be conceptualized (Section 5.2.2).

5.2.2. Determine measurements

The second step is to determine appropriate measurements for the two domains. These measurements are dependent on the elements of the domains, which were identified in the first step (Section 5.2.1). It is not required to solely measure inside the game world, e.g. it is also possible to measure the heart rate of a player to determine the level of arousal [62]. But when not directly measuring the performance on a certain element it has to be made sure that the measurement is solely affected by the this element. Sinclair [43] uses the heart rate to determine if the difficulty of the cycling task in his exergame is to high. Exergames are Serious Games with a physical task as the learning content. If the heart rate exceeds a certain individual threshold the necessary intensity will decrease. Similarly the necessary intensity will increase if the heart rate succeeds a certain threshold. But since the difficulty of the gaming task can also influence the heart rate it is not save to say that changes in heart rate require changes in the "learning" domain. Therefore, for the CDDA to function properly, it has to be made sure that the measurements for a domain are only affected by the perceived challenge of this domain. This is in the most simple case achieved when measuring the performance of the players on the elements of that domain by e.g. measuring the amount of fails, successes, the time it took to complete the element or the amount of actions needed to complete the element.

In *Number to Number Combat* the measurement for the learning domain should be solely dependent on the player's performance in the equation solving task. Therefore the time it takes to solve the element or the number of failed trials are in general candidates for the measurement. Since in the game one equation can only be attempted once, both success or failure will result in a new equation being presented to the player. Therefore the number of failed trials is not applicable here. An appropriate measurement will consist of the time from the equation being presented to the player to the player inserting a solution and the correctness of the solution. By only accounting for one of the two the players will either have infinite time to solve an equation or their answer does not matter at all. The measurement of the gaming domain should account for successful attacks compared to the amount of all attempts of attacks. If the player would not be able to solve a single equation there would be no possibility to attack, therefore solely relying on the amount of successful attacks would create a dependency on the difficulty of the learning domain which should be avoided. Therefore it is necessary to put the number of successful attacks into perspective to all attacks. Additionally, the players have a certain amount of time to perform an attack after solving an equation to give them the possibility to wait for the perfect moment. It can also be considered to measure the time it takes the players to perform the attack. But in our case the measurement will only rely on the percentage of successful attacks because by design the gameplay should focus on precise rather than quick movement, similar to mathematics. But in other cases the time it takes to perform the attack can be a valuable measurement.

In *Oppidum* the measurement for the learning domain should be solely dependent on the player's performance in the trivia quizzes. The number of correctly answered questions compared to all questions in the quiz and the time it takes for completion are appropriate measurements. In this special case only the percentage of correctly answered questions is used as an measurement, because there is no clear advantage gained by using more time to solve the questions. The measurement in the gaming domain should be dependent on the players' current production rates. When starting the game the players can only build very few buildings, but after earning enough materials through production, there are many buildings to choose from and their placement is crucial for their productivity. Because this placement process is not trivial the game also needs to calculate the currently best placement of productivity at each calculation step (Section 5.2.4) and cannot rely on predefined rates. The productivity rate of the player will be put into perspective to best possible productivity rate at the current point in time.

After measurements for both domains were defined the amount of levels of difficulty for both domains have to be determined. Followed by assigning these difficulty levels to each of the elements (Section 5.2.3).

5.2.3. Assign difficulty levels

The next step is to determine an appropriate amount of difficulty levels for both domains. In general, three difficulty levels for both are a good starting point, but when accounting for highly heterogeneous groups of target audience it can be necessary to use five difficulty levels. If static difficulty settings are already implemented for some or all of the elements, they can be utilized. Otherwise the currently implemented difficulty can act as the medium difficulty level. But it has to be made sure that the difficulty levels are consistent for the different elements within a domain. It is not appropriate to include elements that are significantly more difficult than other elements at the same difficulty level. For each of the elements all difficulty levels have to be set. Those difficulty levels should make a noticeable difference in the player's perceived difficulty. In general, appropriate things to adjust in the difficulty levels are the amount of elements, the amount of actions required to solve an element, the time it takes to complete an element and if certain elements are interchangeable, also the kind of element.

For *Number to Number Combat* three difficulty levels for each of the domains seem to be sufficient, in the following referred to as easy, medium and hard. Both domains only contain one adjustable element. The core element of the learning domain are the property of the equations. The difficulty levels of the equations should manipulate the amount of compositions and numbers. Also the choice of compositions are possible candidates, since Addition is easier to calculate than Division. In the easy difficulty the equation should only consist of a single composition and therefore only two numbers. It should only include Addition and Subtraction. The medium difficulty level should also include Multiplication and Division as additional composition kinds. The hard difficulty level should also include multiple compositions and therefore more than two numbers. The gaming difficulty levels also only have to be set for one gaming element. The combat difficulty is determined by the enemies ability to block the player's attack. The enemy will periodically block attacks from the player. By setting different frequencies it is possible to set the three difficulty levels. In the easy gaming difficulty the enemy will block every 2 second, in the medium difficulty every 1 second and in the hard difficulty every 0.5 seconds.

In *Oppidum* three difficulty levels for each of the domains also seems to be sufficient, in the following referred to as easy, medium and hard. Since both domains only contain one adjustable element, only the difficulty settings of those two have to be set. In the learning domain the only learning element are the trivia quizzes. The difficulty of such a quiz can be adjusted by the amount of questions the difficulty of a question and the number of possible answers. For consistency reasons the number of possible answers was set to four in all difficulty settings. The difficulty of the questions is possibly perceived differently by the individual players, therefore the most simple and reliant approach is to adjust the number of questions for each quiz. This is adjusted to 3 questions in the easy, 5 questions in the medium and 7 questions in the hard learning difficulty setting. In the gaming domain the only adjustable element is the building process. There are several possibilities to adjust the difficulty of this process. It is possible to adjust the punishment for incorrect placement, which means a dynamic decrease of productivity when not placed near the correct neighbour, but also increasing the costs for buildings and items can be dynamically changed to indirectly punish an incorrect placement. Adjusting the performance punishment seems to be a more direct and straightforward approach and will therefore be implemented. In the easy gaming difficulty setting the productivity punishment for incorrect placement will be 20%, in the medium gaming difficulty 50% and in the hard difficulty 80%.

After the difficulty levels for all the elements of both domains are determined the intervals in which the difficulties will stay consistent have to be defined (Section 5.2.4).

5.2.4. Determine discrete intervals

An important aspect for DDA and therefore also for CDDA is consistency. In the best case the players should not notice that the difficulties are adjusted. To accomplish this it is necessary to define discrete intervals in which the difficulties will not change. Adjusting the difficulties constantly could result in the difficulty jumping from one setting to another without giving the players the possibility to account for the new situation. These intervals can either be time or action dependent, which depends on the game at hand. In general, action dependent intervals seem to be simpler to implement, because it can be made sure that these actions result in the player entering a new area or section of the game or the view or scene changing. Time dependent actions do not result in such a scene or area change and thus can possibly result in the player noticing changes in the environment and are therefore not that simple to implement correctly.

Since *Number to Number Combat* is a round based fighting game, which ends with the player's or the enemy's health reaching zero, one interval should be equivalent to one round. The end of a round also induces a scene change and the restart of the fighting round. This further encourages the decision to adjust the difficulties after a round, because the players won't immediately recognize these changes and have to explore them in the next round. But adjustments within a round would cause confusion and adjustment in these short amounts of time would also leave only little time for the player to account for the new situation.

In *Oppidum* such rounds can also be identified. The player has the possibility to fight against the enemy in a trivia quiz once per in-game day after all desired building operations are finished. These quizzes also result in a scene change. Thus an interval should end once the trivia quiz is finished. This way it is made sure that in each interval both gaming and learning skills were used. The scene changes also further encourage this decision.

In general, the end of an interval should be placed at the end of logical blocks within the game. Adjusting the difficulties to often or not often enough will result in the difficulties jumping. By tracking the difficulties of an individual for all of the intervals it is possible to identify these flaws and refine the interval structure in a later iteration.

5.2.5. Assign a adjustment function

The last step in implementing the CDDA is to define adjustment functions for both of the domains. These functions are used to calculate the difficulties for the next interval based on the player's performance in the current interval. The most simple approach only accounts for the performance in the previous interval and omits the performance in older intervals. Implementing a formula like this can also result in the difficulties jumping. Taking more intervals into account will smooth the results but will also decreases the functions sensitivity to changes in the player's skill growing rates.

In general such a function should allow for a fast decrease in difficulty, especially at the start of the game to not lose inexperienced player in the first minutes. But an increase in difficulty should only occur ever so slightly to verify that the individual is consistently performing above average. The following simple adjustment function fulfills these demands:

$X \in \{\text{Gaming domain, Learning domain}\}$ d_X = current difficulty of domain X
p_X = performance in the current interval in domain X
l = lower Bound u = upper Bound

$$f_X(d_X, p_X, l, u) = \begin{cases} d_X + (p_X - l) & p_X < l \\ d_X & l \leq p_X \leq u \\ d_X + (p_X - u) & p_X > u \end{cases} \tag{5.1}$$

In this simple example the same function (Equation 5.1) is used to adjust both difficulties. If the performance $p_X \in [0, 1]$ of a player in a domain is between the lower Bound (l) and the upper Bound (u) ($p_X \in [l, u]$) the difficulty of that domain will not change. If the performance is less than the lower Bound ($p_X < l$) the difficulty of the domain will decrease by a maximum of l. If the performance of the player in that domain is greater than the upper Bound ($p_X > u$) the difficulty of the domain will increase by a maximum of $1 - u$.

For both *Number to Number Combat* and *Oppidum* a simple adjustment function seems suitable. Since in both Serious Games the gameplay does not change by a lot from interval to interval, the adjustment can be assumed to be relatively stable. But testing has to show if a simple implementation would result in the difficulties jumping. If this is the case more than the last state can be included into the function. In our case a function that decreases the difficulty if the player's performance is below 60% (lower Bound) of the maximum performance, does not change the difficulty between 60% and 80% and increases the difficulty if the player's performance is higher than 80% (upper Bound) is used. This accounts for a fast relief in perceived challenge in case of a too high starting difficulty and a slow increase of difficulty when constantly performing good.

With the assignment of an adjustment function the implementation of the CDDA, in its basic form, is completed. The CDDA can be further adjusted by the different variants specified in the corresponding steps (Sections 5.2.1, 5.2.2, 5.2.3, 5.2.4, 5.2.5). In addition to the basic CDDA, introducing personalized learning and gaming content can tailor the game even further to an individuals needs (Section 5.3).

5.3. Personalized Content

By introducing dynamic concepts into a Serious Game the game can be tailored towards an individuals needs. CDDA is a powerful tool to achieve an independent adjustment of the difficulties of both domains of a Serious Game to tailor them towards the players skill levels. This tailoring can be further increased by personalizing the content of both domains to an individual. The degree of possible personalization is of course always dependent on the game at hand. But also small degrees of personalization further improve the game towards an optimal learning and gaming experience. Both concepts, personalized gaming (Section 5.3.1) and personalized learning (Section 5.3.2) will be explained in detail in the following.

5.3.1. Personalized Gaming Content

Dependent on the gaming content the game can be personalized in a lot of different ways. The essence is to dynamically adjust the gaming content towards the players needs. This can either be achieved by measuring the players performance on certain elements of the game or use already existing measurements on complete intervals. In doing so, it is possible to dynamically change certain elements of the game similar to the changes within the difficulty brackets (e.g. easy, medium and hard), by replacing them with other elements. This approach is only sufficient if the game includes a minimal amount of interchangeable elements. For example, a logical riddle which is used to challenge the player in the gaming domain, but has no connection to the game world in terms of plausibility can be changed to another logical riddle without any consequences regarding the overall structure of the game. By comparing the player's performance on those "generic" elements to a base value it is possible to determine whether the choice of element is appropriate for the individual player. If a player takes significantly longer on a certain generic element type than the base value the element should be changed to another generic element to further support the player's engagement. Another approach to personalize the gaming content is to dynamically change the overall structure of the environment. This is only appropriate for games that include a certain movement task to allow for those changes. For example, it would not be possible to change the structure of the environment of *Number to Number Combat* or *Oppidum*. But Serious Games which allows for those structure changes can be further personalized towards the players needs by adjusting the "difficulty of the environment". This difficulty mainly refers to way-finding tasks the players have to perform, but on a smaller scale can also refer to maneuvering tasks which can be perceived differently by an individual even though they are considered equally difficult, e.g. in a 2D Platformer jumping two times and then dashing is for most players equally difficult than jumping three times or dashing first and then jumping two times, but a certain individual might struggle with one of these orders even though they are considered equally difficult on a design choice. By constantly having to perform actions that are perceived more difficult than intended an individual can lose motivation and tension can arise. These feelings can be counteracted by changing those tasks to tasks of the players liking. This will benefit the over feelings towards the game and therefore most likely also the player experience.

5.3.2. Personalized Learning Content

In the learning domain the Serious Game can also benefit from personalization. Tailoring the learning content towards a player's needs can not only increase an individuals motivation and improve the feelings towards the game, but can also increase the learning success. Of course this tailoring is dependent on the Serious Game and the learning content at hand. Thus for some Games the degree of personalization might be relatively low, but in general Serious Games, that query the acquired knowledge through repetition throughout the game, will most likely allow for personalized learning content. Changing the order in which certain elements are presented to players is not always possible. For example, in *Number to Number Combat* it is not intended to introduce players to Division before Addition. But after the players have already got in contact with the elements in the desired order the consolidation phase allows for adjustments. In this phase it is possible to query them at different frequencies. For example, in a Serious Game dedicated towards teaching a language, vocabulary, which queries the vocabulary throughout the game to verify the learning progress, can be personalized towards the player's needs. To not annoy players by constantly querying vocabulary which they already know perfectly, the choice of queried vocabulary should be personalized towards the weaknesses of an individual player. This also enhances the learning process by tackling flaws in the players knowledge. By measuring the performance on the repetition of certain learning elements it is possible to identify the players flaws and counteract them presenting these elements more frequently. A possible choice of spaced-repetition technique to choose the personalized learning elements at a given point in time is the *Leitner System* [70]. The system gives every element a score starting at zero. The score increases by correctly solving the element. If the element was not recalled correctly the score goes back to zero. A possible implementation of this system would now only present the learning elements with the lowest scores to the players. This will result in a personalized vocabulary database with different scores for each vocabulary for each individual player.

In *Number to Number Combat* the frequency of certain compositions can be personalized towards the players weaknesses. The compositions Addition, Subtraction, Multiplication and Division are clearly differently difficult and therefore the players should get in contact with them in the order easiest to hardest at first. But when already encountered all of them, measuring the performance on each of the compositions allows to determine the players weaknesses by comparing these results to a base value. This way a certain composition can be shown more frequently to the player if not yet well understood and compositions in which the player performs above the base value can be presented with a lower frequency to not annoy the player. In *Oppidum* the questions of the trivia quizzes can also be personalized by querying those with a low performance rate of the player with a higher frequency than those the player already knows perfectly.

6. HieroQuest: A Serious Game for Learning Middle Egyptian

HieroQuest is a Serious Game dedicated towards teaching Middle Egyptian in an immersive environment. The current version of the game is the product of an iterative refinement circle. While the first prototype of *HieroQuest* only featured the 25 single-literals of Middle Egyptian, the third iteration also includes 40 words and multi-literals and uses the ancient *Story of the Shipwrecked Sailor* to convey the learning content.

In this chapter a short explanation of Middle Egyptian (6.1.1) and the *Story of the Shipwrecked Sailor* (6.1.2) and why Hieroglyphs and Serious Game are a perfect match (6.1.3) are presented in Section 6.1. Followed by a presentation of the implementations and gameplay (6.2.1) and limitations shown by a small pilot-study (6.2.2) of the first iteration of *HieroQuest* in Section 6.2. After that the second iteration of the game will be described in Section 6.3, structured in implementations and gameplay (6.3.1) and limitations shown by a small pre-study (6.3.2). Finally the third iterations of the game is presented in Section 6.4 by describing the implementations (6.4.1) followed by the results of two user studies (6.4.2).

6.1. Background

In this section the historical background of *HieroQuest's* learning content ia presented. Starting with a short introduction to the Middle Egyptian language in Section 6.1.1 followed by an introduction to the *Story of the Shipwrecked Sailor* in Section 6.1.2 and an explanation why Hieroglyphs and Serious Games go hand in hand in Section 6.1.3.

6.1.1. A Short Introduction to Middle Egyptian

The Egyptian history can be categorized into three great epochs each with a different level of Egyptian language. Overall the language was used continuously for 5000 years. While from Old Egyptian (2700-2200 BC) to Middle Egyptian (2200-1350 BC) the grammar got more and more structured and the language as a whole started to follow unified characteristics across Egypt, these standards were again loosened up when Late Egyptian (1350 BC - 200 AD) arose [71]. Since Middle Egyptian follows those more sophisticated structures it acts as the perfect starting point when beginning to learn the Egyptian language.

Since the Egyptian language makes use of simplified representations of real objects it is often mistaken as a solely pictographic script. But also the phonetics of a Hieroglyph are important for their meaning. This is the so called Rebus principle [72]: The meaning of a single Hieroglyph is always dependent on the Hieroglyphs surrounding it. For example the Hieroglyph 𓃀, representing a foot, stands for place when taken alone, but in a word is detached from the object it represents and is only used as the phonetic *b*. Therefore a Hieroglyph can be used for different purposes. There exist different kinds of Hieroglyphs. The basic kind are the single-literals, those are 25 Hieroglyphs, which all translate to a single letter. But considerably more Hieroglyphs are multi-literals, those are Hieroglyphs which translate to two or more letters. Since aesthetics are very important for the Middle Egyptian language words can be written in different ways to ensure a perfect use of the available space. There exist short, medium and long versions of words, which differ in their length. Most of the time the letters contained in a multi-literal are repeated by writing the respective single-literals to ensure that the reader recognizes the correct meaning. The short, medium and long versions of a word are build by repeating none, few or all of the letters in a multi-literal.

6.1.2. The Story of the Shipwrecked Sailor

The *Story of the Shipwrecked Sailor* is an ancient story, which is used in academical teaching. The original story is written in Hieratic, a cursive writing system of the Middle Egyptian language. Since the story is already used in academical teaching it immediately provides relevant vocabulary for the Serious Game. But also the plot of the story can be utilized in a game. It is about a man returning from a failed voyage, who is scared to tell the king about it. His superior tells him a story of annother failed voyage in which he was able to overcome disaster. The superior was once the only survivor of a shipwreck. He was then stranded on an island, were he met a god in form of a snake. He was scared at first, but listened what the god had to say and this way was able to survive. But the attendant does not listen to the story and states that it wont matter to give water to a bird at the night before his slaughter. The moral of the story is to first listen, before jumping to conclusions. In the third iteration of the refinement circle *HieroQuest* teaches the plot of the story as well as the relevant vocabulary of the story (Section 6.4).

6.1.3. Hieroglyphs and Serious Games: A Perfect Match

Hieroglyphs and Serious Game are in general a good combination. In this section two main aspects of this connection are described. The first aspect is the rebus principle. Since a Hieroglyph is the simplified version of a real object, the Hieroglyph can be presented in the most natural way in a 3D game by showing the connection between object and Hieroglyph. Players get the opportunity to interact with the object or look at it from different angles. This allows for a deeper connection with the learning content and promotes remembrance.

The second aspect is that in most cases no pretest is necessary, when researching the learning outcomes, because most people don't know hieroglyphs. In user studies this was always verified after they played the game with a simple yes-or-no question and the results were discarded when a participant had previous knowledge, but this was not necessary at any point in time. The possible bias of a pretest on the study is removed when following the procedure described above. Furthermore, most people have no special interest in learning

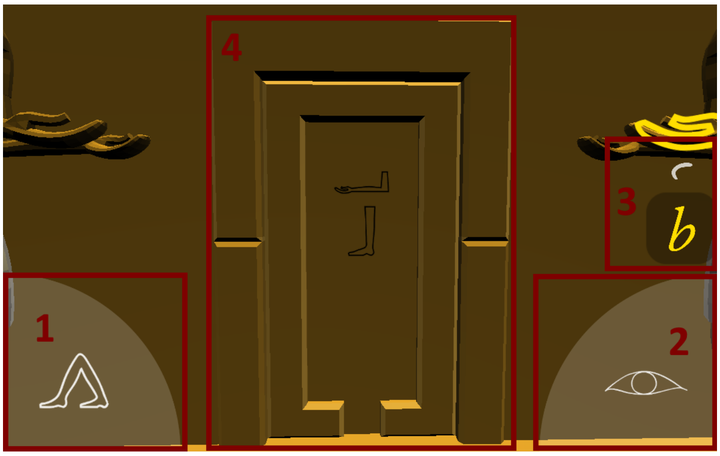


Figure 6.1.: *RiddleDoor*, *LiteralPicker* and Touch-Joysticks

Hieroglyphs and therefore their participation is mostly motivated by the gaming aspect, which allows to test the effectiveness of the game in conveying the learning content with only small motivational bias from intrinsic motivation. This makes research on a Serious Game dedicated towards teaching Middle Egyptian very unique, as it is hard to find learning content in which participants neither have special interest in learning it nor have previous knowledge, which is at the same time not overwhelmingly difficult. For example in mathematics or most languages people have previous knowledge either consciously or unconsciously as e.g. Romance languages are quite similar. Both aspects combined create a symbiosis where one profits form the other.

6.2. HieroQuest: First Iteration

In this section the implementations and the gameplay (6.2.1) and the limitations (6.2.2) of the first iteration of *HieroQuest* are summarized.

6.2.1. Implementations and Gameplay

The first prototype of *HieroQuest* only contained the 25 single literals of the Middle Egyptian language. These Hieroglyphs are taught to the players through various riddles. The game takes place in a fictive Egyptian temple structure. The goal of the player is to escape out of the temple. In order to do so, many different riddles have to be completed. Starting with a small labyrinth where the players get to know how to maneuver (Figure 6.1 annot. 1) and how to control the camera (Figure 6.1 annot. 2) with the Touch-Joysticks. After that the players learn basic interactions with objects by placing vases back into a shelf. Followed by the first kind of riddle with learning content (Figure 6.2 left), in which the players have to repair a broken statue by placing the missing arm and missing foot back onto it. In doing so the players learn the hieroglyphs 𓂝 - for "arm" and the transliteration *A* and 𓃀 - for

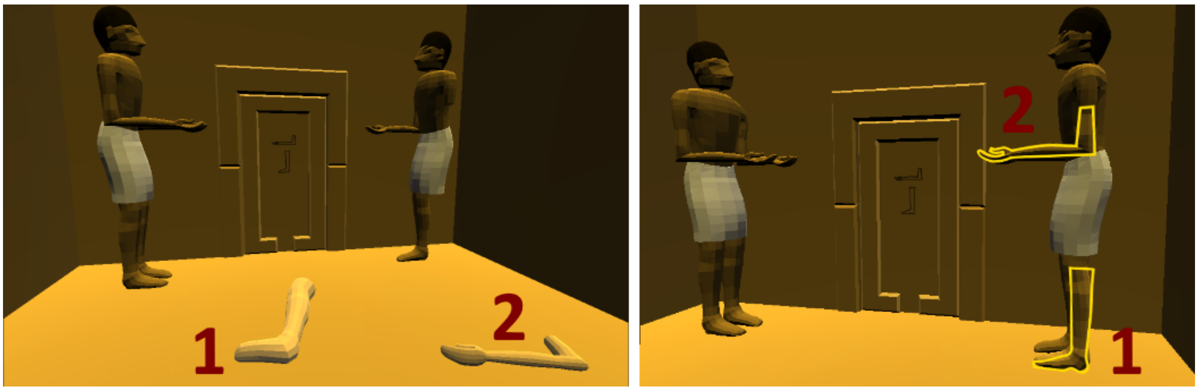


Figure 6.2.: *PickUpRiddle*: left unsolved, right solved

"foot" and the transliteration *b*. This type of riddle, *PickUpRiddle*, teaches the connection between Hieroglyphs and the real object in the most natural way by using the real object and after the riddle is solved placing the Hieroglyph on top of the object to show the connection (Figure 6.2 right).

After the players have learned a Hieroglyph its transliteration gets added to the so called *LiteralPicker* (Figure 6.1 annot. 3), which is the main tool the players use withing the game to interact with Hieroglyphs. The players can scroll through the *LiteralPicker* by dragging on it. The currently selected transliteration glows yellow. When players want to interact with a Hieroglyph or an object which a Hieroglyph represents they have to select the correct transliteration first in order to being able to pick up the object or in case of a Hieroglyph shown on a door solve a so called *TouchRiddle*. The later directly brings us to the next riddle type which is fundamental for the game: *RiddleDoors* (Figure 6.1 annot. 4).

A *RiddleDoor* builds the final challenge of a room. It consists of different *TouchRiddles*, in which the players have to select the transliteration of a hieroglyph shown and then click on it, this process is repeated for every Hieroglyph shown on the door. This makes sure that the players repeat previously learned transliterations and also ensures that the riddles of the room were completely solved by the players, by always querying all the Hieroglyph learned in this room on the door. The advantage of using such a progress control is not having to rely on invisible walls or boundaries to keep the player from moving on when a riddle was not solved.

The second type of natural riddles are the so called *WorldHighlightRiddles*, in which the players have to look at an object from a specific position and in a specific orientation so the Hieroglyph's outline matches the object it represents (Figure 6.3). While far away the Hieroglyph outline is almost invisible (Figure 6.3 annot. 1). When getting closer to the correct position the outline becomes more and more visible (Figure 6.3 annot. 2) until the correct position was found and the outline is fully visible (Figure 6.3 annot. 3). The outline can then be clicked and the solved state of the riddle is indicated by the yellow glow of the outline (Figure 6.3 annot. 4). In this case the players learn the Hieroglyph □ with the transliteration *p*, which represents the top of a stool or the reed mat on top of the stool. In the

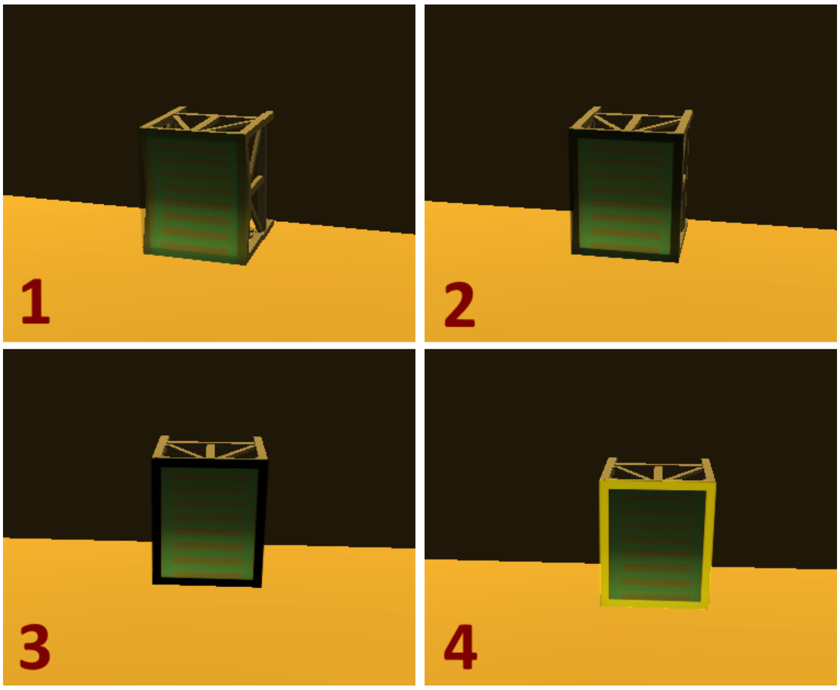


Figure 6.3.: *WorldHighlightRiddle*

special case of this very generic looking Hieroglyph ▯ it would not be easy to understand its origin without showing this connection between object and Hieroglyph.

In addition to the previously shown riddles, which show the connection between object and Hieroglyph in a natural way, also generic riddle types are included in *HieroQuest* because it is not always possible to replicate the object the Hieroglyph represents. The types of riddle are considered generic as it is possible to include any Hieroglyph in the riddle and no special 3D models are used.

In the so called *GenericSlideRiddle* the players have to bring a tiled picture of a Hieroglyph back into the correct order. To create the riddle one tile of the Hieroglyph was removed and then one of the adjacent tiles was moved into the empty spot. The sliding part is repeated until the desired difficulty for the riddle is reached. When approaching the riddle the players first have to recognize which slides have to be performed to solve it (Figure 6.4 annot. 1). The players then have to click on the tiles they want to move into the empty slot. After all the tiles have been moved into the correct spot the sliding part of the riddle is completed (Figure 6.4 annot. 2). In the last step the missing tile has to be inserted back into the empty space to fully complete the Hieroglyph (Figure 6.4 annot. 3).

The second generic riddle is the so called *GenericPickUpRiddle* (Figure 6.5). This riddle type also features a tiled picture of a Hieroglyph. In this case several tiles are missing and have to be put back into the correct spots by the player. This riddle is simpler than the *GenericSlideRiddle* and can therefore be used as an addition to other riddles.

The third generic riddle type shows a Hieroglyph covered by dust, the so called *Generic-*

Figure 6.4.: *GenericSlideRiddle*

DustOffRiddle (Figure 6.6 annot. 1). In order to solve the riddle the players has to swipe over the Hieroglyph to clean it off (Figure 6.6 annot. 2). When the dust is completely removed the riddle is solve and this status is indicated by the yellow glow of the Hieroglyph (Figure 6.6 annot. 3).

The fourth generic riddle focuses on the repetition of already learned transliteration, these are *BreakableVases*. These vases can either block the way of the player or contain items which are needed to solve riddles, therefore the player has to brake them (Figure 6.7 annot. 1). In order to do so the player has to select the transliteration of the Hieroglyph shown on the vase and then click on the vase. If the correct transliteration was selected the vase will brake (Figure 6.6 annot. 2), if a wrong transliteration was chosen an error sound is played, the vase will not brake and the player has to try again.

6.2.2. Pilot-Study and Limitations

To investigate the players acceptance and thoughts towards the game, 19 (14 male, 5 female) computer science students were invited to an informal pilot study. In this study the participants first played the game for approx. 40 minutes and were then asked to fill out a questionnaire. None of the participants had previous knowledge of Hieroglyphs. While the pilot study showed that the game was capable of transferring knowledge of the Middle

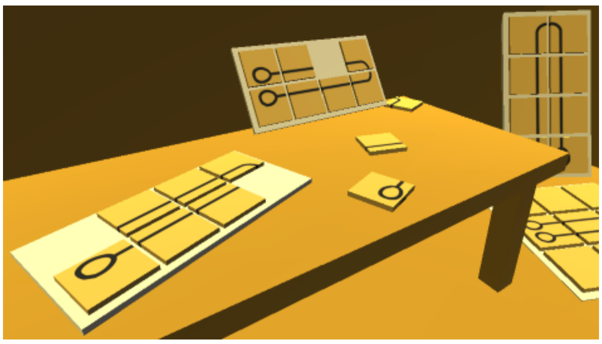


Figure 6.5.: *GenericPickUpRiddle*

Figure 6.6.: *GenericDustOffRiddle*

Egyptian language the limitations predominate.

The participants were especially frustrated with the lack of guidance and not having the possibility to look up Hieroglyphs they have already learned, which resulted in a trial and error method every time they had forgotten one. Furthermore it became evident that Hieroglyphs taught through the natural riddle types were memorized better than those taught through generic riddle types. Therefore in future iterations of the game generic riddles should only be used when the connection between Hieroglyph and object is clear and the riddle should at least be embedded in a meaningful context to hint towards the object. In total, the next iteration of the game has to be improved in terms of guidance and overall player experience.

6.3. HieroQuest: Second Iteration

While the first iteration built the basis the second iteration focused on the overall improvement of the player experience. The small pilot study showed that players felt a lack of guidance and the lack of possibility to repeat forgotten Hieroglyphs. The second iteration of the game also moved away from the temple structure and added a new outdoor level with a Nile boat tour, were the players can walk around in the desert to finally reach a pyramid.

6.3.1. Implementations and Gameplay

To counteract the lack of guidance several hint systems were implemented. The first one is a clickable Hieroglyph representing *Thoth*, the inventor of Hieroglyphs, which appears in

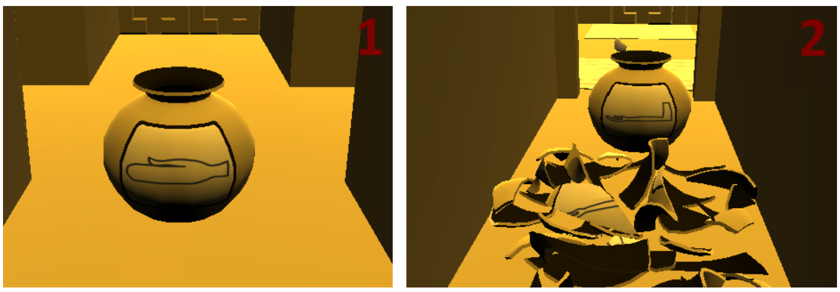


Figure 6.7.: *BreakableVases*

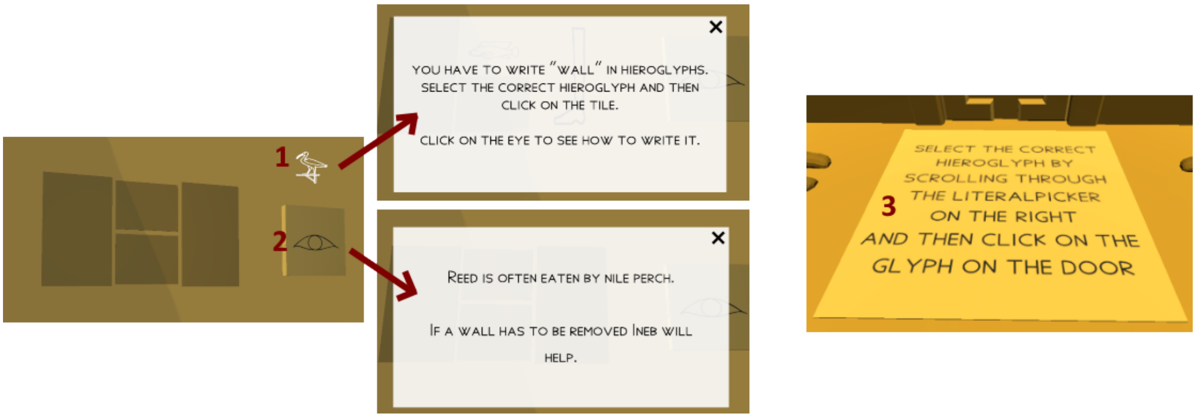


Figure 6.8.: *Thoth*, *MessageTiles* and *TutorialTiles*

the players field-of-view after a wrong action was performed or if no input was registered for a certain amount of time (Figure 6.8 annot. 1). The players can click on *Thoth* to get a hint on which actions they have to perform next. These hints are location based and are appropriate to the players current situation. If a room consists of more than one riddle the hints also account for that by keeping track of the status of the individual riddles. Additionally, *TutorialTiles* were added to the first rooms, which provide additional guidance in the first few rooms (Figure 6.8 annot. 3).

In case the players forget the transliteration of a Hieroglyph in the first prototype they had to try out all transliterations they have unlocked until they find the correct one. In the second iteration an in-game dictionary was added to the game, to help the players. This dictionary can be accessed via the menu and shows the players the Hieroglyphs they have learned so far, with their transliteration, the pronunciation and the object it refers to. With this tool the players are no longer required to use the trial and error method to find the correct transliteration and can instead look it up.

Because the outdoor level gives the players more opportunity to walk around it also increases the difficulty and the chance for the player to get lost. While the room structure of the temple for the most part only yielded one riddle per room, the outdoor level yields multiple riddles in an open space. Since the pilot study showed that players felt a lack of guidance additional hints in form of footprints, which guide the player to the next riddle were provided for the outdoor level. To not patronize the players with always visible footprints the player has to pick up a small statue representing *Ra*, the deity of the sun (Figure 6.9 annot. 1). With the item picked up the players can look through the sun disk on his head for footprints on the floor (Figure 6.9 annot. 2). These will guide the players towards the next riddle they can solve (Figure 6.9 annot. 3).

In the second iteration new riddle types were added to the game, which require the players to write words in Hieroglyphs. To make sure that the players know which words they have to write additional text in form of *MessageTiles* (Figure 6.8 annot. 2) is provided to them. These tiles are not to be mistaken with the hints. While the hints provide detailed instruction for

Figure 6.9.: *Ra* Statue with Guiding Footprints

the players, the *MessageTiles* are a cryptic riddle description and build a riddle in itself. They can also only be read with the transliteration of the Hieroglyph 𓆛 selected. For example in one of the riddles the players have to write the word "wall" in Hieroglyphs: 𓇋𓆛𓈖𓃀. The word contains 𓇋, representing a reed leaf, and 𓆛, representing a Nile perch, which the *MessageTile* hints towards by stating that "reed is often eaten by Nile perch".

The first iteration of *HieroQuest* only included the single-literals of the Middle Egyptian language, the next logical step to extend the learning content is to also include multi-literals and words. The second iteration includes two types of riddles in which the players have to write words. In the first one, the *GenericWordPutDownRiddle*, the players are provided with the Hieroglyphs that are required to build the word (Figure 6.10 annot. 1). So the task for the player is to put them into the correct order. Before being able to pick up the tiles, the respective transliteration needs to be selected in the *LiteralPicker*. The tile can than be picked up and placed onto an empty slot by clicking on the slot (Figure 6.10 annot. 2). If the order of the tiles was correct they will indicate this by glowing yellow (Figure 6.10 annot. 3), else the tiles will move back to their original positions. In the second one, the *GenericWordTouchRiddle*, the players are required to think about the required Hieroglyphs to type the word for themselves. The players need to select the transliteration of the desired Hieroglyph in the *LiteralPicker* and then click on the desired empty spot (Figure 6.11 annot. 2). When the word was built correctly all Hieroglyphs will glow yellow (Figure 6.11 annot. 3), if one of the Hieroglyphs was incorrect or at the wrong position all tiles will go back to the initial blank state and the player has to try again. In this special case the players have to write the name of the famous

Figure 6.10.: *GenericWordPutDownRiddle*

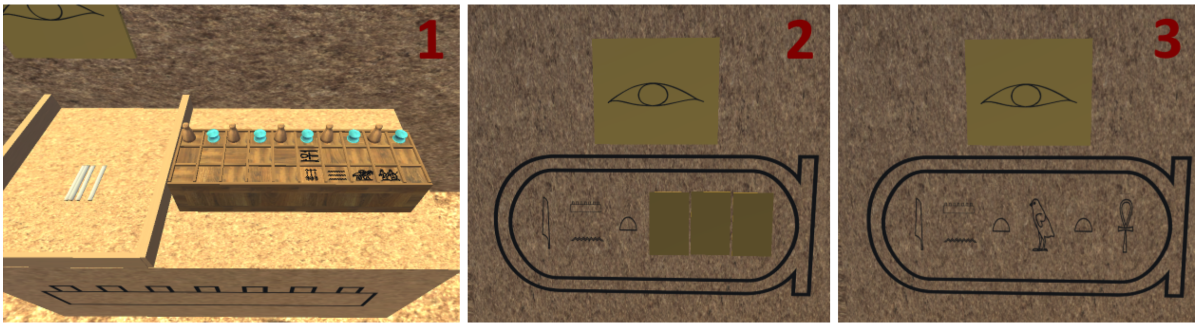


Figure 6.11.: *GenericWordTouchRiddle*

Pharaoh *Tutanchamun* 𓇋𓏠𓈖𓏏𓅱𓏏𓋹 which is written as "Amun-tut-anch". "Amun" means god and is therefore always put first, to write this part of the word the Hieroglyph 𓏠, which represents an old Egyptian board game called *Senet*, is required. In order to unlock this Hieroglyph the players have to win a game of *Senet* against the computer (Figure 6.11 annot. 1).

Another generic riddle type, the so called *GenericFlipRiddle*, was added to the game to include more variety in riddles. In this riddle type the players are presented with a 3x3 tiled Hieroglyph (Figure 6.12). As the name suggest the players have to perform several flips until the Hieroglyph is completely shown. A flip is performed by clicking on one of the tiles, which will cause the clicked tile and all adjacent tiles to turn (Figure 6.12 annot. 1). In order to solve the riddle several of these flips have to be performed (Figure 6.12 annot. 2 and 3). After all tiles show parts of the Hieroglyph and no more backsides are shown the riddle is solved, which is again indicated by a yellow glow (Figure 6.12 annot. 4). In this special case the players learn the Hieroglyph 𓈗, which represents water and therefore the tiles are also floating in water. All of the generic riddles which include tiles are non destructive, which means that the players can at no point bring the riddle into an unsolvable state, but they can make the riddle harder for themselves when performing wrong actions. To give the players the possibility to reset the riddle to its default state, a reset lever was added to all tile-based generic riddles. The players simply have to click on the lever and the riddle will turn back to its default state.

The first iteration of *HieroQuest* had no possibility for the players to save the current state of the game. While this was bearable with a playtime of approx. 40 minutes, with an additional outdoor world the players are not expected to complete the game in a single session. Therefore a saving system was integrated into the game to give the players the possibility to stop playing at any point and later resume from this exact point. A *MainMenu* was needed to let the players choose if they want to continue were they stopped or if they want to start a new game (Figure 6.13 annot. 1). In the course of this a system which lets the players replay certain parts of the game was created, as well. This is the so called *ChapterMode*, which separated the game into small chapters of two to eight Hieroglyphs (Figure 6.13 annot. 2). To give the players even more possibilities to train their knowledge without having to replay the same parts of the game over and over an additional mode was implemented, the so called

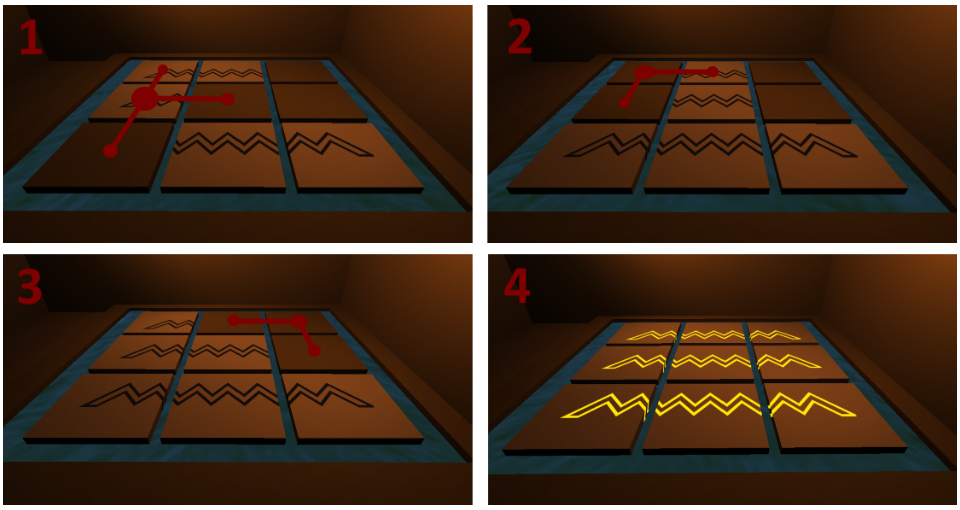


Figure 6.12.: *GenericFlipRiddle*

TrainingMode (Figure 6.13 annot. 3). The players can choose how long they want to train by specifying the amount of rooms they want to play. Each of these rooms includes one riddle and therefore teaches the players one Hieroglyph. These rooms are put together as a linear path, which is dynamically created every time the players enter the *TrainingMode*. The path does not require the players to go back to previous rooms at any point as the path is only build out of left or right turns and simple forward rooms, which makes the path sequential (Figure 6.14). The goal for the players is to get from the starting room (Figure 6.14 annot. 1) to the end room (Figure 6.14 annot. 2).

6.3.2. Pre-Study and Limitations

The more evolved state of the game allowed for a second user study. To evaluate the effectiveness of *HieroQuest* in transferring knowledge of the Middle Egyptian language 18 participants (13 male, 5 female) were invited to play the game. They had a mean age of 24.67 years (SD=2,16). Due to the increase in content, the evaluation was limited to only the single-literals, to keep the play-time at approx. 90 minutes. After they had played the game they answered a questionnaire regarding the learning content and their feelings while playing.

Figure 6.13.: *MainMenu*, *ChapterMode* and *TrainingMode*

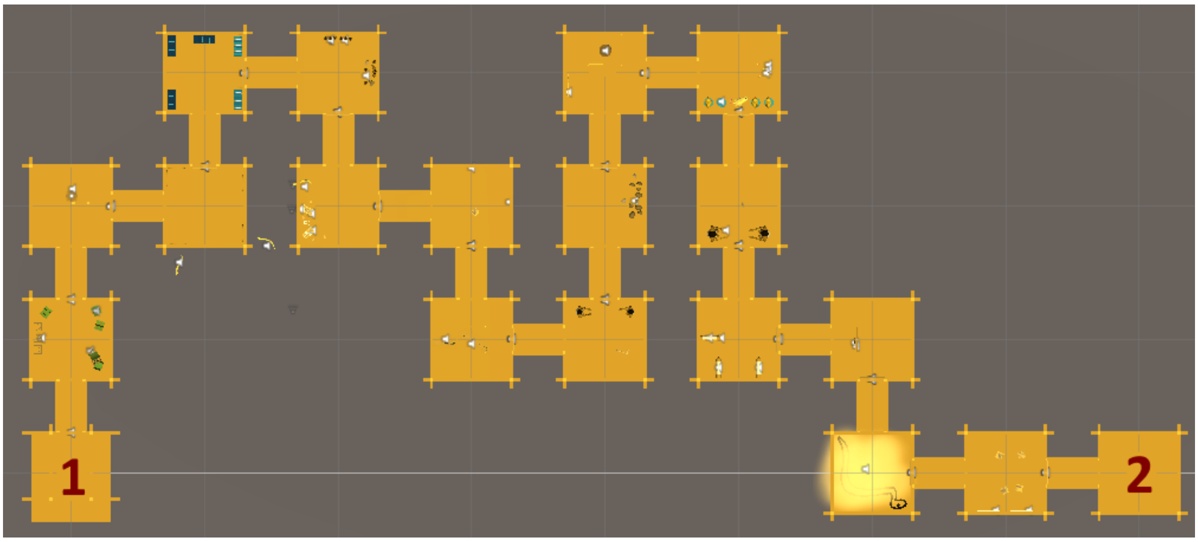


Figure 6.14.: *TrainingMode* Linear Map Generation

The participants were able to remember between 39% and 93% of the Hieroglyphs taught in the game. Which at first confirms that *HieroQuest* is capable of conveying knowledge of the Middle Egyptian language, but the large gap between the lowest and the highest scores also indicates possible flaws. When taking a look at the players feelings while playing the game and other data they provided, it became evident that those who play digital games for more than 60 minutes per week could remember significantly (t(16)=-4.93, p=0.00007) more Hieroglyphs (M=87.3%, SD=0.066, n=9) than the participants with less playtime (M=64.0%, SD=0.116, n=9). Almost the same effect was seen when separating the group of participants into two groups according to their feelings while playing the game, those were 25 Likert-Scale questions (from strongly disagree 1 to strongly agree 5) taken from the Flow State Scale [73, 35]. The average of score of a participant was then used as the flow rating. The participants with a rating higher than 3.5 remembered significantly (t(16)=-3.116, p=0.0033) more Hieroglyphs (M=84.9%, SD=0.082, n=9) than those with a lower rating (M=66.4%, SD=0.146, n=9). Because the groups of higher flow rating and more than 60 minutes of digital playtime per week overlap to 78%, this leads to the assumption that players with greater affinity towards games can focus on the learning content of the game while the others can concentrate less on the learning content because they have to put more effort in playing the game. Since this gap in remembrance rate is not desirable the guidance for novice players needs to be further improved to increase the overall effectiveness of the game.

6.4. HieroQuest: Third Iteration

The third iteration of *HieroQuest* introduced the *Story of the Shipwrecked Sailor* as new learning content. Since the story is also used in academical teaching, using it paves the way for the game to be used as a learning tool for Egyptology students. But before the game can fit this role additional steps towards the overall gaming experience especially for novice players have to be considered.

6.4.1. Implementations and Gameplay

With the overall goal of the game to also function as a learning tool for Egyptology students in mind the second level, which got added in the second iteration, was discarded. While the riddle types and additional guidance are useful additions, the learning content and the more opened up world do not fit this overall goal. The learning content was chosen rather arbitrary and the opened up world puts additional cognitive load on the players, which is especially frustrating for novice players. The third iteration therefore focuses on the single-literals and words and multi-literals from the *Story of the Shipwrecked Sailor* to ensure their relevance, while still keeping the additional guidance and improved player experience of the second iteration.

To enhance the player experience even further a new graphical appearance was introduced. While the first two iterations included only flat colors on walls, floors and objects the third iteration includes Egyptian themed textures to increase the overall immersion (Figure 6.15). Adding onto this, the *Story of the Shipwrecked Sailor* is not only used for the vocabulary, but also the plot of the story itself will be utilized to increase the players immersion. The story is told to the player through the so called *PapyriSystem*. 21 Papyri are placed inside the new level, to present the players the plot of the story in English, which includes the English translation of the words they learn and a pictorial representation of the Hieroglyphs (Figure 6.15 annot. 1 and Figure 6.16 annot. 1). A similar system is used to present the players with grammatical phenomena (Figure 6.15 annot. 2 and Figure 6.16 annot. 2).

Several new riddle types were added to the game, while the two word writing riddles of the second iteration were also kept. The first new riddle type is the so called *ScaleRiddle* (Figure 6.17 annot. 1). The *Story of the Shipwrecked Sailor* also includes several numbers e.g. the length of the ship is 120 cubits. To introduce the players to the Egyptian number system they have to place weights onto a large scale until both sides bear the same weight (Figure 6.17 annot. 2). In doing so the players get to know the Egyptian numbers from 1 to 9999, with the

Figure 6.15.: New Graphics with Story and Grammar Papyri

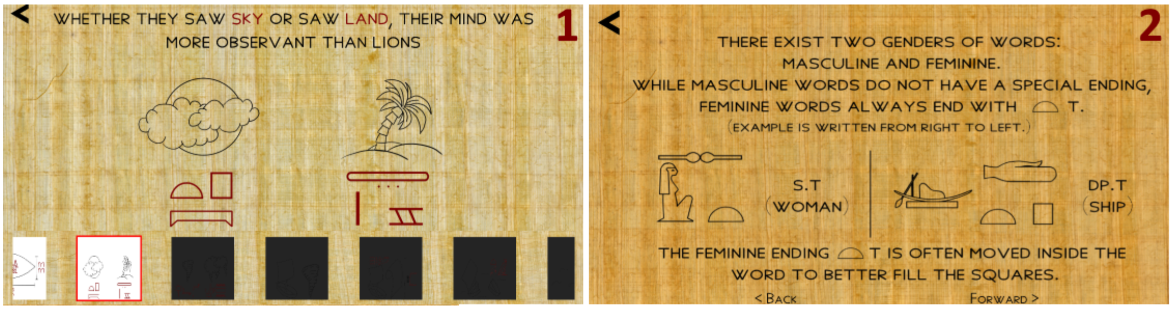


Figure 6.16.: Story and Grammar *PapyriSystem*

four Hieroglyphs 𓏺 (1), 𓎆 (10), 𓍢 (100) and 𓆼 (1000).

The second new riddle type is the so called *HarpRiddle*, this riddle features a large harp which plays a melody indicated by the strings of the harp glowing yellow in a certain order (Figure 6.18 annot. 1). The players have to remember this and than click on each of the strings by paying respect to the order. If they clicked a string correctly it will glow green for a short amount of time and the players can continue with the next one (Figure 6.18 annot. 2), if the string was wrongly chosen it will glow in red and the current progress will be removed (Figure 6.18 annot. 3). If the players want to listen to the correct order of strings again they just have to wait for a certain amount of time. If the complete melody was right the players will in this case be rewarded with the verb "to hear" 𓄔𓅓.

The third new riddle type is the so called *SeesawRiddle*, this riddle features four seesaws which have to be put into the correct position. By moving one of the seesaws one or multiple tiles with Hieroglyphs on them will move, when these tiles are all in the correct positions to build a word the riddle is solved (Figure 6.19) and the players are in this case rewarded with the verb "to stretch" 𓊃𓈖𓂻.

While these three new riddle types can be considered generic they are only used in their respective contexts. But also two generic riddles were added to the game. The first one is the so called *TetrisRiddle*. In this riddle the players have to build a word which is made out of several tiles, the difference between the already existing tile-based *GenericPickUpRiddle* is that

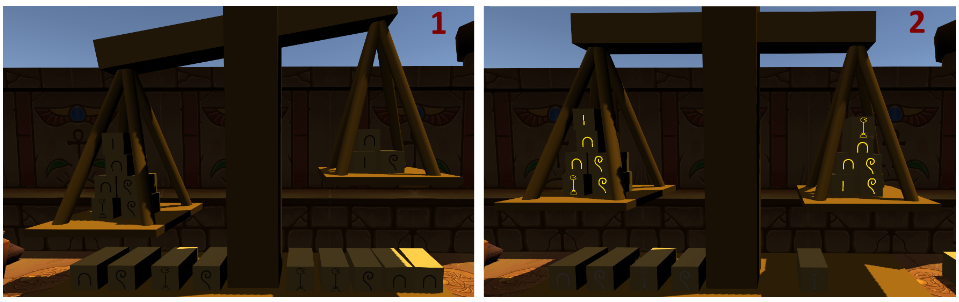


Figure 6.17.: *ScaleRiddle*

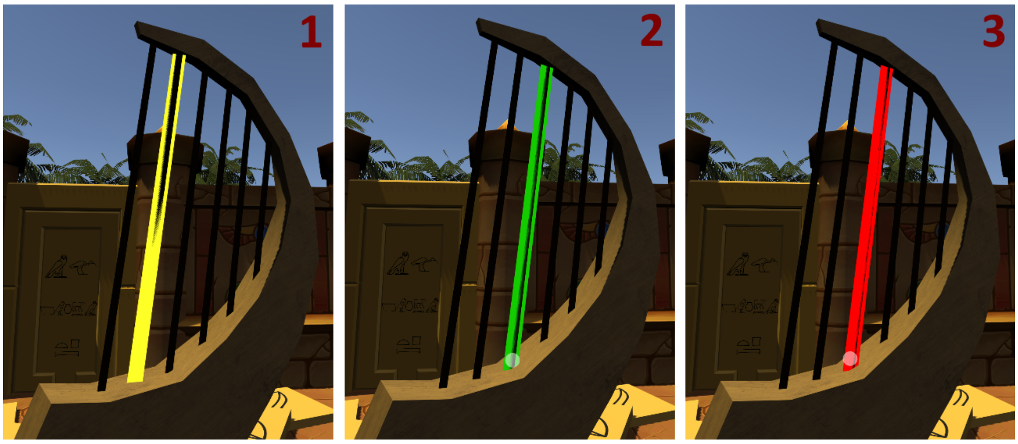


Figure 6.18.: *HarpRiddle*

the tiles have different shapes and therefore only tiles with the correct shape and the correct content can be inserted into an empty spot. The riddle will accept any tile with the correct shape regardless of the content at first, but when the word is fully built will verify the content of the tiles and if one is wrong all tiles will be placed back into their original positions. If the content of all tiles is correct the tiles will glow yellow and in this special case the players will be rewarded with the noun "snake" . If a player recognizes a mistakenly placed tile, it is possible to pick it back up and this way remove it from the riddle.

The second new generic riddle type is the so called *MemoryRiddle*. As the name states in this riddle the players have to play memory. When entering the room they are shown all tiles with the Hieroglyph side upwards, but after a certain amount of time all the tiles will flip and now the players have to find matching pairs until all tiles are again facing upwards. After the riddle is completed only certain Hieroglyphs will glow yellow and these then have to be inserted into a word riddle, in this special case the player will learn the verb "to know" .

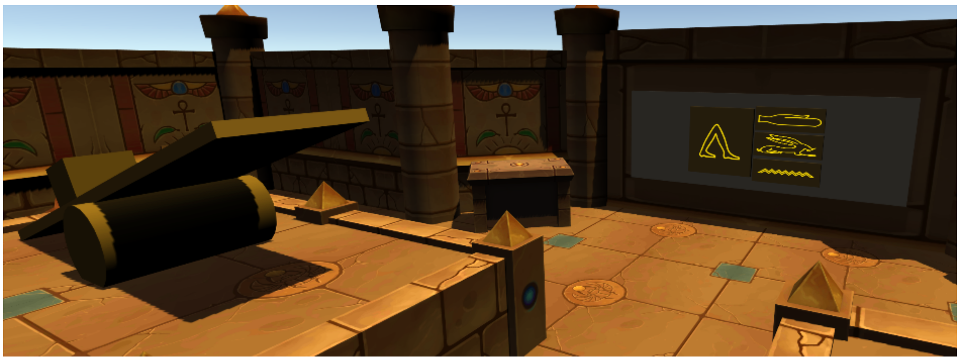


Figure 6.19.: *SeesawRiddle*

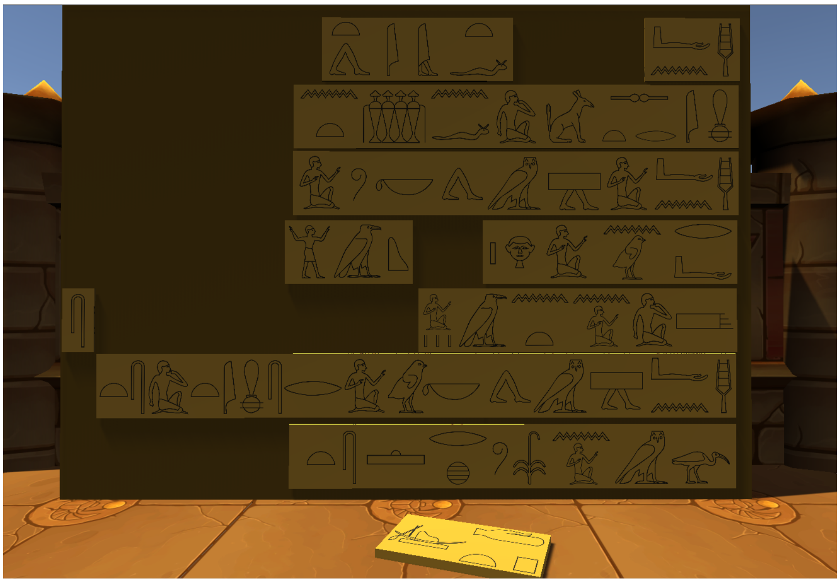


Figure 6.20.: *StoryRiddle*

The third new generic riddle type is the so called *StoryRiddle*. In this riddle type the players have to fill empty spots in the *Story of the Shipwrecked Sailor* itself (Figure 6.20). The missing pieces are obtained through various different riddles and only consist of words the players have already learned and therefore the correct transliteration of the words has to be selected in the *LiteralPicker* before the players are able to pick them up. This riddle is therefore used as a repetition.

Because the length of the game increased to a total play-time of approx. four hours a special evaluation level consisting of single-literals, multi-literals and words was implemented to make future user studies still possible. In this special level 13 words and the 14 single-literals used to write them are combined with 10 story papyri. Overall the evaluation level has a playtime of approx. 90 minutes.

6.4.2. User Study and Limitations

The third iteration of *HieroQuest* improved the game in many aspects, which made a larger user study possible. This study is split into two studies with participants from different backgrounds. While the participants of Study I (Section 6.4.2) were Computer Science students, which had general interest in gaming, but had no previous knowledge in Middle Egyptian, the participants of Study II (Section 6.4.2) were Egyptology students, which had previous knowledge in Middle Egyptian, but no general interest in gaming. In both studies the participants played two different versions of the evaluation level in a between-subject study design. The different versions of the evaluation level only differed in one game element. While in version A the participants were provided with a permanent hint, showing the corresponding Hieroglyph to the currently selected transliteration, in version B this hint was not visible and the participants had to rely on the in-game Dictionary. The participants of

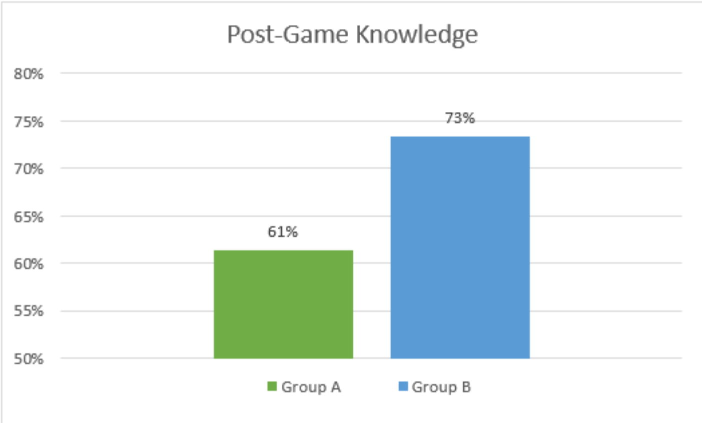


Figure 6.21.: Post-Game Knowledge Study I

version A were also free to use the Dictionary, but could also solely rely on the permanent hint. In general, the participants of version B could chose for themselves at which point in time they wanted to get a hint by opening the vocabulary and the participants of version A had a hint at all points in time [74]. The participants of both studies were invited to play the evaluation level on a tablet device. Since in Study I none of the participants had previous knowledge in Middle Egyptian, no pre-test before playing the game was necessary. In Study II a knowledge test before playing was necessary, as all of the participants had previous knowledge. This pre-game questionnaire consisted of 27 questions querying knowledge about the 27 Hieroglyphs included in the evaluation level and 2 questions querying knowledge about the story of the shipwrecked sailor (Section A). In any case the participants then played the evaluation level until they were finished or the time-frame of 90 minutes was exceeded. After playing they filled out the post-game questionnaire consisting of the 33 questions of the Game Experience Questionnaire (GEQ) Core Module [47], 27 questions querying knowledge about the 27 Hieroglyphs included in the evaluation level and the 2 questions querying knowledge about the story of the shipwrecked sailor (Section B). To keep any possible bias caused by the pre-game questionnaire as low as possible the order of the questions and the order of the answers regarding the hieroglyphs and the story were changed [74].

User Study I

In Study I 30 participants without previous knowledge of Middle Egyptian, but with digital gaming experience, Computer Science students, were separated randomly in Group A (male=13, female=2) and Group B (male=11, female=4). Group A played version A of the evaluation level, with the permanent hint and Group B played version B, without the hint. The participants of Group A had a mean age of 22.73 (SD=2.08) and on average play 11.33 (SD=9.29) hours of digital games per week. The participants of Group B had a mean age of 22.53 (SD=2.58) and on average play 11.27 (SD=7.63) hours of digital games per week.

The knowledge of the participants after they played the game was directly derived from

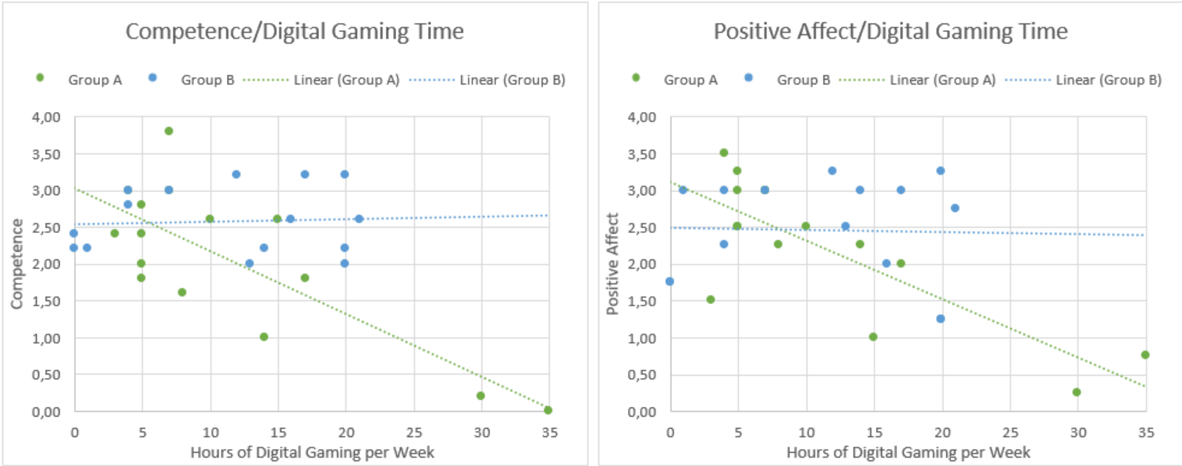


Figure 6.22.: Competence and Positive Affect to Digital Gaming per Week

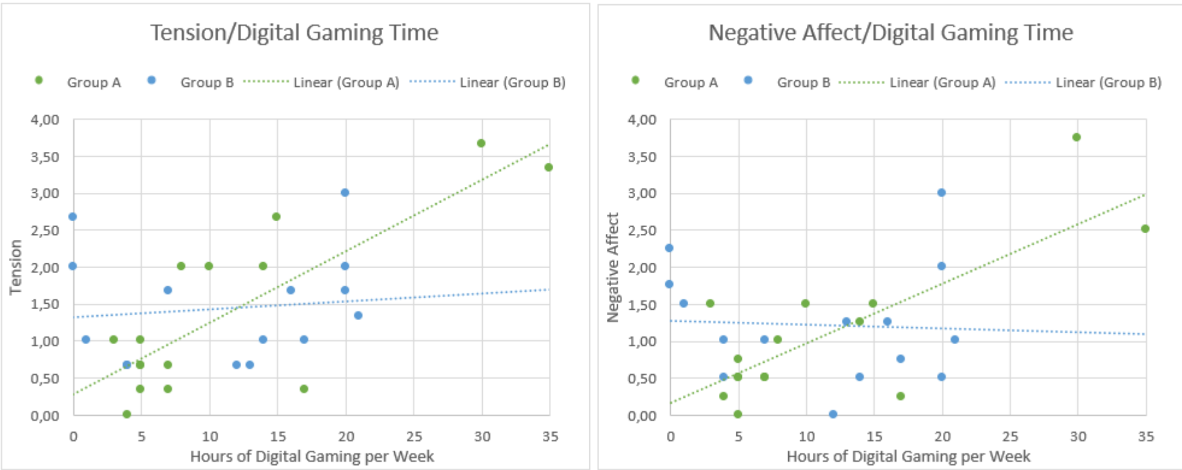


Figure 6.23.: Tension and Negative Affect to Digital Gaming per Week

the post-game questionnaire because none of the participants had knowledge of Middle Egyptian before playing the game. While Group A achieved a score of 61.4% (SD=0.12), Group B achieved a score of 73.3% (SD=0.09) (Figure 6.21). The game is therefore able to convey knowledge of Middle Egyptian to its players. The score of Group B was significantly higher than the score of Group A (t(28)=1.70, p=0.003), which shows the positive impact of the absence of a permanent hint and instead the possibility for the participants to autonomously decide at which point to get a hint was beneficial for the knowledge transfer [74].

Taking a closer look at the questions of the GEQ and the respective categories they belong to (Positive categories: competence, immersion, flow and positive affect; Negative categories: tension, challenge and negative affect)[47] and the playtime of digital games per week shows several correlations. The difference in correlations between *Competence and Gaming Experience* is significant (Group A: r=-0.789; Group B: r=0.06; z=-2.764, p=0.003) (Figure 6.22 left). For

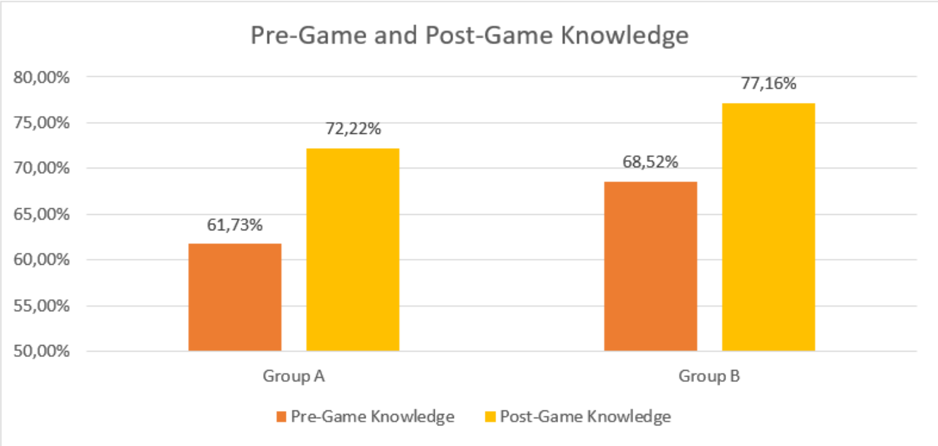


Figure 6.24.: Pre-Game and Post-Game Knowledge Study II

Positive Affect and Gaming Experience a similar and also significant difference in correlation can be found (Group A: r=-0.797; Group B=-0.003; z=-2.661, p=0.004) (Figure 6.22 right) [74].

A reversed but yet also significant difference can be found for two of the negative categories. The difference in correlations between *Tension and Gaming Experience* is significant (Group A: r=0.803; Group B: r=0.105; z=2.452, p=0.007) (Figure 6.23 left). For *Positive Affect and Gaming Experience* a similar and also significant difference in correlation can be found (Group A: r=0.783; Group B: r=0.05; z=2.702, p=0.003) (Figure 6.23 right) [74].

These significant differences all show the same phenomena, while the players of version B show almost no correlation between the four categories and their gaming experience, the players of version A with previous gaming experience did not feel challenged by the game. The controls and logical riddles of *HieroQuest* in general do not challenge players with a lot of experience because the game should also be usable by novice players. Therefore in both versions players with a lot of gaming experience are expected to be under challenged by the logical riddles. In version B this lack of challenge in the logical riddles cannot be seen in the correlations between the four categories and the gaming experience. But since the two version of the game only differ in the Hieroglyph hint the increased difficulty in the learning domain most likely counteracted the effect. Overall these correlations are not desirable and are a clear limitation of the game in the third iteration [74].

User Study II

In Study II 12 participants with previous knowledge of Middle Egyptian, but almost no digital gaming experience, Egyptology students, were separated randomly in Group A (male=1, female=3, diverse=2) and Group B (male=2, female=4). Group A played version A of the evaluation level, with the permanent hint and Group B played version B, without the hint. The participants of Group A had a mean age of 30.5 (SD=2.3) and on average play 2.0 (SD=2.31) hours of digital games per week. The participants of Group B had a mean age of 25.5 (SD=5.56) and on average play 3.67 (SD=3.5) hours of digital games per week.

The knowledge of the participants after they played the game, the post-game questionnaire, had to be put into perspective with the participants previous knowledge because all of the participants had knowledge of Middle Egyptian before playing the game. Group A had a previous score of 61.7% (SD=0.10), Group B had a previous score of 68.5% (SD=0.12) (Figure 6.24). For both groups this score significantly increased (Group A: t(6)=2.02, p=0.01; Group B: t(6)=2,015 p=0.004)) to a score of 72.2% (SD=0.14) for Group A and for Group B to a score of 77.2% (SD=0.11) (Figure 6.24). The game is therefore able to convey knowledge of Middle Egyptian to players with previous knowledge of Middle Egyptian. The difference between the gained knowledge of the two groups was not significant (Group A: 10.5%, SD=0.07; Group B: 8.64%, SD=0.05; t(10)=1.81, p=0.31). This is probably caused by the lack of previous gaming experience and the intrinsic motivation to learn the Hieroglyphs as they are needed for their course of study [74].

7. HieroQuest as a Dynamic Learning Environment

In the following chapter the transformation of *HieroQuest* from a static to a dynamic learning environment will be shown. Since *HieroQuest* is an already existing Serious Games several steps had to be performed to make this transformation under the criteria of the CDDA possible (Section 5). The process of identifying the respective domains of the riddles and elements within the game up to a game level which fully implements CDDA will be explained in detail in Section 7.1. Followed by a simple and a complex example of the CDDA implementation in two rooms of *HieroQuest* in Section 7.2. To fully support the dynamic character of the CDDA the previously existing *TrainingMode* (Section 6.3) was completely reworked to function as a dynamic learning environment with replayability, explained in detail in Section 7.3.

7.1. CDDA in HieroQuest

Implementing CDDA in an already existing Serious Game first requires the identification of dynamically adjustable elements and the categorization of theses into the learning or gaming domain (Section 5.2.1). This is a non trivial process. The riddles in *HieroQuest* , which require the knowledge of Hieroglyphs and therefore require the player to first select a transliteration in the *LiteralPicker* before being able to interact with the element or riddle, are classified into the learning domain (Table 7.1). Those which do not require the selection of a transliteration are classified into the gaming domain (Table 7.2). This classification is inline with the CDDA guidelines, were the element should be classified into the domain they mostly contribute to and secures a consistent separation (Section 5.2.1).

The second step of the CDDA is the definition of measurements to determine the player's perceived challenges in the two domains (Section 5.2.2). An in-game measurement for both domains, which directly measures the performance on every dynamically adjustable element was chosen, to ensure that the measured performance of a domain is only determined by the elements of the domain. The measurement is based on the number of failed attempts or the time it took the players to solve the element. A detailed explanation is provided in step five, when determining the update function.

The third step of the CDDA is the determination of the different difficulty levels for the riddles regarding their domain (Section 5.2.3). To ensure the consistency the number of actions needed to solve a riddle was used to change the difficulty level of that riddle. Three difficulty levels for both domains were implemented: easy ($[0, 1[$), medium ($[1, 2[$) and hard ($[2, 3]$). The already existing riddle difficulties and their number of actions required were used as the medium difficulty for that riddle and the easy and hard difficulty were derived from that by decreasing/increasing the number of required actions. It is important that the change in the number of required actions is compliant with the actions needed for the riddle. For example it is not comparatively to define three difficulty levels for one of the riddles as 2, 10 and 12 necessary actions because the change from easy to medium is much larger than the change from medium to hard. In addition, since the difficulty of a certain riddle can increase drastically when requiring an additional action, while for annother riddle this will barely be the case, the different difficulty levels have to be defined separately for each of the riddles and have to be consistent for that riddle type throughout the game. In this process it also needs to be ensured that the overall difficulty of a room does not differ by a large amount compared to a room with the same difficulty level for both domains, while a linear increase in difficulty within a certain difficulty level over the course of the game is intended and indispensable as the players get introduced to more complex content in both domains. Furthermore it is not always possible to dynamically change the difficulty of a certain riddle type, if the riddle is elementary and only requires one action. In the gaming domain this is the case for the *GenericDustOffRiddle* (Table 7.2), which only requires one swiping action. In the learning domain this is the case for *BreakableVases*, *TouchRiddles* and tiles showing a complete Hieroglyph, *TileRiddles*, (table 7.1), which all only requires the selection of one transliteration. It is only possible to increase or decrease the difficulty level of these elementary riddle types by increasing or decreasing their number of occurrence within a room. For example, this is done on the Riddle Door of a room, which requires different amounts of *TouchRiddles*, dependent on the learning difficulty level. For the complete level the different amounts of actions required sum up to approx. 80 for easy, 120 for medium and 180 for hard regarding the learning domain and approx. 100 for easy, 160 for medium and 220 for hard regarding the gaming domain.

The fourth step of the CDDA creation is the separation of the level into disjunctive parts. The difficulties of a part in the learning and the gaming domain are calculated according to the performance of the players in the previous part (Section 5.2.4). Since the levels of *HieroQuest* feature a temple structure with several rooms and for the most part one riddle per room, it makes sense to use the rooms as the disjunctive parts. A room is fully completed

Riddle Type	Dynamic Difficulty	Dynamic Hieroglyph
BreakableVase	No	Yes
TouchRiddle	No	Both
TilesRiddle	No	No
RiddleDoor	Yes	Yes
GenericWordPutDownRiddle	Yes	No
GenericWordTouchRiddle	Yes	No
ScaleRiddle	Yes	No
StoryRiddle	Yes	No

Table 7.1.: Learning Domain Riddles

Riddle Type	Dynamic Difficulty	Dynamic Hieroglyph	Switchable
GenericPickUpRiddle	Yes	No	Yes
GenericSlideRiddle	Yes	No	Yes
GenericFlipRiddle	Yes	No	Yes
TetrisRiddle	Yes	No	No
PickUpRiddle	Yes	No	No
SeesawRiddle	Yes	No	No
WorldHighlightRiddle	Yes	No	No
HarpRiddle	Yes	No	No
MemoryRiddle	Yes	Yes	No
GenericDustOffRiddle	No	No	Yes

Table 7.2.: Gaming Domain Riddles

when the door, which leads to the next room, is opened by the player. Since the main riddle of the room has to be completed before the players are able to open the door this will act as the action based transition two parts (Section 5.2.4). Since rooms can contain riddles, e.g. *BreakableVases*, that are only placed inside the room to cause distraction and will maybe not be solved by the players, the performance of the players will only be calculated according to the solved riddles.

The fifth step is the definition of a formula to calculate the overall room difficulty for both domains (Section 5.2.5). This formula is then used by the game in the exact second the player correctly solves the last *TouchRiddle* on the *RiddleDoor* of a room to calculate both difficulties for the next one. At the start of the game the difficulty levels for both domains are set to 1.5, corresponding to the medium difficulty. First the mean solve rate has to be calculated for both domains. This is done by first counting the number of all attended riddles for both domains, those are the riddles the player tried to solve. Then for each of the attended riddles the solve rate is calculated: If a player has failed a riddle more than 4 times before solving it the solve rate of that riddle is 0. If a player manages to solve the riddle on the first try the solve rate of that riddle is 1. In between the solve rate decreases by 0.2 per failed attempt. These solve rates are than summed up and the result is divided by the amount of attended riddles. All of the calculations are performed for both domains separately, but with the same formula, where one uses the attended riddles of the learning domain (L) and the other uses the attended riddles of the gaming domain (G) ($X \in \{L, G\}$):

Per X in the current room:

$n_X =$ number of failed trials on this X

$$g_X(n_X) = \begin{cases} (5 - n_X) * 0,2 & n_X < 5, \\ 0 & n_X \geq 5. \end{cases}$$

Mean X result function of the current room:

$$\vec{v_X} = \begin{pmatrix} g_1(n_X) \\ \vdots \\ g_m(n_X) \end{pmatrix}$$

$$d_X = \frac{\sum_{k=1}^{m} \vec{v_{X_k}}}{m}$$

The difficulty of the learning domain in the next room is then calculated by using the mean solve rate and the learning difficulty of the current room. If the mean solve rate is below 0.6 the learning difficulty will decrease according to the difference between the mean solve rate and 0.6. If the mean solve rate is between 0.6 and 0.8 the learning difficulty will not change and if the mean solve rate is above 0.8 the learning difficulty will increase according to the difference between the mean solve rate and 0.8, which is calculated by the following formula:

$$x_L = \text{Learning difficulty of the current room}$$

$$f_L(x_L, d_L) = \begin{cases} x_L + (d_L - 0,6) & d_L < 0,6, \\ x_L & 0,6 \leq d_L \leq 0,8, \\ x_L + (d_L - 0,8) & d_L > 0,8. \end{cases} \quad (7.1)$$

The difficulty of the gaming domain in the next room is then calculated by using the mean solve rate and the gaming difficulty of the current room. Additionally, the time it took the player to complete the attended riddle within the gaming domain is used because some of the riddles cannot be failed and therefore solely the time it took to complete them can be used. If the mean solve rate is below 0.6 the gaming difficulty will decrease according to the difference between the mean solve rate and 0.6. If the mean solve rate is between 0.6 and 0.8 the gaming difficulty will not change and if the mean solve rate is above 0.8 the gaming difficulty will increase according to the difference between the mean solve rate and 0.8, which is calculated by the following formula:

$$x_G = \text{Gaming difficulty of the current room}$$

$$t_{G_{mod}}\left(t_{G_p}, d_G\right) = \begin{cases} \frac{t_{G_p}}{t_{G_e}} & \lfloor x_G \rfloor = 0, \\ \frac{t_{G_p}}{t_{G_m}} & \lfloor x_G \rfloor = 1, \\ \frac{t_{G_p}}{t_{G_h}} & \lfloor x_G \rfloor = 2. \end{cases}$$

t_{G_p}=time it took to finish all G riddles of the room t_{G_e}=G riddle reference time difficulty easy

t_{G_m}=G riddle reference time difficulty medium t_{G_h}=G riddle reference time difficulty hard

$$d_{G_{total}}\left(d_G, t_{G_{mod}}\right) = \begin{cases} d_G + (1 - t_{G_{mod}}) & t_{G_{mod}} < 0,8 \\ d_G & 0,8 \leq t_{G_{mod}} \leq 1,2 \\ d_G + (1,2 - t_{G_{mod}}) & t_{G_{mod}} > 1,2 \end{cases}$$

$$f_G\left(x_G, d_{G_{total}}\right) = \begin{cases} x_G + (0,6 - d_{G_{total}}) & d_{G_{total}} < 0,6 \\ x_G & 0,6 \leq d_{G_{total}} \leq 0,8 \\ x_G + (d_{G_{total}} - 0,8) & d_{G_{total}} > 0,8 \end{cases} \tag{7.2}$$

The formulas f_L (Equation 7.1) and f_G (Equation 7.2) calculate the difficulty for the next room in both domains. The formulas are designed according to the guidelines of the CDDA (Section 5.2.5), where a decrease in difficulty should occur significantly faster than an increase in difficulty because overwhelmed players should reach their appropriate difficulty level within the first minutes of play. Therefore an increase in difficulty can only be achieved when constantly performing good, mean solve rates of more than 80%, and a decrease in difficulty is already caused by mean solve rates below 60% in both domains.

7.1.1. Personalized Gaming Riddles

Since previous studies on *HieroQuest*, in particular the open-ended questions in the user studies of the second and third iteration querying what the player especially like and disliked about the game, have shown that some of the participants do not like the *GenericFlipRiddles* or the *GenericSlideRiddles* (Section 6.3.2 and Section 6.4.2) and since these riddles types are anyways generic, they can be switched dynamically on the personal preference of the player. The guidelines of the CDDA suggest that certain tasks can be perceived differently difficult by players of the same overall skill level, which can especially be the case for logical riddles, where certain players do not have a solution strategy (Section 5.3). To counteract possible bias due to frustration caused from these riddle types, because they may be perceived much harder than intended, a system was implemented, which tracks the players performance on these two riddles types and dynamically changes the occurrence of them accordingly. At those spots within the level where a *GenericFlipRiddle* or a *GenericSlideRiddle* is placed the game dynamically decides which one to place according to the players previous performance on them. In general it would also be possible to place a *GenericPickUpRiddle* instead, but since its difficulty is much lower compared to the other two this would violate the intended linear difficulty (Chapter 5). At the start of the game either of the two will occur with equal probability. After the player has solved one of them the time it took them will be compared to a reference time and according to the players performance the probability of occurrence of this riddle type will either be increased or decreased.

7.1.2. Personalized Learning Content

By implementing the CDDA in *HieroQuest* the amount of riddles in a room within both domains are dependent on the respective difficulty level. When increasing the amount of riddles or actions required to solve a riddle in the learning domain the number of transliterations, which the players have to select in order to solve the riddles, increases. For example, a *RiddleDoor* in the medium learning difficulty requires the player to solve three *TouchRiddles*, which in total requires the selection of three correct transliteration. In the hard learning difficulty the same door would require four and in the easy learning difficulty only two correct transliteration. Since the *RiddleDoors* act as the progress control of a room and therefore only has to contain the Hieroglyphs taught in this room, most of the time one, the other Hieroglyphs can be chosen dynamically. To increase the potential learning success of a player the Hieroglyphs used in riddle types that allow for a dynamic choice were chosen dynamically. This is mainly the case for *RiddleDoors* and *BreakableVases*. Because the choice of Hieroglyphs should not be arbitrary but rather focus on the weaknesses of the players a spaced repetition technique the "Leitner System" [70], was used to choose the appropriate Hieroglyphs. This systems gives every Hieroglyph a certain score, starting at zero. When the transliteration of this Hieroglyph was correctly chosen at the first time of attending the riddle the score will increase by one, but when the transliteration is wrongly chosen, even once, the score of that Hieroglyph will again be set to zero. This will lead to a personalized Hieroglyph database for an individual player. When the difficulty of the next room is calculated also the appropriate Hieroglyphs used in the room are chosen. This is achieved by identifying how many dynamic Hieroglyphs are needed for the next room and than the Hieroglyphs with the lowest scores are chosen and placed inside the riddles. When there are more Hieroglyphs with the lowest score than needed for the next room they will be randomly sampled out of those. When there are less Hieroglyphs with the lowest score than needed for the next room the next score bracket will also be included. If this will increase the amount of Hieroglyphs to more than necessary, all Hieroglyphs with the lowest score are included and the missing Hieroglyphs will be randomly sampled from the next bracket. This dynamic choice of Hieroglyphs tries to maximize the learning success, while decreasing the possible frustration caused by having to insert the transliterations of Hieroglyphs which are already well known again and again.

7.1.3. Other Dynamics

The hint system which was introduced in the second iteration of *HieroQuest* (section 6.3), was also made dynamic in the course of implementing the CDDA. Since the hints only help with riddles of the gaming domain they are only dependent on the gaming difficulty of a room. When the gaming difficulty of the current room is set to easy the hints provided by the game are more descriptive and try to help the players more than the hints provided by the game when the gaming difficulty is set to hard. Nevertheless, the hints still follow a cascading system, were not all the information is provided in the first hint, so the players are still required to think for themselves. The hints will contain more information as the players

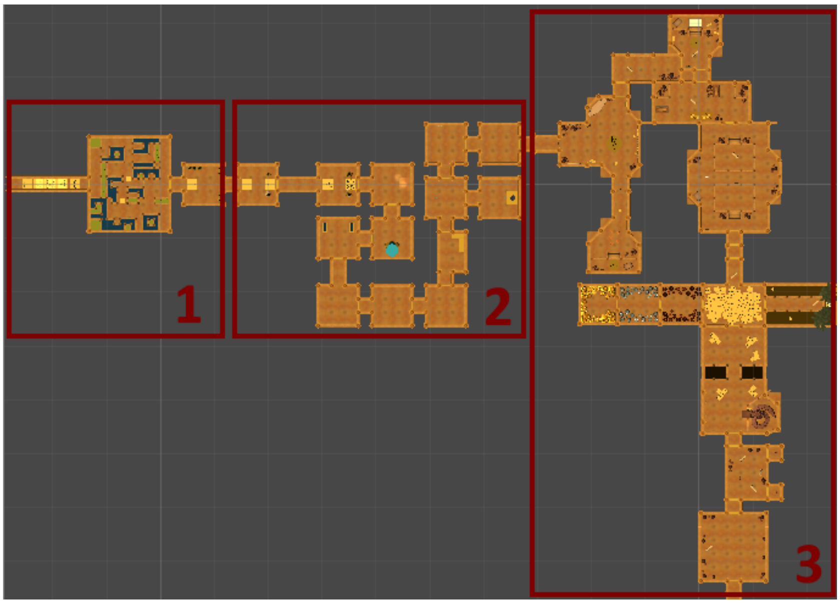


Figure 7.1.: Dynamic Evaluation Level Complete Map

continue to fail the riddle, but still remain appropriate to the current difficulty.

7.2. Concrete Implementation

In this section the concrete implementation of the CDDA will be explain by using two rooms, one simple (*StoolRoom*) and one complex room (*ShipRoom*), as an example. Since the difficulty is supposed to linearly increase within a respective difficulty bracket from start till the end a balancing of each individual room is necessary by taking the difficulty of each of the riddles within the room into account. Since the *StoolRoom* is one of the earlier rooms located in the single-literal section of the level (Figure 7.1 annot. 2) the overall difficulties of the room within the brackets should be lower than the ones in the *ShipRoom* because this one is located in the word section of the level (Figure 7.1 annot. 3). The rooms within the first section of the level (Figure 7.1 annot. 1) are the tutorial rooms. These are used to determine the correct gaming difficulty level, since no Hieroglyphs are introduced in these rooms. In the first of the rooms the player has to maneuver through a labyrinth, which immediately shows if a player is overwhelmed by the controls of the camera and the player movement, this way the gaming difficulty can be decreased according to the time it took the players to complete the room. An increase in gaming difficulty is not possible within these first rooms. In the second room the player has to pick up four vases and put them back into the shelves before being able to move on to the next room. In this room the gaming difficulty can either be easy or medium according to the player's performance in the labyrinth. The amount of required actions for the easy difficulty is not reduced by reducing the amount of vases but rather placing the vase immediately into shelf when the player clicks on it, rather than having to pick it up and than

Figure 7.2.: Dynamic *StoolRoom* Gaming Difficult Set to Hard

clicking on the correct spot on the shelf (medium difficulty). This way inexperienced players do not get overwhelmed at the start and are slowly introduced to the concept of picking up items and than placing them. From than on the single-literal section (Figure 7.1 annot. 2) starts, which is than followed by the word section (Figure 7.1 annot. 3) of the level.

StoolRoom: One of the rooms in the single-literal section is the *StoolRoom* (Figure 7.2). In this room the players are introduced to the Hieroglyph ▯ which represents the top of a stool or the reed mat on top of it. It is the fifth room of the level and therefore the overall difficulty level should not be to high. It is also the first time the players get in contact with the *WorldHighlightRiddle* (Figure 6.3), which is a very unique concept and can be overwhelming at first. Therefore the room only includes the *WorldHighlightRiddle* in the gaming domain and only the *TouchRiddles* on the *RiddleDoor* in the learning domain. The overall gaming domain difficulty of the room is balanced by the amount of single *WorldHighlightRiddles*: While in the easy gaming difficulty only two of the riddles are shown (Figure 7.2 red annot. 1 and 2), the medium gaming difficulty requires the player to solve three (Figure 7.2 red annot. 1-3) and the hard gaming difficulty requires the player to solve four (Figure 7.2 red annot. 1-4). To help the players getting used to the new riddle type blue markers on the floor show the correct position the players have to stand in the easy and medium difficulty. The overall learning difficulty of the room is balanced by the amount of *TouchRiddles* on the *RiddleDoor* (Figure 7.2 blue annot. 1). The door always contains the freshly learned ▯ to act as a progress control and additionally contains different amounts of dynamic Hieroglyphs: While the easy learning difficulty adds one dynamic Hieroglyph to the door, the medium learning difficulty adds two dynamic Hieroglyphs and the hard learning difficulty three. Since this room is the fifth room of the game the complexity is rather simple and straightforward for the players and also in terms of CDDA.

ShipRoom: One of the rooms in the word section is the *ShipRoom* (Figure 7.3). In this room the players get in contact with their first words, respectively the none "ship" and the verb "to see". Since this room is the first room of the word section the overall difficulty has increased when comparing the respective difficulty levels to those of the *StoolRoom*, for example. The room includes different gaming (Table 7.3) and learning riddles (Table 7.4).

The verb is taught to the players mostly through riddles from the gaming domain. More specifically through four generic tile based riddles (Figure 7.3 red annot. 1-3). The difficulty of each of these riddles is balanced through the actions required to solve it. The tiles necessary to solve the *GenericPickUpRiddle* and the *GenericSlideRiddle* are partially lying on the ground and are partially contained in *BreakableVases* ensuring a seamless integration. Neither the *GenericFlipRiddle* nor the *GenericSlideRiddle* in this room are affected by the riddle personalisation (Section 7.1.1), because this is the first room the players get in contact with both of these riddle types the room is only used to determine the performance of the players on these riddle types. The *GenericSlideRiddle* (Figure 7.3 red annot. 1) requires four sliding steps and two additional steps for picking up the missing piece and placing it into the missing spot in the easy difficulty. While the two additional steps do not change the sliding steps are increased to eight in the medium and twelve in the hard difficulty (Table 7.3 1). Both *GenericPickUpRiddles* (Figure 7.3 red annot. 2) combined require six steps for picking up and placing the missing pieces back into the riddle in the easy, eight in the medium and ten in the hard gaming difficulty (Table 7.3 2). This riddle is in general easier than the *GenericSlideRiddle* and *GenericFlipRiddle*. The *GenericFlipRiddle* (Figure 7.3 red annot. 3) requires two steps in the easy, three in the medium and four in the hard gaming difficulty setting (Table 7.3 3). While this amount of steps does not seem to be a large increase, the difficulty increase in the puzzle by introducing one additional step is immense. Since these are the only riddles regarding the gaming domain this adds up to 14 required actions in the easy, 21 in the medium and 28 in the hard gaming difficulty for this specific room (Table 7.3 Total).

The noun is taught the players trough a *GenericWordPutDownRiddle* (Section 6.3), which in this case does not indicate the placement of the four tiles (Figure 7.3 annot. 3) through four separate tile moulds but rather only shows a big white space where the players have to build the word (Figure 7.3 annot. 2). The correct order of the tiles is shown to the players when they have learned the verb "to see" and then activate the *TouchRiddle* on the platform (Figure 7.3 annot. 4). After performing these steps the players will see the word "ship" written in Hieroglyphs on the 3D ship model in the room, when standing on the activated platform. Now the players have to remember the correct order of the single Hieroglyphs used

Gaming Domain Riddles (G)	Easy	Medium	Hard
1: *GenericSlideRiddle*	6	10	14
2: *GenericPickUpRiddle*	6	8	10
3: *GenericFlipRiddle*	2	3	4
Total	14	21	28

Table 7.3.: Gaming Domain Riddles *ShipRoom*

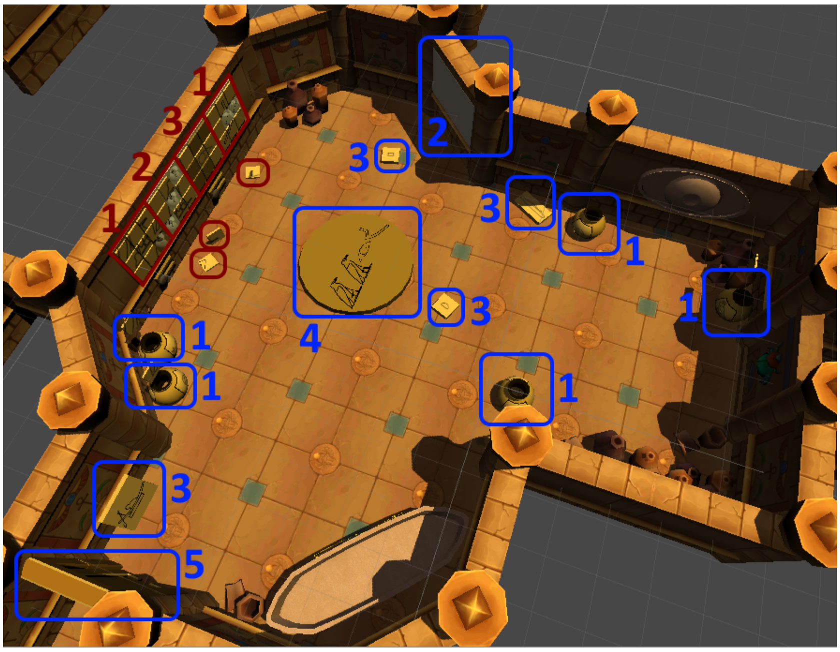


Figure 7.3.: Dynamic *ShipRoom* Gaming and Learning Difficult Hard

to build the word, pick them up and place them onto the white word building space one by one (Figure 7.3 annot. 2-3). Picking up the tiles requires the players to select the correct transliteration of the Hieroglyphs depicted on them in the *LiteralPicker*, which are in total three, one tile is depicting the ship itself and therefore no transliteration has to be selected for this tile, required actions in the learning domain for picking up the tiles (Table 7.4 3). Additionally four actions in the learning domain are required to place them correctly onto word writing space (Table 7.4 2). The difficulty of the riddle is in this case not balanced by the amount of required actions, since the player should build the complete word regarding the difficulty level. The riddle is balanced through additional help provided to the players via an annotation showing them the correct placement of the single Hieroglyphs on the word writing space. Therefore the required actions are the same on all learning difficulty levels for this riddle (Table 7.4 2-4). It is also not possible to dynamically change the Hieroglyphs on the tiles or the platform, as they are part of the word writing process. The remaining riddles regarding the learning domain are the *BreakeableVases* (Figure 7.3 annot. 1) and the *TouchRiddles* on the *RiddleDoor* (Figure 7.3 annot. 5), the difficulty of those is balanced through their amount of occurrence. In the easy learning difficulty two, in the medium three and in the hard five *BreakableVases* are placed inside the room (Table 7.4 1). To continue to the next room the players are required to solve two *RiddleDoors*, with a total of four *TouchRiddles* in the easy, six in the medium and eight in the hard learning difficulty setting (Table 7.4 5). This adds up to 14 riddles regarding the learning domain in the easy, 17 in the medium and 21 in the hard (Table 7.4 Total).

Both of the implementation examples highlight the additional balancing work required to

Learning Domain Riddles (L)	Easy	Medium	Hard
1: *BreakableVase*	2	3	5
2: *GenericWordPutDownRiddle*	4	4	4
3: *TileRiddle*	3	3	3
4: *TouchGlyph*	5	7	9
Total	14	17	21

Table 7.4.: Learning Domain Riddles *Ship Room*

generate a linear increasing difficulty for both domains for the three difficulty brackets. It is not possible to simply determine a difficulty level for each of the riddles and assume that this will ensure a balance in the difficulty between the rooms and also generate a linear increasing difficulty over the course of the level.

7.3. Training Mode - Long Term Learning Environment

The implementation of the CDDA in the evaluation level of *HieroQuest* paved the way for a dynamic training environment. The Training Mode that was introduced in the second iteration of the game (Section 6.3) only included linear paths, no possibility to change the difficulty and only supported a random sampling of already learned Hieroglyphs. Therefore the mode was not as dynamic as is could be and was completely reworked by introducing new graphics, a new path generation algorithm and the possibility to set desired starting difficulties for both domains and also the possibility to keep none, one or both difficulties static. By creating an appropriate path according to the players gaming difficulty and also using the Hieroglyph scores from the dynamic evaluation level a personalized long term learning environment was created. The process of map generation (Section 7.3.1) and the generation of the dynamic rooms (Section 7.3.2) will be explained in detail in this section.

7.3.1. Dynamic Map Generation

While in the dynamic evaluation level highly dynamic concepts like the personalized riddles or the dynamic *RiddleDoors* were introduced, it was not possible to make full use of the concept within the static structure of the level. The interior of the rooms changed dynamically towards the players needs but the structure of the level and the path the players have to follow remains the same. Replayability is only partially given, since player will eventually see the progress they made by possibly recognizing harder riddles than the last time they played the level, but will most likely also get bored by the static one-dimensional path they have to follow. Besides that, no rewards are given to the players when replaying the dynamic level. The training mode tries to closes this gap by generating a long term learning environment, which not only includes dynamic interior of the rooms, but also a dynamic path according to the players gaming difficulty.

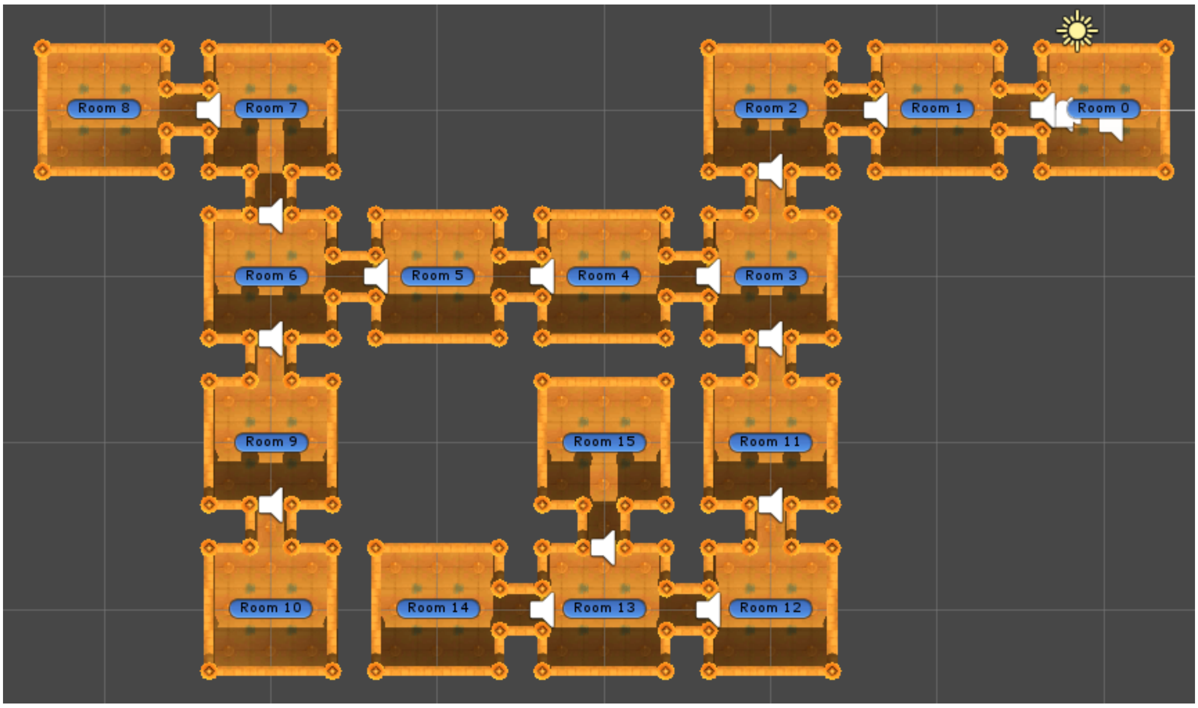


Figure 7.4.: Randomly Generated Training Path Hard Gaming Difficulty

The gaming difficulty of the players is either set to a intended starting difficulty or is known from the dynamic evaluation level or previous training sessions. When entering the *TrainingMode* a path with the following criteria will be generated: An easy gaming difficulty will only result in a linear path, a medium gaming difficulty will also introduce intersections at a certain probability, where the players eventually have to backtrack as they are not able to continue with the path, and a hard gaming difficulty will cause those intersections to occur at an even higher probability. The length of the path is specified before entering the training mode with a maximum length of the current amount of unlocked Hieroglyphs.

In the easy gaming difficulty, a path is generated by expanding on the previous room, in the first step from the starting room (Figure 7.4 Room 0), with one of three possible choices. The next room will either continue in a forward direction or make a left or right turn (Figure 7.4 Room 1, 2, 4, 5, 7, 9, 11 and 12). This way a linear path will be generated. In the medium and hard gaming difficulties it is also possible that at a certain probability intersection rooms will be included. These rooms either introduce a new room in forward and left direction, forward and right direction or left and right direction (Figure 7.4 Room 3, 6 and 13). In the course of these additional paths also dead ends are introduced, from which the players will have to track back (Figure 7.4 Room 8, 10 and 14).

After an intersection room was created the path will continue growing using breadth-first expansion. When the path was generated the order in which the rooms have to be visited by the players is determined using depth-first search, this makes sure that the players do not have to constantly backtrack and can follow a certain path until the end. It is of course possible to also use a combination of breadth-first search and depth-first search to order the rooms, to generate a even more complex path, which is for now not included.

7.3.2. Dynamic Room and Dynamic Door Generation

After the path was generated and the rooms are ordered as desired the interiors have to be placed inside the rooms. First the Hieroglyphs have to be determined by taking the players personal Hieroglyph database provided by the dynamic evaluation level or previous training sessions into account. The Hieroglyphs will be chosen in a similar way to the one described in Section 7.1. At first the Hieroglyphs of the lowest score bracket will be used, if these are more than the intended path length, they will be randomly sampled. If these are not enough they will all be used and additionally Hieroglyphs from the next bracket will be included, either all of them if they are still not enough or randomly sampled if more than needed. This process is repeated until the necessary amount of Hieroglyphs was drawn. The order of the Hieroglyphs within a certain score bracket are randomized but the overall order of the brackets is kept, to make sure that the players train on a variety of Hieroglyphs even when just playing a few rooms in each of the training sessions.

Since in this new *TrainingMode* not only single-literals but also words which require knowledge of the Hieroglyphs used to build them are included it has to be made sure that all the Hieroglyphs necessary to build a word are taught to the player in a room before the actual word is encountered. Of course the players already know and might still remember the Hieroglyphs and their transliteration from previous training sessions or the dynamic evaluation level, because only already unlocked Hieroglyphs can be trained on, but since the players start each training session with an empty *LiteralPicker* they have no chance to select the correct transliteration even though they might know it. Therefore the order of the Hieroglyph list may has to be changed by the game if certain Hieroglyphs require the knowledge of other Hieroglyphs. If this is the case than those Hieroglyphs are inserted at random positions before the Hieroglyphs they are needed for. If those Hieroglyphs again require previous Hieroglyphs this process has to be repeated. For example, if the word "ship" should be placed at room 0, this would not be possible. Since the word itself requires the Hieroglyphs 𓊪, 𓏏, 𓂧 and also the word "to see" is needed to solve the riddle in the ship room, which again requires the knowledge of the 𓄿 Hieroglyph, the first position the ship room can be placed is room 5. To ensure that the path is solvable for the player the systems has to check for these dependencies and therefore might has to change the order.

General Function In general, in the *TrainingMode* the same functions for calculating the difficulties in the two domains are used as in the dynamic evaluation level (Section 7.1). This is also compliant with the guidelines of a consistent CDDA (Chapter 5). The only difference is that the path which is generated before the player enters the *TrainingMode* (Section 7.3.1) is calculated with the current gaming difficulty or a defined starting difficulty, so the complexity of the path wont change throughout the training session even when the players gaming difficulty will increase or decrease, but of course the interiors of the rooms are still affected.

Figure 7.5.: *UshabtiRiddle*

Since the mode is intended to train the players knowledge, in all rooms two *BreakableVases* are placed in the hard, one in the medium and none in the easy learning difficulty setting in addition to the four (hard), three (medium) or two (easy) dynamic *TouchRiddles* on the *RiddleDoor*.

Decoration To increase the immersion and overall user experience all of the rooms were decorated with the same Egyptian artefacts that are used thought the game. These decorations create the illusion that the rooms are not dynamically generated. When first testing the functionality of the algorithm it became clear that one could clearly tell that the rooms are not designed by a game designer but are rather put together by a computer. To counteract these feelings different and suitable decorations were placed inside the rooms.

7.3.3. New Dynamic Riddle Types

The dynamic generation of the training environment required additional riddles types. One of these riddles is used to guide the players in case of a dead end room, the so called *UshabtiRiddle*. Another one is used to show the players that they have reached the end of the training session and at the same time also reward them with collectibles, the so called *SarcophagusRiddle*.

UshabtiRiddle Since the intersection rooms from the medium and hard gaming difficulty setting introduced dead ends (Figure 7.4 Room 8, 10 and 14) to the paths, a new riddle type had to be conceptualized to indicate to the players that there are still riddles yet to be solved. For this purpose a riddle that in general behaves like a *RiddleDoor*, but first requires the player to insert a missing item, was implemented (Figure 7.5 annot. 1). The riddle consists of six *Ushabtis*, those are in general small statues, 10 to 20cm, in the form of a mummy, but also life size exemplars exist. One of the *Ushabtis* is missing and has to be found by the players. The missing one is always placed inside a *BreakableVase* inside the respective dead end. In this special case the players learn the ▯ Hieroglyph in the dead end, therefore the *Ushabti* is

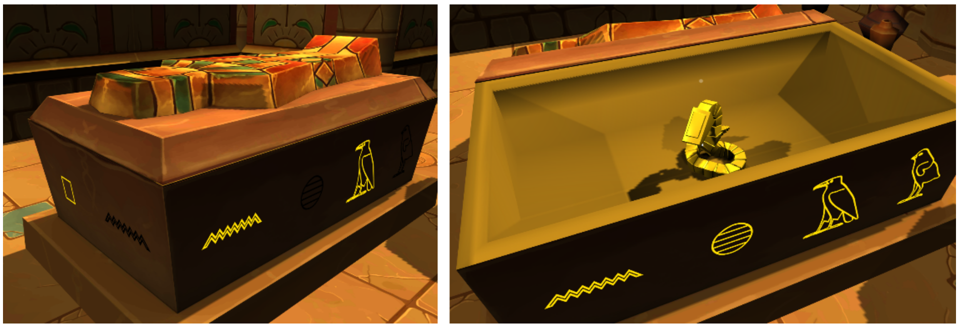


Figure 7.6.: *SarcophagusRiddle* with Special Reward

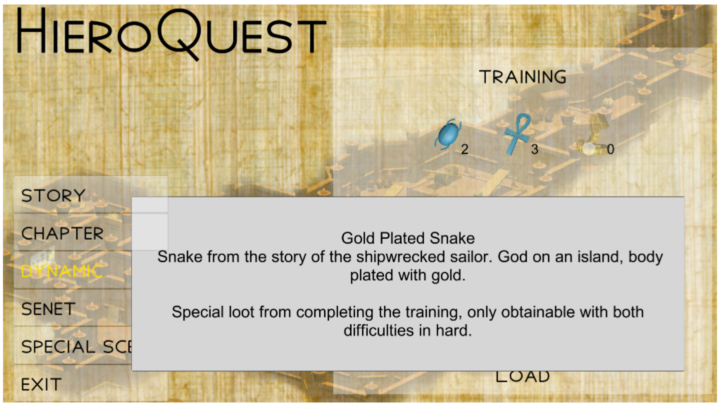


Figure 7.7.: Collected Rewards in the Main Menu

contained in a *BreakableVase* depicting the ▯ Hieroglyph, which makes sure that the players have solved all riddles in the room before being able to continue (Figure 7.5 annot. 2). When the vase is broken the *Ushabti* can be picked up by the players (Figure 7.5 annot. 3). Then they have to carry it to the intersection room and place it back into the shrine of different *Ushabtis*. In doing so Hieroglyphs on the larger *Ushabtis* will be exposed, which the players now have to use the correct transliteration on, like a *RiddleDoor* (Figure 7.5 annot. 4)). A correctly solved one is indicated by a yellow glow, in this special case three of four riddles are already solved. Since the Hieroglyphs are only shown to the player once the *Ushabti* was inserted the Hieroglyphs on the door can be dynamically adapted in content as well as amount just like a standard dynamic *RiddleDoor*. When all *TouchRiddles* on the *Ushabtis* are solved they will vanish into the ground and the player can continue to the next room. By creating a riddle in which the players have to carry a certain item from the dead end to the intersection room it is made sure that the player will notice that there is still a path left to explore. Since the *TouchRiddles* are dynamically determined after the player has finished the dead end room it is not possible to show any Hieroglyphs on the door beforehand. Therefore the *UshabitRiddle* also ensures that in the intersection room no empty door without any Hieroglyphs on it is shown to the players. Instead the players recognize the riddle they can't solve yet.

SarcophagusRiddle and Rewards At the end of the path the final challenge is awaiting the player. A Sarcophagus (Figure 7.6 annot. 1) was placed in the middle of the room containing a dynamic amount of *TouchRiddles*. The upper bound of the amount is capped by the amount of Hieroglyphs taught in the training session, but in general includes four Hieroglyphs in the easy, eight Hieroglyphs in the medium and twelve Hieroglyphs in the hard learning difficulty setting. No additional riddles are present in the room. The players need to select the transliteration of each of the Hieroglyph depicted on the Sarcophagus and then click on the respective Hieroglyphs. A solved riddle is indicated by the yellow glow of the Hieroglyph. If the player fails to select the correct transliteration of one of the Hieroglyphs the progress on this side of the Sarcophagus will reset and has to be repeated. This ensures a challenge for the players, while not creating to much tension reseting the progress on the complete Sarcophagus. If all of the *TouchRiddles* are solved the Sarcophagus will open and the player will be rewarded with a collectable item. The training session will end and the players are brought back to the *MainMenu* of the game. Here they can also view their collectables (Figure 7.7). Three different rewards, a Scarab, a Ankh symbol and a Golden Snake, can be collected by the players, additional information is provided to each of them when hovering over the icon (Figure 7.7 Box). The different rewards can only be obtained when certain thresholds are met, for example while the Scarab is the basic reward and can therefore be obtained at any of the difficulty levels in both domains the Golden Snake can only be obtained when finishing the training with both difficulty settings set to hard. This type of reward systems further encourages replayability by motivating the player to collect all of the pieces.

8. User Study

Due to the pandemic it was not possible to conduct a user study, therefore in this chapter the theoretic methodology and procedure of the user study will be described without the results. To fully understand the impact of the two domains on the learning success four variants of the CDDA have to be tested against each other: a) dynamic learning difficulty and dynamic gaming difficulty, b) dynamic learning difficulty and static gaming difficulty, c) static learning difficulty and dynamic gaming difficulty and d) static learning difficulty and static gaming difficulty. Only verifying the positive impact of the completely dynamic difficulty compared to the completely static difficulty would not preclude that the two partially dynamic difficulties could still outperform the completely dynamic difficulty. Therefore to verify the positive impact of the CDDA on learning success and player experience all four variants have to be investigated.

In order to evaluate these effects several hypotheses were formulated (Section 8.1). Those were than investigated in two different user studies. The first user study (Section 8.2) focuses on the short-term effects of the CDDA while the second user study (Section 8.3) focuses on the long-term effects.

8.1. Hypotheses

In the design process of the CSGB several assumptions were made, mainly the equal contribution of both domains to the overall player experience (Chapter 3). The CDDA implementation in *HieroQuest* is based on these assumptions (Chapter 7). To investigate the impact of the CDDA implementation on the learning success and the overall player experience the following hypotheses were formulated:

H1: Dynamic Learning Difficulty and Dynamic Gaming Difficulty will result in a higher learning success rates than the completely or partially Static Difficulty approaches. To verify the positive impact of the CDDA on the learning success in *HieroQuest* dynamic learning and dynamic gaming difficulty has to perform significantly better than the partially or fully static approaches in terms of learning success.

H2: Dynamic Learning Difficulty and Static Gaming Difficulty will result in no significantly different learning success rates than Static Learning Difficulty and Dynamic Gaming Difficulty. To verify the equal contribution of both domains to the overall learning success and player experience dynamic learning and static gaming difficulty should not lead to significantly different results than static learning and dynamic gaming difficulty.

H3: Static Learning Difficulty and Static Gaming Difficulty will result in significantly lower learning success rates than the completely or partially Dynamic Difficulty approaches. To verify the positive impact of the CDDA on the learning success even when only one domain is dynamic the completely static difficulty has to perform significantly worse than partially or completely dynamic difficulty.

H4: Static Learning Difficulty and Static Gaming Difficulty will result in significantly higher correlations between previous experience with digital games and the categories of the GEQ Core Module compared to completely or partially Dynamic Difficulty. In the user study of the third iteration of *HieroQuest* (Section 6.4.2) significant differences between the two groups in several correlations became evident. Mainly the previous experience with digital games (play-time of digital games per week) and several GEQ Core Module [47] categories showed these differences between the two groups. It is assumed that these differences in correlations can also be shown when comparing completely static difficulty to partially or completely dynamic difficulty, because they are assumed to be caused by an imbalance in perceived challenges and skills.

H5: Dynamic Learning Difficulty and Dynamic Gaming Difficulty will result in significantly lower correlations between previous experience with digital games and the categories of the GEQ Core Module compared to completely or partially Static Difficulty. Following the reasoning for hypothesis 4 (H4) the completely dynamic difficulty should significantly reduce the correlations between previous experience with digital games and the GEQ categories, because the imbalance between perceived challenges and skills is minimized by dynamic difficulty adjustment.

8.2. User Study I

The goal of User Study I is to investigate the short-term effects of the CDDA on the learning success and the player experience in *HieroQuest*.

Participants: For User Study I 60 participants without any previous knowledge of the Middle Egyptian language were invited. Since the participants have to be separated into four groups to play the four different versions, 60 was identified as a minimum number of participants. They did not have any previous knowledge in Middle Egyptian, which was verified by a yes-no question. In the user study of the third iteration of *HieroQuest* (Section 6.4.2) participants without previous knowledge and participants with previous knowledge of Middle Egyptian were invited. This showed several interesting implications and is overall compliant with the goal of the CDDA to make Serious Games effectively usable by players with any level of previous experience in both domains. In this case it was decided to only invite participants without previous knowledge to simplify the procedure of this first user study on CDDA. Since the participants have no previous knowledge of the language, no pre-test is required. Additionally, without previous knowledge the results of the post-test directly show the learning success and do not have to be compared to previous results. In a later stage also participants from various backgrounds have to be included (Chapter 9), but at this early stage of investigation a unified group of participants was identified to be sufficient.

Questionnaire: To measure the learning success and the player experience after the gaming session, a post-game questionnaire was designed. This questionnaire contains the Game Experience Questionnaire (GEQ) Core Module [47], multiple-choice questions regarding the serious topic and several feedback questions (Appendix B): The GEQ Core Module consists of 33 questions regarding the players feelings while playing the game. These are answered on a five-point Likert scale, ranging from 0 <not-at-all> to 4 <extremely>; The 29 questions regarding the serious topic consist of 27 questions that query the 27 Hieroglyphs taught in the special evaluation level (Section 8.2 Game Level) and 2 questions that query knowledge about the *Story of the Shipwrecked Sailor* (6.1.2). Since the plot of the story is an immanent learning goal of *HieroQuest* it also needs to be queried in the questionnaire; The rest of questions consist of open-ended feedback and demographic questions. Especially the previous experience with digital games (play-time per week) and the elements the players especially liked/disliked are of interest.

Game Level: Because the average play-time of the complete game is not feasible in one session, a special evaluation level was created (Section 6.4.2). Later on CDDA was implemented into the level (Section 7.1). This level will now be used to evaluate the impact of the CDDA on the learning success and the player experience. In order to do so, four different versions of the level, which only vary in the use of dynamic and static difficulties were created. The first version uses a static medium learning and a static medium gaming difficulty (SLSG). The second version uses a static medium learning and a dynamic gaming difficulty (SLDG). The third version uses a dynamic learning difficulty and a static medium gaming difficulty (DLSG). The fourth version uses a dynamic learning and a dynamic gaming difficulty (DLDG). Since the different variants of the evaluation level only differ in this single aspect it is possible to directly compare the results of the different groups. The evaluation level first introduces the player's to a selection of single-literals and then move on to multi-literals, words and the plot of the *Story of the Shipwrecked Sailor*. The evaluation level has an average play-time of approx. 90 minutes.

Procedure: The 60 participants were invited to play the game in a laboratory environment in the university. The user study was conducted on a single day with four different time slots. Beforehand the participants were randomly separated into four groups, which leaves each group with 15 participants. Each group is assigned one version of the game level (Section 8.2 Game Level). The participants of Group A played the game with the SLSG settings, the participants of Group B played the game with the SLDG settings, the participants of Group C played the game with the DLSG settings and the participants of group D played the game with the DLDG settings. Every participant only plays one version of the game. At each time slot the 15 participants were seated in front of one hand-held tablet device each. These devices were already prepared with the game set up in the correct version for the group. After some procedural instructions the participants started playing the game. The sounds of the game were turned off to minimize distraction. The participants were not allowed to talk during the whole study. Instructors were present to ensure these conditions, but only interfered in case of technical difficulties. Instructors did not provide any advice regarding the gameplay.

After a participant had finished the game the questionnaire (Section 8.2 Questionnaire) is shown on the tablet device. After 90 minutes of play-time the questionnaire is shown to all players regardless of their progress in the game. Since it is an requirement on the evaluation level to be completable in the time frame of 90 minutes, the participants are still queried all Hieroglyphs taught in the level in case of non-completion even when they did not encounter them. After the participants had completed the questionnaire they were free to leave.

Evaluation: The results of the questionnaires are automatically inserted into a database after completion. Since the questionnaires are directly answered on the table device it is possible to also attach technical data to the results. For example the learning and gaming difficulties of a participants over the course of the game and the time spend in each interval are also included. This allows for a fast evaluation of the data, since no additional matching task has to be performed. The data is then analyzed according to the five hypotheses (Section 8.1). The average learning success of each group can directly be compared to the other groups, because none of the participants had previous knowledge in the Middle Egyptian Language. The correlations between the participants' previous experience with digital games and the GEQ Core Module can also be directly compared. This is possible even when groups have different average previous experiences with digital games, because only the correlations are compared.

8.3. User Study II

The goal of User Study II is to investigate the long-term effects of the CDDA on the learning success and the player experience in *HieroQuest*.

Participants: For User Study II the 60 participants of User Study I (Section 8.2 Participants) were invited to three additional gaming session. Since the participants have to be familiar with the evaluation level of the game before they are able to train their knowledge in the *TrainingMode*, they are required to play the evaluation level first. User Study II is therefore an extension of User Study I. The participants remain in the four groups throughout the whole study.

Questionnaire: User Study II makes use of the same questionnaire as User Study I (Section 8.2 Questionnaire) with a randomized question order on each session. An example of a randomized question and answer order can be seen by taking a look at the pre-test of the user study of the third iteration of *HieroQuest* (Appendix A). By utilizing the same questionnaire it is possible to compare the results from the first gaming session (User Study I) to the results of the additional three sessions (User Study II). Since the questionnaire queries every Hieroglyph taught in the game it is not possible to specifically learn for the questionnaire by remembering the questions. Since the game level used for the additional three sessions does not include the plot of the story (Section 8.3 Game Level) these two questions are not considered in the evaluation step of User Study II (Section 8.3 Evaluation).

Game Level: The *TrainingMode* (Section 7.3) was used as the game level of User Study II. While in User Study I the participants still had to learn the Hieroglyphs they are now able to train their knowledge. The *TrainingMode* is also set up to support the four different versions to ensure comparability: static medium learning and a static medium gaming difficulty (SLSG), static medium learning and a dynamic gaming difficulty (SLDG), dynamic learning difficulty and a static medium gaming difficulty (DLSG) and dynamic learning and a dynamic gaming difficulty (DLDG).

The hieroglyphs taught in the training mode are dependent on the players weaknesses and strengths, similar to the choice of dynamically adjustable Hieroglyphs in the evaluation level (Section 7.3). To ensure that the game can use the personal Hieroglyph database of a participant (Section 7.1.2), the tablet devices are numbered. The participants are required to write down the number of the tablet device they used in User Study I. To further ensure that the devices do not get mixed up each participant is presented a random number, which is used to log into the *TrainingMode*. This authentication process is necessary since the possible mix up of personal Hieroglyph databases would compromise the results of the study. In the *TrainingMode* used for this study the players play 15 rooms, with one Hieroglyph per room, in one session which results in an average playtime of 30 minutes per session.

Procedure: The first gaming session was already conducted in User Study I (Section 8.2 Procedure). In this session the 60 participants were separated into four groups and played the corresponding version of the evaluation level of the game: SLSG, SLDG, DLSG, DLDG. In the three following sessions the participants play the *TrainingMode*, which is also set up in the four different versions. The participants remain in their group for the course of the user study and therefore play in the same version as before.

The gaming sessions are on a three day basis, which means two free days between each session. Similar to the procedure of User Study I one group of participants is invited at a time. The 15 participants of a group are then seated in front of their personal tablet, ensured by a two step authentication process (Section 8.3 Game Level). They then play the *TrainingMode* until they are finished with the 15 rooms or the time frame of 30 minutes is over. The personalized hieroglyph Database of a participant carries over from the different sessions (Section 7.1.2). In the case of dynamic difficulty in one or both domains the dynamic difficulty level also carries over. In the case of static difficulty the difficulty levels remains at medium over the whole course of the study. After the participants are finished with the gaming session they are presented with the questionnaire (Section 8.3 Questionnaire). Then the participants are free to leave. In total a participant is invited to four gaming sessions, playing the evaluation level in the first session and the *TrainingMode* in the remaining three. After the fourth session the user study is finished.

Evaluation: The results of all four questionnaires of a participant are automatically inserted into the database after completion. Additionally also the technical data of each gaming session is included, e.g. learning and gaming difficulties over the course of each session and the time it took. The data of a group is put into perspective to the changes over the four sessions and to the data of the other groups, according to the hypotheses (Section 8.1). The changes of the average learning success of each group is analyzed over the four sessions. The learning success can again directly be taken from the results of the questionnaires, since the only learning input was the game itself. Also the changes of the correlations between the participants' previous experience with digital games and the GEQ Core Module are analyzed over the four sessions.

9. Limitations and Future Work

In this chapter limitations of the proposed work and possible future directions are presented. The chapter is structured as follows. First the limitations and future directions of the CSGB model are shown, followed by the CDDA. Then the limitations and future directions of *HieroQuest*, the CDDA implementation in *HieroQuest* and the user study are summarized.

The CSGB model is a purely theoretical model. The general idea is based on the findings from Plecher et al. [74], in which players with a lot of previous experience were annoyed by a version of the game with a low learning difficulty, while the players with a lot of previous experience of a version with a higher learning difficulty were not. Plecher et al. [74] reasoned that the learning difficulty counteracted the imbalance in the gaming domain. One could argue that basing a model on this assumption is not enough to fully support it, but the model is not solely based on the assumption and is rather inspired by it. Several other researchers already identified the need for a separation of learning and gaming domain (e.g. [40, 41]). One could also argue that the model does not provide any kind of measurements to verify the current state of the player, but the model is per definition only a theoretical one and on its own not intended to verify the player experience. It should only be used to reason about Serious Games design and to base further implications on it.

Another aspect is the CDDA. One could argue that the CDDA does not provide detailed guidelines for the separation of the elements into either the learning or gaming domain. The guidelines are rather vague in this point as it is highly dependent on the game at hand and therefore the guidelines have to be generalized. The two highly different example games and the implementation in *HieroQuest* try to counteract this limitation, but it could be argued that more examples should be provided in the future.

The Serious Game *HieroQuest* has several limitations itself. The learning content of the game needs to be further improved to function as a tool for Egyptology students or intermediate learners of the Middle Egyptian language. Possible future directions are the integration of grammar and the writing of Hieroglyphs, this can also be enhanced by integrating new types of riddles. For example, a 2D riddle that requires the players to write Hieroglyphs or rearrange single Hieroglyphs to words or even whole sentences can be implemented. Also the integration of Augmented Reality can be used to recognize the players' handwriting of Hieroglyphs and overlay the correct writing. This can also be connected to the player's personal Hieroglyph database (Section 7.1.2). Another aspect is the replayability of *HieroQuest*. While the *TrainingMode* encouraged the replayability by a dynamically created the game world, additional motivations in terms of rewards and collectables should also be implemented in the standard game modes.

Also limitations of the implementation of the CDDA in *HieroQuest* can be shown. The main limitation of the implementation is the reliable separation of the elements into their respective domains. Since the game is a 3D game all of the learning tasks still require some form of interaction with the game world, e.g. manoeuvring or clicking on colliders. Even though the elements were carefully assigned one could still argue that the domains are not fully separated, which is something that can never be fully avoided.

The user study also shows a few limitations. the main limitation of the user study is that only the theoretical process was described. Due to the pandemic it was not possible to conduct a study without putting the participants at risk. Another limitation of the user study is the background of the participants. While the CDDA is intended to enable a Serious Game to be applicable to a broad group of target audience, a large group of possible users was not considered, since none of them had previous experience with Hieroglyphs. At this stage of research the primary goal was to show that the CDDA is capable of improving the learning outcome and player experience for this group of participants, before investigating the impact of the CDDA for other groups. Furthermore one could argue that the number of participants is relatively low with 15 per group, which has to be increased in future studies.

Additionally the CDDA needs to be applied to different games and there the impact needs to be investigated, as well. It is also necessary to compare a Serious Game with CDDA to traditional learning approaches to investigate the short-term and long-term learning success and motivation.

10. Conclusion

In this chapter the presented work is summarized and at the same time a detailed conclusion is given. Serious Games have the possibility to enhance learning processes all over the world, especially in areas where teaching is not available. But before they can become applicable to a heterogeneous group of target audience, they have to become effectively be usable for them. Not only for players that are in general interested in games, but also for learners that are interested in the serious topic. In this thesis two main approaches were conceptualized that try to enhance the Serious Game design process and Serious Game research for them to become dynamic, personalized, attractive and effective learning environments, by shedding light on the differentiation between the learning and the gaming domain. Serious Games try to achieve to contrary goals at the same time. They try to maximize the fun (attractiveness) and the learning outcome (effectiveness), which already hints towards the necessity of a differentiated consideration of both domains.

To achieve this overall goal of creating models for personalized learning environments, smaller sub goals were defined in Section 1.3. The first of these sub goals was the creation of a theoretical model for Serious Game balancing, that takes the differences between the learning and the gaming domain into account: *Componentwise Serious Game Balance (CSGB)* (Chapter 3). Before the model was conceptualized the theoretical basis in terms of Serious Game definitions, learning theories, learning environments, motivation and flow had to be investigated (Chapter 2). With this information being available it was possible to create the CSGB model. The model encapsulates the difference between the two domains of a Serious Game and the balance between the perceived challenges and skills of an individual player. It is not intended to be used as a criteria or measurements for good Serious Game design, but should rather be used as a basis for reflection on balancing in Serious Games and future research. The model on the one hand simplifies the concept of player experience by only considering the basic principle of an enjoyable experience, the balance between challenge and skill, but on the other hand also complicates the concept by introducing the learning domain as an additional factor. Overall the model creates a new viewpoint on Serious Game research and design.

The second sub goal was the conceptualization of a dynamic difficulty adjustment for Serious Games, which is based on the CSGB model: *Componentwise Dynamic Difficulty Adjustment (CDDA)* (Chapter 5). Taking a look at already existing solutions for balancing in entertainment games and Serious Games the basis for the model was built (Chapter 4). The CDDA fills the gap between the theoretical CSGB model and the practical implementation. With CDDA it is possible to enhance already existing Serious Games or create new ones that take the skill levels of a player into account and are therefore tailored towards an individual. The process of building such a Serious Game is described by a five step plan (Section 5.2): 1)

Identify dynamically adjustable elements; 2) Find measurements; 3) Assign difficulty levels; 4) Identify intervals; 5) Assign an update function. Additionally, a Serious Game can be further tailored towards an individual even further by personalizing the gaming and learning content (Section 5.3).

The third sub goal was the implementation of the CDDA in the already existing Serious Game *HieroQuest* (Chapter 5). The complete design process of *HieroQuest* itself, with the three refinement circles the game went through, were analyzed first, to get an insight on the game (Chapter 6). To implement the CDDA the five step plan (Section 5.2) was used. Implementing the model in *HieroQuest* not only showed that it is possible to use the plan to create a dynamic and personalized learning environment out of an already existing static Serious Game, but also allowed to test the impact of the CDDA on the learning success and the player experience in an user study (Chapter 8).

The fourth sub goal was to utilize the dynamic property of the CDDA to the fullest and create a new game mode within *HieroQuest* to increase replayability and also give players the possibility to train their knowledge (Section 7.3). In this new *TrainingMode* they game world is build based on the player's skill levels, which are measured beforehand in the standard game mode or in real-time while training their knowledge. Furthermore it was possible to effectively train the player's knowledge by building a personalized Hieroglyph database, to tackle flaws in the player's knowledge in the learning domain (Section 5.3). Additionally, the creation of this new mode made long-term investigations on the impact of CDDA on the learning success and player experience possible.

The fifth sub goal was the conduction of a user study to evaluate the short-term and long-term impacts of the CDDA (Chapter 8). Sadly, due to the pandemic at the time of the creation of this thesis, it was not possible to conduct a user study without putting the participants at risk. Therefore only the theoretical methodology of the user study was presented. To account for both short- and long-term effect two user studies were described. While in the first one the participants played the special evaluation level of the game with CDDA, in the second user study the participants additionally played the *TrainingMode* of the game on three occasions. In order to fully investigate the impact of the CDDA the participants had to be separated into four groups for both studies. Each group played a different version of the game: A) Dynamic learning difficulty and Dynamic gaming difficulty; B) Dynamic learning difficulty and Static gaming difficulty; C) Static learning difficulty and Dynamic gaming difficulty; D) Static learning difficulty and Static gaming difficulty; Only by considering all four versions it is possible to identify whether the CDDA (Dynamic Learning and Dynamic Gaming Difficulty) outperforms the other versions in terms of learning success and player experience.

Limitations of the proposed work and possible future directions for the different aspects of the work were presented in (Chapter 9). The main drawback of the presented work is the still yet to be confirmed assumption that the CDDA has a significantly positive impact on the learning success and player experience of a Serious Game. The user study would have built a first step towards this assumption by confirming it for the Serious Game *HieroQuest*, which could sadly not be conducted. To investigate if this assumption holds and to also fortify the equal contribution of learning and gaming domain to the overall player experience and learning success within a Serious Game assumed in the CSGB model, CDDA needs to be implemented in different Serious Games. Only then it is possible to confirm theses positive impacts.

This thesis build the basis for a new viewpoint on Serious Game design and Serious Game research by shedding light on a differentiated consideration of learning and gaming domain. It paved the way for future research on this direction of Serious Game balancing, which will hopefully in the future increase the popularity and effectiveness of Serious Games, so they can become powerful and personalized tools which enhance learning processes across the world.

A. Pre-Game Questionnaire

HieroQuest Questionnaire

Pre-Game Questionnaire

1. Select the correct Transliteration corresponding to the Hieroglyph shown:

t | *d* | *b* | NONE

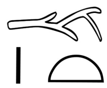
2. Select the correct Transliteration corresponding to the Hieroglyph shown:

p | *g* | *d* | NONE

3. Select the correct Transliteration corresponding to the Hieroglyph shown:

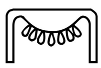
st | *m* | *p* | NONE

4. Select the correct Transliteration corresponding to the Hieroglyph shown:



jni | *rkḥ* | *ẖ.t* | NONE

5. Select the correct Transliteration corresponding to the Hieroglyph shown:



km | *nbw* | *ḫt* | NONE

6. Select the correct Transliteration corresponding to the Hieroglyph shown:

š | *w* | *f* | NONE

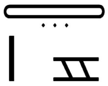
7. Select the correct Transliteration corresponding to the Hieroglyph shown:

ḏꜥ	*ḏd*	*sḏm*	NONE

8. Select the correct Transliteration corresponding to the Hieroglyph shown:

t	*k*	*b*	NONE

9. Select the correct English translation corresponding to the Hieroglyph shown:



SHIP SEE ISLE NONE

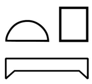
10. Select the correct English translation corresponding to the Hieroglyph shown:

SNAKE VIPER BASIN NONE

11. Select the correct English translation corresponding to the Hieroglyph shown:

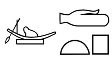
WOOD MOUTH BASIN NONE

12. Select the correct English translation corresponding to the Hieroglyph shown:



SHIP BREAD STRETCH NONE

13. Select the correct English translation corresponding to the Hieroglyph shown:



SHIP SKY WATER NONE

14. Select the correct English translation corresponding to the Hieroglyph shown:

QUAILCHICK VULTURE OWL NONE

15. Select the correct English translation corresponding to the Hieroglyph shown:

OWL VULTURE QUAILCHICK NONE

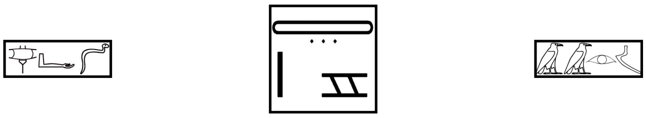
16. Select the correct Hieroglyph corresponding to the Transliteration shown:

ḥ

NONE

17. Select the correct Hieroglyph corresponding to the Transliteration shown:

nšn.j



NONE

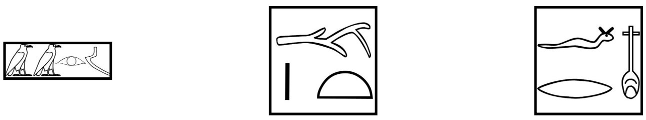
18. Select the correct Hieroglyph corresponding to the Transliteration shown:

rḥ

NONE

19. Select the correct Hieroglyph corresponding to the Transliteration shown:

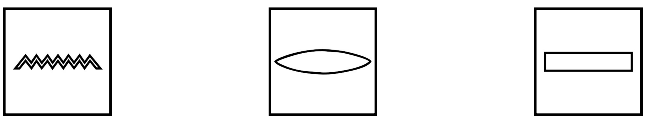
mꜣꜣ



NONE

20. Select the correct Hieroglyph corresponding to the Transliteration shown:

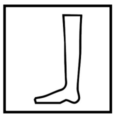
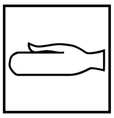
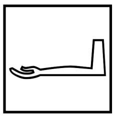
š



NONE

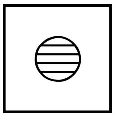
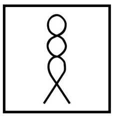
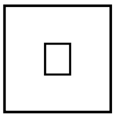
21. Select the correct Hieroglyph corresponding to the Transliteration shown:

ʿ

   NONE

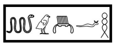
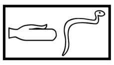
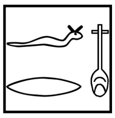
22. Select the correct Hieroglyph corresponding to the Transliteration shown:

ḥ

   NONE

23. Select the correct Hieroglyph corresponding to the Transliteration shown:

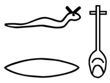
f

   NONE

24. Select the correct English translation corresponding to the Hieroglyph shown:

100 10 1000 NONE

25. Select the correct English translation corresponding to the Hieroglyph shown:



SAY GOOD KNOW NONE

26. Select the correct English translation corresponding to the Hieroglyph shown:

HEAR SEE STRETCH NONE

27. Select the correct English translation corresponding to the Hieroglyph shown:

WATER WELL BASIN NONE

B. Post-Game Questionnaire

HieroQuest Evaluation

Post-Game Questionnaire

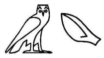
1. Describe your game experience from <not at all - 0> to <extremly - 4>:

	0	1	2	3	4
I felt content					
I felt skilful					
I was interesed in the game's story					
I thought it was fun					
I was fully occupied with the game					
I felt happy					
It gave me a bad mood					
I thought about other things					
I found it tiresome					
I felt competent					
I thought it was hard					
It was aesthetically pleasing					
I forgot everything around me					
I felt good					
I was good at it					
I felt bored					
I felt successful					

2. Describe your game experience from <not at all - 0> to <extremly - 4>:

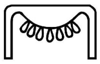
	0	1	2	3	4
I felt imaginative	○	○	○	○	○
I felt that I could explore things	○	○	○	○	○
I enjoyed it	○	○	○	○	○
I was fast at reaching the game's targets	○	○	○	○	○
I felt annoyed	○	○	○	○	○
I felt pressured	○	○	○	○	○
I felt irritable	○	○	○	○	○
I lost track of time	○	○	○	○	○
I felt challenged	○	○	○	○	○
I found it impressive	○	○	○	○	○
I was deeply concentrated in the game	○	○	○	○	○
I felt frustrated	○	○	○	○	○
It felt like a rich experience	○	○	○	○	○
I lost connection with the outside world	○	○	○	○	○
I felt time pressure	○	○	○	○	○
I had to put a lot of effort into it	○	○	○	○	○

3. Select the correct Transliteration corresponding to the Hieroglyph shown:



ḏd	sḏm	ḏꜥ	NONE

4. Select the correct Transliteration corresponding to the Hieroglyph shown:

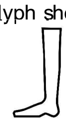


nbw	km	ḫt	NONE

5. Select the correct Transliteration corresponding to the Hieroglyph shown:

jni	rkḥ	ẖ.t	NONE

6. Select the correct Transliteration corresponding to the Hieroglyph shown:



t	d	b	NONE

7. Select the correct Transliteration corresponding to the Hieroglyph shown:

g	p	d	NONE

8. Select the correct Transliteration corresponding to the Hieroglyph shown:

t	b	k	NONE

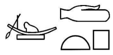
9. Select the correct Transliteration corresponding to the Hieroglyph shown:

m	p	st	NONE

10. Select the correct Transliteration corresponding to the Hieroglyph shown:

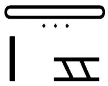
w	*š*	*f*	NONE

11. Select the correct English translation corresponding to the Hieroglyph shown:



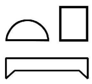
SHIP SKY WATER NONE

12. Select the correct English translation corresponding to the Hieroglyph shown:



ISLE SHIP SEE NONE

13. Select the correct English translation corresponding to the Hieroglyph shown:



SHIP BREAD STRETCH NONE

14. Select the correct English translation corresponding to the Hieroglyph shown:

VIPER SNAKE BASIN NONE

15. Select the correct English translation corresponding to the Hieroglyph shown:

VULTURE OWL QUAILCHICK NONE

16. Select the correct English translation corresponding to the Hieroglyph shown:

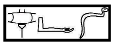
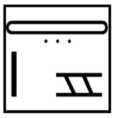
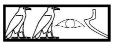
VULTURE OWL QUAILCHICK NONE

17. Select the correct English translation corresponding to the Hieroglyph shown:

BASIN MOUTH WOOD NONE

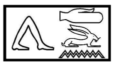
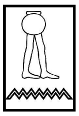
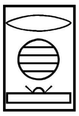
18. Select the correct Hieroglyph corresponding to the Transliteration shown:

nšn.j

   NONE

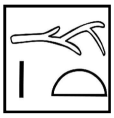
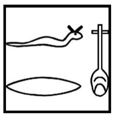
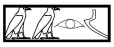
19. Select the correct Hieroglyph corresponding to the Transliteration shown:

rḫ

   NONE

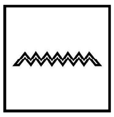
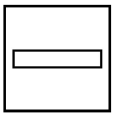
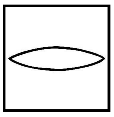
20. Select the correct Hieroglyph corresponding to the Transliteration shown:

mꜣꜣ

   NONE

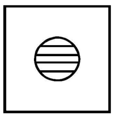
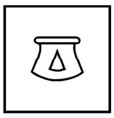
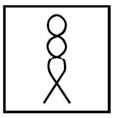
21. Select the correct Hieroglyph corresponding to the Transliteration shown:

š

   NONE

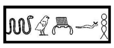
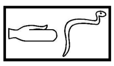
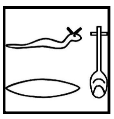
22. Select the correct Hieroglyph corresponding to the Transliteration shown:

ḫ

   NONE

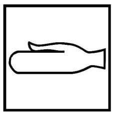
23. Select the correct Hieroglyph corresponding to the Transliteration shown:

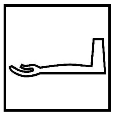
f

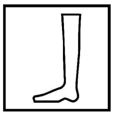
   NONE

24. Select the correct Hieroglyph corresponding to the Transliteration shown:

ʿ



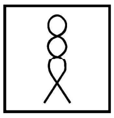


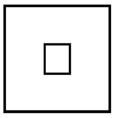


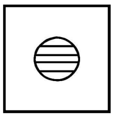
NONE

25. Select the correct Hieroglyph corresponding to the Transliteration shown:

ḥ







NONE

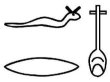
26. Select the correct English translation corresponding to the Hieroglyph shown:

SEE HEAR STRETCH NONE

27. Select the correct English translation corresponding to the Hieroglyph shown:

1000 100 10 NONE

28. Select the correct English translation corresponding to the Hieroglyph shown:



GOOD KNOW SAY NONE

29. Select the correct English translation corresponding to the Hieroglyph shown:

BASIN WELL WATER NONE

30. With which metal is the snake's body plated?

- Bronze
- Silver
- Gold
- None

31. How many sailors died in the storm?

- 12
- 120
- 140
- None

32. What is your gender?

- Male
- Female
- Prefer not to say

33. What is your age?

____

34. How many hours you spend (aprox.) playing digital games per week?

____

35. Which digital games in particular?

____

36. What is your course of study?

____

37. To which degree do you agree with the following statements from <not at all - 0> to <extremly - 4>?

	0	1	2	3	4
I think the game is usefull to learn hieroglyphs	⬭	⬭	⬭	⬭	⬭
I would prefer this game over books or flashcards	⬭	⬭	⬭	⬭	⬭
I think the game is fun to play	⬭	⬭	⬭	⬭	⬭
I think i learned something	⬭	⬭	⬭	⬭	⬭
The tutorial successfully explained the basic controls	⬭	⬭	⬭	⬭	⬭
The hints provided useful information	⬭	⬭	⬭	⬭	⬭
The lexicon is a useful tool	⬭	⬭	⬭	⬭	⬭
I used the lexicon often	⬭	⬭	⬭	⬭	⬭

38. Which elements of the game you especially liked?

39. Which elements of the game you especially disliked?

List of Figures

List of Tables

Bibliography

[1] B. Schmitz, R. Klemke, and M. Specht. "Effects of mobile gaming patterns on learning outcomes: A literature review". In: *International Journal of Technology Enhanced Learning* 4.5-6 (2012). cited By 26, pp. 345–358. DOI: 10.1504/IJTEL.2012.051817.

[2] R. Dörner, S. Göbel, W. Effelsberg, and J. Wiemeyer. *Serious Games: Foundations, Concepts and Practice*. Jan. 2016. ISBN: 978-3-319-40611-4. DOI: 10.1007/978-3-319-40612-1.

[3] C. C. Abt. *Serious games*. University press of America, 1987.

[4] U. Ritterfeld, M. Cody, and P. Vorderer. *Serious games: Mechanisms and effects*. cited By 308. 2009, pp. 1–530. DOI: 10.4324/9780203891650.

[5] F. Bellotti, R. Berta, A. De Gloria, and L. Primavera. "Enhancing the educational value of video games". In: *Computers in Entertainment* 7.2 (2009). cited By 50. DOI: 10.1145/1541895.1541903.

[6] L. A. Annetta. "The "I's" Have It: A Framework for Serious Educational Game Design". In: *Review of General Psychology* 14.2 (2010), pp. 105–113. DOI: 10.1037/a0018985.

[7] E. Anderson, L. McLoughlin, F. Liarokapis, C. Peters, P. Petridis, and S. de Freitas. "Developing serious games for cultural heritage: A state-of-the-art Review". In: *Virtual Reality* 14.4 (2010). cited By 156, pp. 255–275. DOI: 10.1007/s10055-010-0177-3.

[8] M. Zyda. "From visual simulation to virtual reality to games". In: *Computer* 38.9 (2005), pp. 25–32.

[9] I. Caponetto, J. Earp, and M. Ott. "Gamification and education: A literature review". In: vol. 1. cited By 92. 2014, pp. 50–57.

[10] F. Bellotti, M. Ott, S. Arnab, R. Berta, S. De Freitas, K. Kiili, and A. De Gloria. "Designing serious games for education: From pedagogical principles to game mechanisms". In: vol. 2011-January. cited By 23. 2011, pp. 26–34.

[11] J. Torrente, P. Moreno-Ger, B. Fernández-Manjón, and Á. Del Blanco. "Game-like simulations for online adaptive learning: A case study". In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 5670 LNCS (2009). cited By 24, pp. 162–173. DOI: 10.1007/978-3-642-03364-3_21.

[12] N. Shaker, S. Asteriadis, G. Yannakakis, and K. Karpouzis. "Fusing visual and behavioral cues for modeling user experience in games". In: *IEEE Transactions on Cybernetics* 43.6 (2013). cited By 37, pp. 1519–1531. DOI: 10.1109/TCYB.2013.2271738.

[13] C. Grappiolo, Y.-G. Cheong, J. Togelius, R. Khaled, and G. Yannakakis. "Towards player adaptivity in a serious game for conflict resolution". In: cited By 20. 2011, pp. 192–198. DOI: 10.1109/VS-GAMES.2011.39.

[14] W.-Y. Chan, J. Qin, Y.-P. Chui, and P.-A. Heng. "A serious game for learning ultrasound-guided needle placement skills". In: *IEEE Transactions on Information Technology in Biomedicine* 16.6 (2012). cited By 27, pp. 1032–1042. DOI: 10.1109/TITB.2012.2204406.

[15] M. Popescu, S. Arnab, R. Berta, J. Earp, D. S, M. Romero, I. Stefan, and M. Usart. "Serious games in formal education: Discussing some critical aspects". In: *5th European Conference On Game Bases Learning* 2011 (Oct. 2011).

[16] P. Mishra and M. J. Koehler. ""Technological pedagogical content knowledge: A framework for teacher knowledge"". English (US). In: *Teachers College Record* 108.6 (June 2006), pp. 1017–1054. ISSN: 0161-4681. DOI: 10.1111/j.1467-9620.2006.00684.x.

[17] B. S. Bloom et al. "Taxonomy of educational objectives. Vol. 1: Cognitive domain". In: *New York: McKay* 20 (1956), p. 24.

[18] R. Garris, R. Ahlers, and J. E. Driskell. "Games, Motivation, and Learning: A Research and Practice Model". In: *Simulation & Gaming* 33.4 (2002), pp. 441–467. DOI: 10.1177/1046878102238607.

[19] L. Hinojosa. "Contributions of Educational Psychology to University Education". In: *Psychology* 06 (Jan. 2015), pp. 177–190. DOI: 10.4236/psych.2015.63017.

[20] J. Lave and E. Wenger. *Situated Learning: Legitimate Peripheral Participation*. Learning in Doing: Social, Cognitive and Computational Perspectives. Cambridge University Press, 1991. DOI: 10.1017/CBO9780511815355.

[21] A. Y. Kolb and D. A. Kolb. "Learning Styles and Learning Spaces: Enhancing Experiential Learning in Higher Education". In: *Academy of Management Learning & Education* 4.2 (2005), pp. 193–212. ISSN: 1537260X. DOI: 10.2307/40214287.

[22] J. Bransford, A. Brown, and R. R. Cocking. "How people learn : brain, mind, experience, and school". In: 1999.

[23] U. Bronfenbrenner. *The ecology of human development*. Harvard university press, 1979.

[24] M. Radovan and D. Makovec. "Relations between students' motivation, and perceptions of the learning environment". In: *Center for Educational Policy Studies Journal* 5.2 (2015), pp. 115–138.

[25] J. Sweller. "Cognitive load theory, learning difficulty, and instructional design". In: *Learning and Instruction* 4.4 (1994), pp. 295–312. ISSN: 0959-4752. DOI: 10.1016/0959-4752(94)90003-5.

[26] R. Ryan and E. Deci. "Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being." In: *The American psychologist* 55 1 (2000), pp. 68–78.

[27] C. Ames. "Classrooms: Goals, structures, and student motivation." In: *Psychology* 84(3) (1992), pp. 261–271. DOI: 10.1037/0022-0663.84.3.261.

[28] L. J. Rawsthorne and A. J. Elliot. "Achievement Goals and Intrinsic Motivation: A Meta-Analytic Review". In: *Personality and Social Psychology Review* 3.4 (1999), pp. 326–344. DOI: 10.1207/s15327957pspr0304_3.

[29] K. Ohtani, R. Okada, T. Ito, and M. Nakaya. "A Multilevel Analysis of Classroom Goal Structures' Effects on Intrinsic Motivation and Peer Modeling: Teachers' Promoting Interaction as a Classroom Level Mediator". In: *Psychology* 04 (Jan. 2013), pp. 629–637. DOI: 10.4236/psych.2013.48090.

[30] M. Csikszentmihalyi. *Beyond boredom and anxiety*. San Franciso: Jossey-Bass, 1975.

[31] P. Sweetser and P. Wyeth. "GameFlow: a model for evaluating player enjoyment in games". In: *Computers in Entertainment (CIE)* 3.3 (2005), pp. 3–3.

[32] M. Csikszentmihalyi. *Flow: The Psychology of Optimal Experience*. Harper & Row, 1990.

[33] M. Csikszentmihalyi. *Flow and the Foundations of Positive Psychology. The Collected Works of Mihaly Csikszentmihalyi*. Springer Netherlands, 2014.

[34] M. Csikszentmihalyi and I. S. Csikszentmihalyi. *Optimal experience: Psychological studies of flow in consciousness*. New York: Cambridge University Press, 1988.

[35] M. Csikszentmihalyi. *Finding Flow: The Psychology of Engagement with Everyday Life*. Finding Flow: The Psychology of Engagement with Everyday Life. New York, NY, US: Basic Books, 1997. ISBN: 978-0-465-04513-6.

[36] F. Massimini and M. Carli. "The systematic assessment of flow in daily experience". In: *Optimal Experience: Psychological Studies of Flow in Consciousness*. Ed. by M. Csikszentmihalyi and I. S. Csikszentmihalyi. Cambridge University Press, 1988, pp. 266–287. DOI: 10.1017/CBO9780511621956.016.

[37] P. Sweetser, D. M. Johnson, and P. Wyeth. "Revisiting the GameFlow model with detailed heuristics". In: 2012.

[38] P. Sweetser, D. M. Johnson, P. Wyeth, A. Anwar, Y. Meng, and A. Ozdowska. "GameFlow in Different Game Genres and Platforms". In: *Computers in Entertainment* 15 (2017), 1:1–1:24.

[39] V. Nagalingam, R. Ibrahim, and R. Che. "EDUGXQ: User Experience Instrument for Educational Games' Evaluation". In: *International Journal of Advanced Computer Science and Applications* 11 (Jan. 2020). DOI: 10.14569/IJACSA.2020.0110170.

[40] A. Hoblitz. *Spielend Lernen im Flow. Die motivationale Wirkung von Serious Games im Schulunterricht*. Springer VS, Wiesbaden, 2015.

[41] K. Kiili, S. [Freitas], S. Arnab, and T. Lainema. "The Design Principles for Flow Experience in Educational Games". In: *Procedia Computer Science* 15 (2012). 4th International Conference on Games and Virtual Worlds for Serious Applications(VS-GAMES'12), pp. 78–91. ISSN: 1877-0509. DOI: https://doi.org/10.1016/j.procs.2012.10.060. URL: http://www.sciencedirect.com/science/article/pii/S1877050912008228.

[42] D. Pavlas. "A Model Of Flow And Play In Game-based Learning The Impact Of Game Characteristics, Player Traits, And Player States". In: 2010.

[43] J. Sinclair, P. Hingston, M. Masek, and K. Nosaka. "Testing an exergame for effectiveness and attractiveness". In: Jan. 2011, pp. 1–8. DOI: 10.1109/ICEGIC.2010.5716909.

[44] V. V. Abeele, K. Spiel, L. Nacke, D. Johnson, and K. Gerling. "Development and validation of the player experience inventory: A scale to measure player experiences at the level of functional and psychosocial consequences". In: *International Journal of Human-Computer Studies* 135 (2020), p. 102370. ISSN: 1071-5819. DOI: https://doi.org/10.1016/j.ijhcs.2019.102370. URL: http://www.sciencedirect.com/science/article/pii/S1071581919301302.

[45] J. H. Brockmyer, C. M. Fox, K. A. Curtiss, E. McBroom, K. M. Burkhart, and J. N. Pidruzny. "The development of the Game Engagement Questionnaire: A measure of engagement in video game-playing". In: *Journal of Experimental Social Psychology* 45.4 (2009), pp. 624–634. ISSN: 0022-1031. DOI: https://doi.org/10.1016/j.jesp.2009.02.016. URL: http://www.sciencedirect.com/science/article/pii/S0022103109000444.

[46] M.-T. Cheng, H.-C. She, and L. Annetta. "Game immersion experience: its hierarchical structure and impact on game-based science learning". In: *Journal of Computer Assisted Learning* 31.3 (2015), pp. 232–253.

[47] W. A. IJsselsteijn, Y. A. W. de Kort, and K. Poels. "The Game Experience Questionnaire". English. In: (2013).

[48] C. Jennett, A. L. Cox, P. Cairns, S. Dhoparee, A. Epps, T. Tijs, and A. Walton. "Measuring and defining the experience of immersion in games". In: *International Journal of Human-Computer Studies* 66.9 (2008), pp. 641–661. ISSN: 1071-5819. DOI: https://doi.org/10.1016/j.ijhcs.2008.04.004. URL: http://www.sciencedirect.com/science/article/pii/S1071581908000499.

[49] K. Poels, Y. de Kort, and W. IJsselsteijn. *D3.3 : Game Experience Questionnaire: development of a self-report measure to assess the psychological impact of digital games*. English. Technische Universiteit Eindhoven, 2007.

[50] V. Vanden Abeele, L. E. Nacke, E. D. Mekler, and D. Johnson. "Design and Preliminary Validation of The Player Experience Inventory". In: *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts*. CHI PLAY Companion '16. Austin, Texas, USA: Association for Computing Machinery, 2016, pp. 335–341. ISBN: 9781450344586. DOI: 10.1145/2968120.2987744. URL: https://doi.org/10.1145/2968120.2987744.

[51] E. L.-C. Law, F. Brühlmann, and E. D. Mekler. "Systematic Review and Validation of the Game Experience Questionnaire (GEQ) - Implications for Citation and Reporting Practice". In: *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play*. CHI PLAY '18. Melbourne, VIC, Australia: Association for Computing Machinery, 2018, pp. 257–270. ISBN: 9781450356244. DOI: 10.1145/3242671.3242683. URL: https://doi.org/10.1145/3242671.3242683.

[52] L. Nacke. "Affective ludology: Scientific measurement of user experience in interactive entertainment". PhD thesis. Blekinge Institute of Technology, 2009.

[53] D. Pavlas, K. Heyne, W. Bedwell, E. Lazzara, and E. Salas. "Game-based learning: The impact of flow state and videogame self-efficacy". In: *Proceedings of the human factors and ergonomics society annual meeting*. Vol. 54. 28. SAGE Publications Sage CA: Los Angeles, CA. 2010, pp. 2398–2402.

[54] J. Schell. *The Art of Game Design: A book of lenses*. CRC press, 2008.

[55] J. Seering, R. Mayol, E. Harpstead, T. Chen, A. Cook, and J. Hammer. "Peer Feedback Processes in the Game Industry". In: Oct. 2019, pp. 427–438. ISBN: 978-1-4503-6688-5. DOI: 10.1145/3311350.3347176.

[56] J. Tremblay, B. Bouchard, and A. Bouzouane. "Adaptive Game Mechanics for Learning Purposes - Making Serious Games Playable and Fun". In: *CSEDU*. 2010.

[57] J. Chen. "Flow in games (and everything else)". In: *Communications of the ACM* 50.4 (2007), pp. 31–34.

[58] S.-W. Um, T. Kim, and J.-S. Choi. "Dynamic Difficulty Controlling Game System". In: *2007 Digest of Technical Papers International Conference on Consumer Electronics* (2007), pp. 1–2.

[59] M. Zohaib. "Dynamic Difficulty Adjustment (DDA) in Computer Games: A Review". In: *Advances in Human-Computer Interaction* 2018 (Nov. 2018), pp. 1–12. DOI: 10.1155/2018/5681652.

[60] A. Baldwin. "Balancing Act: The Effect of Dynamic Difficulty Adjustment in Competitive Multiplayer Video Games". PhD thesis. Queensland University of Technology, 2016.

[61] C. Bailey and M. Katchabaw. "An Expreimental Testbed to Enable Auto-dynamic Difficulty in Modern Video Games". In: 2005.

[62] M. Ninaus, K. Tsarava, and K. Moeller. "A Pilot Study on the Feasibility of Dynamic Difficulty Adjustment in Game-Based Learning Using Heart-Rate". In: *Games and Learning Alliance: 8th International Conference, GALA 2019, Athens, Greece, November 27–29, 2019, Proceedings*. Springer Nature. 2019, p. 117.

[63] C. Liu, P. Agrawal, N. Sarkar, and S. Chen. "Dynamic difficulty adjustment in computer games through real-time anxiety-based affective feedback". In: *International Journal of Human-Computer Interaction* 25.6 (2009). cited By 135, pp. 506–529. DOI: 10.1080/10447310902963944.

[64] R. Hunicke and V. Chapman. "AI for dynamic difficulty adjustment in games". In: vol. WS-04-04. cited By 128. 2004, pp. 91–96.

[65] A. Zook and M. Riedl. "A temporal data-driven player model for dynamic difficulty adjustment". In: cited By 36. 2012, pp. 93–98.

[66] N. Shaker, G. Yannakakis, J. Togelius, M. Nicolau, and M. O'neill. "Evolving personalized content for super mario bros using grammatical evolution". In: cited By 26. 2012, pp. 75–80.

[67] A. Denisova and P. Cairns. "Adaptation in digital games: The effect of challenge adjustment on player performance and experience". In: cited By 20. 2015, pp. 97–102. DOI: 10.1145/2793107.2793141.

[68] J. Hamari, D. J. Shernoff, E. Rowe, B. Coller, J. Asbell-Clarke, and T. Edwards. "Challenging games help students learn: An empirical study on engagement, flow and immersion in game-based learning". In: *Computers in Human Behavior* 54 (2016), pp. 170–179. ISSN: 0747-5632. DOI: https://doi.org/10.1016/j.chb.2015.07.045. URL: http://www.sciencedirect.com/science/article/pii/S074756321530056X.

[69] D. A. Plecher, C. Eichhorn, A. Köhler, and G. Klinker. "Oppidum - A Serious-AR-Game About Celtic Life and History". In: *Games and Learning Alliance - 8th International Conference, GALA 2019, Athens, Greece, November 27-29, 2019, Proceedings*. Vol. 11899. Lecture Notes in Computer Science. Springer, 2019, pp. 550–559. DOI: 10.1007/978-3-030-34350-7_53.

[70] S. Leitner. *So lernt man lernen: Der Weg zum Erfolg (How to learn to learn)*. Freiburg im Breisgau, Herder Verlag, 2002, 1972.

[71] M. Höveler-Müller. *Hieroglyphen Lesen und Schreiben in 24 einfachen Schritten*. C.H. Beck oHG, 2014.

[72] C. Ziegler. *The Pharaohs*. Thames & Hudson, 2002.

[73] H. W. M. Susan A. Jackson. "Development and Validation of a Scale Measure Optimal Experience: The Flow State Scale". In: (1996).

[74] D. A. Plecher, F. Herber, C. Eichhorn, A. Pongratz, G. Tanson, and G. Klinker. "HieroQuest - A Serious Game for Learning Egyptian Hieroglyphs". In: *Journal on Computing and Cultural Heritage (JOCCH)* (). (accepted).