Digital Educational Games: Methodologies for Evaluating the Impact of Game Type

Thesis submitted for the degree of
Doctor of Philosophy
at the University of Leicester

by
Stephanie Heintz
Department of Computer Science
University of Leicester

2016

DIGITAL EDUCATIONAL GAMES: METHODOLOGIES FOR EVALUATING THE IMPACT OF GAME TYPE
- BY STEPHANIE HEINTZ

Abstract

The main research question addressed in this thesis is how the choice of game type influences the success of digital educational games (DEG), where success is defined as significant knowledge gain in combination with positive player experience. Games differ in type if they differ at least by one game feature.

As a first step we identified a comprehensive set of unique game features, summarised in the Game Elements-Attributes Model (GEAM), where elements are the defining components that all games share (e.g. Challenges) and attributes are their possible implementation (e.g. time pressure).

To deepen the understanding of relationships amongst game features, we conducted a survey based on the GEAM and received 321 responses. Using hierarchical clustering, we grouped 67 games, selected by the survey respondents, in terms of similarity and mapped the identified clusters on a 2D space to visualise their difference in distance from each other. On the resulting Game Genre Map, five main areas were detected, which proved to conform mostly to a selection of existing game genres. By specifying their GEAM attributes, we redefined these genres: Mini-game, Action, Adventure, Resource, and Role-play.

Based on the aforementioned groundwork, two empirical studies were conducted. Study 1 compared three DEGs of the Mini-game genre, differing in a single GEAM attribute - time pressure vs. puzzle solving and abstract vs. realistic graphics. Study 2 compared DEGs of different genres which vary in the implementation of several GEAM attributes. For both studies, statistically significant differences were found in learning outcome, for Study 2 also in the player experience dimensions: Flow, Tension, Challenge, and Negative Affect. However, the influences of the covariates - learning and play preconditions, learning style, and personality traits - were not confirmed. Further research based on the methodological frameworks developed is needed.

Acknowledgements

First of all I want to express my deepest gratitude to my supervisor Dr. Effie Law. Her guidance and advice was inspiring and pointed me in the right direction without confining me to specific approaches. In many fruitful discussions her input led to new ideas and her feedback shaped my research and writing skills.

I would also like to thank the students and staff of the Computer Science department at the University of Leicester, for making it such a nice place to work at. It was particularly great to have received all the support for my research from Dr. Michael Hoffman, Dr. Rob van Stee, and Dr. Tom Ridge who let me carry out my surveys and experiments during their Java sessions, from Gilbert Laycock who gave me great technical support, and from fellow PhD students who volunteered to test my games. My thanks also go to the Computer Science department of Furtwangen University, especially to the professors Jirka Dell'Oro-Friedl, Wilhelm Walter, and Bruno Friedman.

I am particularly thankful to everyone who participated in the surveys and studies which I conducted during the course of my research. Without them, this work would not have been possible and so my heartfelt thank-you goes to each individual participant who was willing to spend time to take part.

Moreover I am thankful for my fellow PhD students, with whom I could share the experience of aiming for the degree. Thanks to them, I found Leicester to be a welcoming and enjoyable place. I am particularly glad to have met Gabi, who became my flatmate and made Leicester feel like home, but also Octavian, who always cheered me up.

Finally I would like to thank my family for being my reliable support: my parents and especially my brother Matthias who was always available for discussions and any support needed. But most important was the support from my husband Sebastian. His interest in my work and constant encouragement helped me to pursue the completion of this thesis and I am grateful to have him by my side.

Contents

List of Tables

List of Figures

Abbreviations

2D two-dimensional.

AC Abstract Conceptualisation.

AE Active Experimentation.

AI artificial intelligence.

ANCOVA analysis of covariance.

ANOVA analysis of variance.

App application.

AR augmented reality.

BFI Big Five Inventory.

C cluster.

CE Concrete Experience.

CharQ learner characteristics questionnaire.

CompQ comparison questionnaire.

CS+Mgt Computing with Management course at the University of Leicester.

CS-media Computer Science in Media course at the Furtwangen University.

CS-online Online Media course at the Furtwangen University.

CS-prac Computing course at the University of Leicester (more practice oriented).

CS-theo Computer Science course at the University of Leicester (more theory oriented).

CV covariate.

DEG digital educational game.

Design Media Design course at the Furtwangen University.

DV dependent variable.

ExpQ experience questionnaire.

GCS Game Classification Survey.

GEAM Game Elements-Attributes Model.

GEQ game experience questionnaire.

GER2012 pre-study, conducted in 2012 at the Furtwangen University (Germany).

GER2014 Study 2, conducted in 2014 at the Furtwangen University (Germany).

GS general statement; part of ExpQ.

HC hierarchical clustering.

I/O Input/Output.

ILS index of learning style (Felder and Soloman, 2015).

Interv interview.

IV independent variable.

KnowQ knowledge questionnaire.

M mean value.

MANCOVA multiple analysis of covariance.

MANOVA multiple analysis of variance.

MBTI Myer-Briggs type indicator.

MDS multidimensional scaling.

MMORPG massive multi-player online RPG.

N number of cases included in the statistical analysis.

NPC non-player character.

ProgQ pre-study programming questionnaire.

Q Question.

Qn Questionnaire.

RGT repertory grid technique.

RO Reflective Observation.

RPG Role-play games.

RQ research question.

SD standard deviation.

UK2012 pre-study, conducted in 2012 at the University of Leicester (UK).

UK2013/14 Study 1, conducted in 2013/14 at the University of Leicester (UK).

Glossary

Action genre is one of the five redefined game genres of the Game Genre Map.

Adventure genre is one of the five redefined game genres of the Game Genre Map.

ABSTRACT game is a DEG designed for the comparison with other DEGs in Study 1.

Action is one of the game elements of the Game Elements-Attributes Model.

ACTION 2 game is a DEG designed for the comparison with another DEG in Study 2.

ACTION game is a DEG designed for the comparison with other DEGs in Study 1.

ADVENTURE game is a DEG designed for the comparison with another DEG in Study 2.

Challenge is one of the game elements of the Game Elements-Attributes Model.

Digital games are games which are based on and mediated by a computer.

Educational games are Serious Games with the purpose to educate.

Game attributes are a set of related concepts for each game element that developers can act on.

Game Classification is a set of criteria by which games are differentiated.

Game elements are abstract building blocks of a game which define games and form the fundamental architecture or skeleton of every game.

Game Elements-Attributes Model is a game model with two layers of abstraction, game elements and game attributes, which summarizes and structures the different features of games.

Game features is a generic term used to refer to differences and similarities between games, which is further refined by the terms "game elements" and "game attributes" in Chapter 3.

Game Genre Map shows game clusters of similar games mapped on a 2D space to visualise their differences, resulting in the identification of five main areas, which lead to the redefinition of the five game genres: Action, Adventure, Mini-game, Resource, and Role-play.

Game genres are a subset of game types; in combination game genres form a popular game classification. In Chapter 4 we redefine five game genres and thereafter the term refers to these.

Game types are groups of games, differentiated based on one or more game features.

Goal is one of the game elements of the Game Elements-Attributes Model.

Input/Output is one of the game elements of the Game Elements-Attributes Model.

Mini-game genre is one of the five redefined game genres of the Game Genre Map.

Perspective is one of the game elements of the Game Elements-Attributes Model.

Player is one of the game elements of the Game Elements-Attributes Model.

PUZZLE game is a DEG designed for the comparison with other DEGs in Study 1.

Resource genre is one of the five redefined game genres of the Game Genre Map.

Role-play genre is one of the five redefined game genres of the Game Genre Map.

Reward/Penalty is one of the game elements of the Game Elements-Attributes Model.

Serious Games are games which are designed to serve a purpose other than pure entertainment.

Serious Gaming is the use of games for a purpose other than pure entertainment.

Setting/World is one of the game elements of the Game Elements-Attributes Model.

Structure is one of the game elements of the Game Elements-Attributes Model.

TOOL is an e-learning tool designed for the comparison with DEGs in Study 1.

Tripartite Educational Game Model is a model of the three core components of educational games - game, learning content, and learner/player - and their relation.

Chapter 1

Introduction

1.1 Motivation

In recent years the digital games market has been growing and is now one of the leading industries in the entertainment sector. This growth is accompanied by an interest in using the motivating and engaging nature of games for educational purposes. Processing information by engaging in a game induces learning. Integrating learning content into a game should therefore be a viable option for making learning more effective and enjoyable. While more and more digital educational games (DEGs) are developed for supporting learning and teaching, typically developers give no or only weak rationale for why they design a game in a particular way. Considering the large variety of entertainment games there should be a range of possible solutions for building a DEG. Thus for existing DEGs it is unclear if the same content could have been learned more effectively if the DEGs were designed differently.

To give an example, a DEG for teaching how to access fields of a 2D array using the programming language Java can be a game where a player has to read code and find the accessed fields in limited time. This game could either use simple shapes as graphics or provide the player with a small background story that the identification of the correct target field is to guide planes through fog. Designing appropriate graphics for the plane story is more demanding, but may be more appealing, increasing motivation and focus, with possibly positive impact on the learning. Not only the graphics but the whole game concept can be different. Instead of training the player to be able to read code correctly in limited time, using multiple similar tasks, the game could also use a story to explain the access of the array in more detail, requiring only some small tasks to evaluate the understanding. Each of the described games can be entertaining as well as educational, but may differ in effectiveness and efficiency or enjoyability. In order to understand the full potential of DEGs, research is needed to explore possible design choices, especially how they influence the success of DEGs in terms of learning outcome and player experience.

Addressing this question is essential for several reasons. Recommendations need to be given to game designers as to how their choices influence the success of a DEG. To satisfy the learners, their game preferences may be taken into account, but only if this does not compromise the learning outcome. Also, if similar results can be achieved with less expensive games as compared to more expensive ones, this is of economic relevance. Finally, for existing findings from research on DEGs, it is important to know if results could have been substantially different based on a game with an alternative design.

1.2 Research Goal and Approach

The main research question of this thesis is how design choices made when building a digital educational game (DEG) influence the success of this game in terms of learning outcome and player experience. The measures learning outcome and player experience are chosen based on the two main purposes of an educational game: the player is able to learn something from the game, which has the positive qualities of being motivating, engaging, and fun.

In order to create a successful DEG some general design rules have to be followed, like choosing an appropriate level of difficulty, giving feedback to the player and building a usable interface (further discussed in Chapter 5). Beyond that, the diversity of well-designed entertainment games shows the numerous choices regarding the overall idea of a game. With the aim to optimise the learning outcome, it is important to question the influence of these choices. The main issue in answering this question is the huge variety of games and associated number of options. Given the inherent constraints, this thesis does not aim to give an absolute answer. Instead, it aims to lay a foundation for a systematic approach with some first results. To address the main research question, the following sub research questions (RQs) need to be answered.

1. Game features

   - *RQ 1a:* What are fundamental game features, representing similarities and differences between games?
   - *RQ 1b:* How can we ensure to have found a comprehensive set of features?
   - *RQ 1c:* How can we structure the identified features to facilitate the comparison of games?

2. Game types

   - *RQ 2a:* What relations exist between the game features?
   - *RQ 2b:* How can we make use of these relations to identify significantly different types of games?
   - *RQ 2c:* Can we rely on game genre as a commonly used classification system?

3. Educational game design

   - *RQ 3a:* How can we approach the design of digital educational games?
   - *RQ 3b:* How can learning content be incorporated into a game?
   - *RQ 3c:* What are possible design solutions for DEGs teaching the access of cells, rows, columns, and areas in 2D arrays with Java?

4. Comparison of DEGs with different type

   - *RQ 4a:* How can the success of a DEG in terms of learning outcome and player experience be measured?
   - *RQ 4b:* What additional characteristics of the learner may have an impact on the success of a DEG?
   - *RQ 4c:* What is the impact of game type for DEGs which only differ by one game feature?
   - *RQ 4d:* What is the impact of game type for DEGs which differ by more than one feature?

In this section the general approach to answering these questions is explained and divided into main steps.

An educational game has the aim to entertain, and teach at the same time. Its success is therefore measured by positive player experience in combination with attained learning outcome. To achieve both, the relation between educational content and game is deemed essential. Since the game is the vehicle that delivers the learning content, it defines the possible presentation of the latter, depending on how deep the content is embedded into the game. The content can be mostly separated or else integrated into the game, becoming an essential part of it. For example, when learning how to solve a certain type of math equations, an equation could pop up every time the player attacks an opponent in the game. If answering the question has no effect on the game, it does at most cause some distraction, due to the interruption. However, if the correctness of the answer determines whether the opponent will be defeated or not, it influences the course of the game. For the first scenario, the game could easily be replaced by a different game, e.g. questions come up whenever the player collects a treasure, walks past a certain object or finishes a level. Hence, the research question of how the game type influences the success of a DEG does not apply for games where the content is only loosely connected, as for those the design choices will only influence the attractiveness of the game, which is a question of general game design (not educational game design). Therefore only DEGs with integrated content are considered for this research. It can be argued that these kinds of games have more potential, since they make use of the different game aspects, like challenges faced by the player, levels requiring increasing skill, or stories including information to support learning (for further detail see Section 5.3.3).

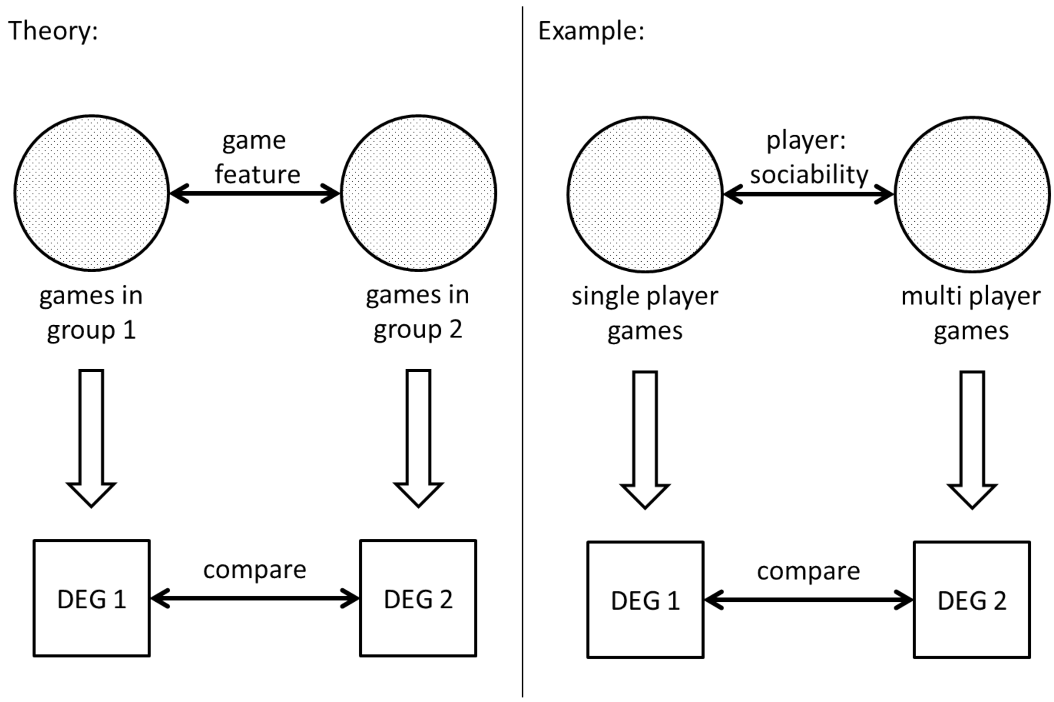


Figure 1.1: Comparison of DEGs based on one feature.

If the content is integrated into the game, its presentation is closely linked to the design of the game. To further analyse the influence of the design choices, the different game features need to be identified which can be controlled by the designer.

If the distinguishing features of games are varied in an experiment their influence can be systematically analysed. The approach would be to divide games into groups, which differ by one feature (Figure 1.1), design a DEG for each group and compare them. For example the feature could be the sociability supported by a game: single-player vs. multi-player (with possible further differentiations in terms of cooperation and competition). A single-player game can then be compared with a multi-player game, teaching the same learning content, to see how the feature influences the presentation of the content and the success of the game.

This approach has one major flaw. One single-player game will not be able to represent all possible single-player games. Testing more than one game is only manageable, if the games are grouped by similarity in order to restrict the numbers of options. For one representative of each group, a comparison could then be conducted between the same game instantiated as single- and multi-player version. However, it raises another issue, since some game ideas may only work either as single or as multi-player game. Both issues can be solved by studying the relations between features. Games sharing a similar set of features can be grouped forming a game type defined by these features. The resulting limited number of game types can either be compared with each other, or since they share similar features, it is also possible to compare single features between sub-types within one type (Figure 1.2).

In general we describe games that differ at least by one feature as games of different type, as will be explained in Chapter 4. The terms sub-type and type in the preceding description are used to differentiate between games that differ more significantly (type) and between games that differ only by one feature (sub-type).

To summarize, the research on how the design of a DEG influences its success will be approached with the following main steps, which each represents a chapter of the thesis (more details are given in Section 1.4):

- Construction of a game model by collecting and structuring the features to represent main similarities and differences among games. (RQ 1a - 1c)
- Identification of significantly different types of games by studying the relations among the identified features of games. (RQ 2a - 2c)
- Development of DEGs with consideration of design rules and discussion on how to integrate learning content into a game. (RQ 3a - 3c)
- Implementation of two studies to compare DEGs of types, differing by only one feature, and DEGs of significantly different types, differing by a set of features. (RQ 4a - 4d)

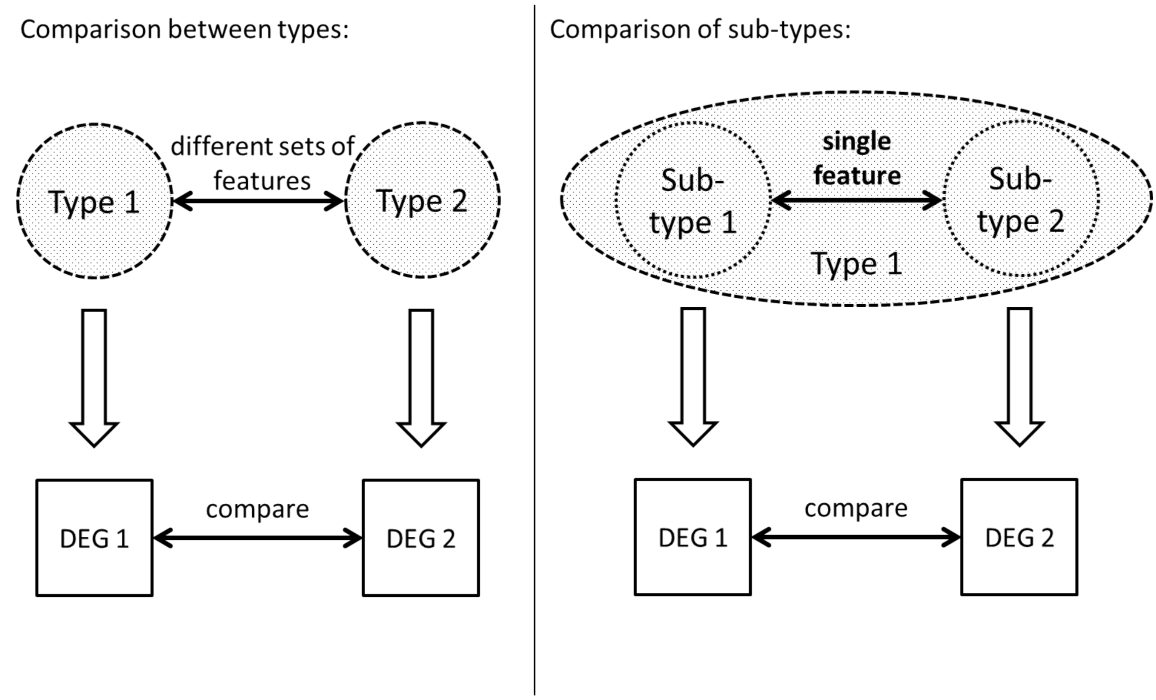


Figure 1.2: Comparing games features between or within types.

1.3 Originality of the Research Work and Main Contributions

Literature addressing the choice of game type for educational games is sparse. Here we highlight key findings of the literature review to argue for the originality of the research work presented in this thesis.

One article authored by Amory et al. (1999) has been found that seems to be closely related to the proposed research question as indicated by its title: "The use of computer games as an educational tool: identification of appropriate game types and game elements". The authors questioned 20 students about four different game types after playing a representative of each type, with the aim to "discover the games type they found most enjoyable; identify game elements that contribute to the fun aspect of the games; and to evaluate student opinion relating to the use of games in education" (p.313). The compared game types are strategy, shoot-em-up, simulation, and adventure games, where strategy and adventure appeared to be favoured. Their research focused only on the interest of the target group, but did not address the applicability of a game type for the inclusion of learning content or the implementation of learning theories. Therefore, this paper gives only a general direction of the research that can be done on this subject.

Frazer et al. (2008) and Frazer (2010) analysed the features different game genres (a form of game types, as further explained in Chapter 4) offer for learning, but their list of genre is not exhaustive, nor are the genres clearly defined. Although the requirements for good learning environments derived from different literature sources are interesting, the game genre analysis is limited, as only three example games represent a whole genre, which leads to flaws, e.g. for RPG/Adventure the games "Final Fantasy X", "Grand Theft Auto III", and "Oblivion" were analysed and found to offer no conversation option, while especially RPG games are popular multi-player online games with plenty of communication among players. Additionally the research is based purely on the observation of entertainment games, so no recommendation or example is given on how to design educational games and thus on applicability and effect of the choice of genre to prove the theoretical findings.

The work of Prensky (2001), Chong et al. (2005), and Rapeepisarn et al. (2008) is also purely theoretical. Prensky (2001, p.156) assigns "learning activities" and suitable "possible game styles" to different kinds of learning content (e.g. questions in form of a quiz show can be used to learn facts). His suggestions are certainly useful, but there is a lack of empirical studies to confirm and complete his findings. While Prensky's game style decisions depend purely on the learning content, Chong et al. (2005) focused on the learner by matching different learning styles with game genres. However, their results are limited as they are based on analyses of only three games from supposedly different genres. Rapeepisarn et al. (2008) aimed to combine both approaches by linking Chong et al.'s (2005) learning styles with Prensky's (2001) learning activities.

One paper was found to actually compare two different educational games teaching the same content (Hwang et al., 2013). With the attempt to support a sequential and a global learning style (Felder and Spurlin, 2005), one game guides the player on a clear path through the game, while the other allows the player to freely choose the order of activities in the game. Results indicate that playing the game version supporting players' learning style would be beneficial, while given the choice between both games, players do not necessarily select the one suited for their style of learning. However, for the pre- and post-test used to measure the learning achievement only the post-test was about the topic taught in the game, while the pre-test was about general knowledge of the course content. Thus it was possibly falsely assumed that all participants had the same prior knowledge of what was taught in the game.

To summarize, our research aims to address the following limitations identified in the previous research (note that "game type" is a more general term and subsumes the term "game genre"):

- *Previous research comparing game types for DEGs, which we had identified and reviewed, was based on unjustified selections of game types.* We solved this issue by identifying a limited set of five game types which cover the landscape of digital games while overlapping as little as possible. For refining the game types, games were analysed and grouped by similarity. Differences between groups were then visualised on a two dimensional map. The areas furthest apart (left, right, top, bottom, middle) on the map each form a game type. We therefore called it a Game Genre Map, which also shows how the five game types are related (presented in Chapter 4). Although we did not compare all five game types in our empirical study, the Game Genre Map allowed us to make an informed decision on which game types are appropriate for teaching the topic selected for the study.

- *Previous studies lack precise definitions for the game types they compare.* We derived the five game types from the Game Genre Map based on a collection of game features by identifying which set of features best characterises each game type. With the aim to find a comprehensive set of game features we identified the core elements of a game and searched for how these elements are implemented in different games. The features were summarised in the Game Elements-Attributes Model (presented in Chapter 3).

- *Previous comparison studies are limited to a single aspect and do not discuss the full range of design options for a DEG.* To the best of our knowledge Hwang et al. (2013) has hitherto been the only study comparing different types of DEGs which teach the same topic. They purely focus on the difference in learning style, but do not further justify their remaining game design decisions. Additionally they did not assess the prior knowledge students had about the topic taught in the games and also did not consider player experience. We are perhaps the first to conduct an empirical study of this extent, comparing DEGs of different game types with regards to the learning outcome and player experience.

1.4 Outline of the Thesis

Chapter 2: First it needs to be clarified what a digital educational game is and how similar terms like serious games, gamification and technology enhanced learning are related to it. Starting with the definition of games the different terms are explored, providing the basis for the research in subsequent chapters.

Chapter 3: In order to approach the main research question, the main features of games need to be explored, particularly the ones which are possibly useful for learning. Identifying the basic elements that games share helps structuring these features. The result is a game model which supports further research. A slightly abbreviate version of this chapter has been published and presented at the DiGRA 2015 conference (Heintz and Law, 2015b).

Chapter 4: Based on the game model developed in the previous chapter, differences and similarities between games can be identified and similar games can be grouped together. Being able to classify games into groups is an important step for approaching the overall research question. By identifying the main features that characterise each group, the features can be researched regarding how they may facilitate teaching certain content. Knowing the defining features of different types also makes it possible to build educational games which represent a certain type of game and can be compared in terms of their learning outcome and player experience. A slightly abbreviated version of this chapter has been published and presented at the CHI PLAY 2015 conference (Heintz and Law, 2015a).

Chapter 5: The next step is the development of educational games. Based on a literature review on the design of educational games, insights into how to build these games have been gained. Multiple educational games have been built for the comparison study presented in Chapter 6. Choosing different game types, which provide different game features, the design decision process and development of the games are described. This includes the consideration of the targeted group of learners, who have been asked in advance to identify appropriate learning content and their requirements for a learning tool. Parts of this chapter have been published and presented at the BCS HCI conference (Heintz and Law, 2012).

Chapter 6: To approach the research question of this thesis, two studies have been conducted, comparing different types of games for teaching the same educational content. The first study, with a between-subject design, analyses two game features throughout three different types of games and compares them with a learning tool, used by the control group. The second study, with a within-subject design, compares two games differing in several game features, with the assumption that more significant results can be obtained, illustrating the extent of the impact the choice of game type may have. Overall, differences between game types have been found in terms of the learning outcome (in both studies) and the player experience (in the second study).

Figure 1.3 gives an overview of the topics covered in each chapter and shows how they are related.

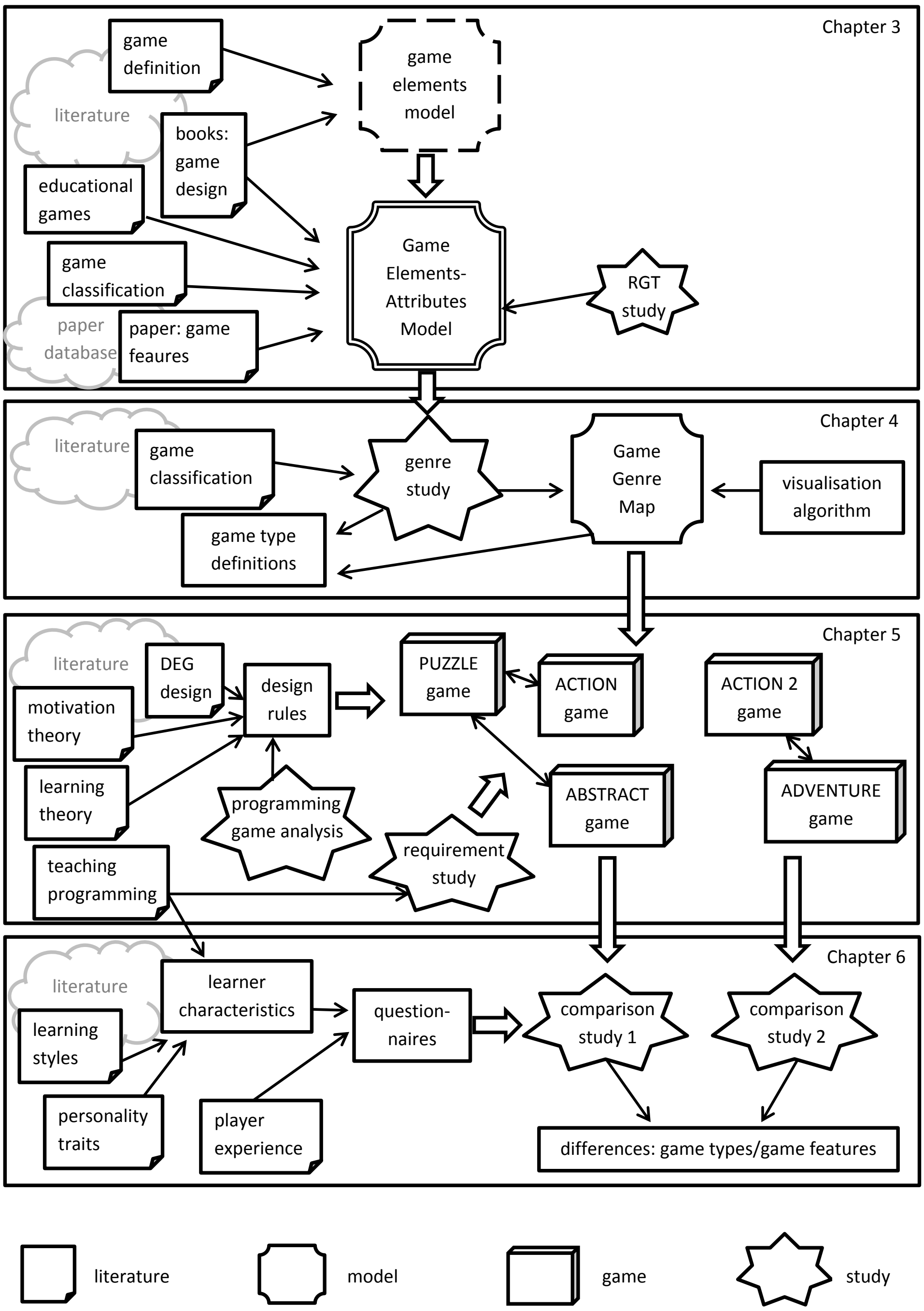


Figure 1.3: Summary and overview of the main research steps.

Chapter 2

Educational Games

This chapter aims to find a suitable definition for educational games and to clarify its relation to similar and sometimes overlapping terms like serious games, game-based learning, edutainment and gamification. To approach the definition of educational games, the term game will be discussed first, followed by the term serious games, which appears to be close in meaning, but more generic, as attaining educational goals is a serious purpose.

2.1 Games

In order to identify what qualifies a game to be called educational or serious, the term game needs to be clarified first. Our understanding of games is mainly based on two sources, both considering and thoroughly discussing a collection of definitions for "game" from different areas.

Laying a theoretical foundation to game design, Salen and Zimmerman (2004) analyse and compare eight definitions for games (Abt, 1970; Avedon, 1971; Caillois, 1961; Costikyan, 1994; Crawford, 1982; Huizinga, 1949; Parlett, 1999; Suits, 1990). Extracting the main defining elements they formulate their own definition:

> "A game is a system in which players engage in an artificial conflict, defined by rules, that results in a quantifiable outcome." (Salen and Zimmerman, 2004, p.80)

Besides stating four elements to be found in every game: player, conflict, rules, and quantifiable outcome, the definition also includes three different perspectives on games. The game itself can be understood as a system. Widening the view and taking the player's perspective on the game into account adds the emotional level of player experience. An even broader view considers the border towards the real world, describing the game as being artificial. Salen and Zimmerman (2004) call these different perspectives schemas, naming them rules (formal schemas), play (experiential schemas), and culture (contextual schemas) and visualise them in three concentric circles. Figure 2.1 shows a slightly adjusted graphic with the three perspectives being renamed to 'system', 'player' and 'real-world', which resonates with the view of Juul (2011) (see below).

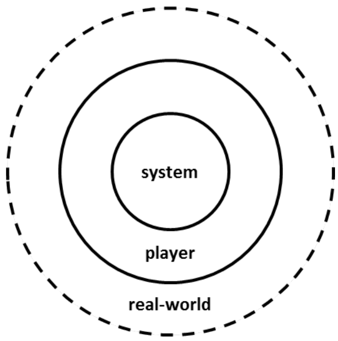


Figure 2.1: Three perspectives on games as visualised by Salen and Zimmerman (2004) with slightly different terms.

Juul (2011) uses the three perspectives to analyse mainly the same game definitions (Avedon, 1971; Caillois, 1961; Crawford, 1982; Huizinga, 1949; Kelley, 1988; Suits, 1990) as Salen and Zimmerman (2004), including theirs. As a result Juul (2011) presents a new definition with six features, each describing aspects of one or multiple perspectives:

> "A game is a rule-based system with a variable and quantifiable outcome, where different outcomes are assigned different values, the player exerts effort in order to influence the outcome, the player feels emotionally attached to the outcome and the consequences of the activity are negotiable."(Juul, 2011, p.36)

While mentioning the same four basic elements of games: player, player effort (requiring conflict or challenge), rules and quantifiable outcome, this definition includes more detail on the player and real-world perspectives than Salen and Zimmerman's (2004). From the player perspective, the challenge in a game demands effort, and the outcome evokes emotional reactions depending on the values attached to it. Besides the emotional effect and effort invested by the player (e.g. time, energy), a game has, according to Juul (2011), negotiable and thus optional real-life consequences (e.g. losing money in a bet a player agreed to).

2.2 Serious Games

Although sometimes criticised for combining two apparently contradicting words, the term serious games is now commonly used. A search for "serious game" on Google returns more than 100 million hits. Djaouti et al. (2011) gave an overview on the origins of the term, identifying Abt's (1970) book with the title "Serious games" as supposedly first source, followed by Sawyer (2002) who used it (only) in the title of his paper that seems to have started the trend of its current popularity.

2.2.1 Comparing definitions for Serious Games

Numerous, somewhat conflicting definitions for serious games exist. A list of identified definitions is presented and discussed below:

- Abt (1970): "Games may be played seriously or casually. We are concerned with serious games in the sense that these games have an explicit and carefully thought-out educational purpose and are not intended to be played primarily for amusement. This does not mean that serious game are not, or should not be, entertaining."
- Michael and Chen (2005): "The simplest definition of serious games, then, is games that do not have entertainment, enjoyment, or fun as their primary purpose." [p.21]
- Zyda (2005): "Serious game: a mental contest, played with a computer in accordance with specific rules, that uses entertainment to further government or corporate training, education, health, public policy, and strategic communication objectives."
- Bergeron (2006): "A serious game is an interactive computer application, with or without a significant hardware component, that has a challenging goal; is fun to play and/or engaging; incorporates some concept of scoring; imparts to the user a skill, knowledge, or attitude that can be applied in the real world."
- Raybourn (2007): "A serious game is defined in this paper as the use of interactive digital technologies for training and education in private, public, government, and military sectors."
- Susi et al. (2007): "Serious Games: The application of gaming technology, process, and design to the solution of problems faced by businesses and other

organizations. Serious games promote the transfer and cross fertilization of game development knowledge and techniques in traditionally non-game markets such as training, product design, sales, marketing, etc."

- Ritterfeld et al. (2009): "As a starting point we define serious games as any form of interactive computer-based game software for one or multiple players to be used on any platform and that has been developed with the intention to be more than entertainment."
- Barbosa and Silva (2011): "They are referred to as games that engage the user and simultaneously contribute to the achievement of a certain objective other than just entertainment, whether the user is aware of that fact or not."

We identified three drawbacks of the above listed definitions, which should be clarified before moving forward with finding an appropriate definition for serious games:

- *A serious game should still be a game, as otherwise the term would be misleading.* Raybourn's (2007) and Susi et al.'s (2007) definitions are not limited to games, as the "use of interactive digital technology" is not unique to games and similarly "the application of gaming technology, process and design" can also include non-game applications making use of gaming technologies. The other authors either mentioned games directly or included a definition of game (Bergeron, 2006; Zyda, 2005).
- *Serious games should not be limited to digital games*, as for example a calculation competition game played in a classroom should also be counted as a serious game. However, Zyda (2005), Bergeron (2006), Raybourn (2007) and Ritterfeld et al. (2009) explicitly limited the term to computer-based or digital games.
- *Serious games should not be limited to certain areas or sectors of application.* We believe that they can be useful in any area and not just for "businesses and other organizations" (Susi et al., 2007). Even the more extensive lists included in Raybourn's (2007) and Zyda's (2005) definitions are problematic as they still imply restriction. Considering the area of application can however be useful for the classification of serious games

Nearly all authors agree that serious games are defined by being entertaining (fun, amusement) while also serving an additional purpose or objective. The only exception are Raybourn (2007) and Susi et al. (2007), for whom the use of game technologies in areas other than gaming is the basic requirement for being a serious game, not requiring it to be a game, which we already criticised above.

According to Abt's (1970) and Ritterfeld et al.'s (2009) definition, serious games have to be specifically developed and intended to be used for a purpose other than entertainment. There are two arguments that make us question this requirement. Firstly, if the serious purpose must be intended by the developer, this excludes the use of "off-the-shelf games" as serious games (Anagnostou and Pappa, 2011, p.950). These are games which are originally designed as entertainment games but found to have a serious value. Secondly, the intention of the developer, if not stated clearly in the game, can be difficult to verify. To determine if a game is serious or not would then rely, to a certain degree, on an assumption.

Similar issues arise from Michael and Chen's (2005) and Zyda's (2005) definition, who defined serious games by their (primary) purpose or objective. The question is: Who determines the purpose of a game? If it is the developer, we point to the same arguments as above. If anyone can identify a purpose other than pure entertainment, then every game is a potential serious game. If someone plays football to exercise, or "Call of Duty" to relax from a day at work[1], this serves as much a serious purpose as if someone trains their balance playing "Soccer heading" from the Wii fit series[2], which is advertised as having a purpose beyond entertainment.

The same conclusion is drawn from the remaining two definitions (Barbosa and Silva, 2011; Bergeron, 2006), which demanded the actual achievement of the objective. Again, every game can potentially meet this requirement. Bergeron's (2006) definition for example says that a serious game "imparts to the user a skill [...] that can be applied in the real world". Every game that uses a mouse as interaction device trains the player's skill to work with this device, which can be useful for the interaction with other serious applications. Of course most people will never play a game for this reason. A solution would be to consider the actual use of a game to determine whether it is serious or not.

2.2.2 Towards a definition of Serious Games

Our argumentation leads to a very broad understanding of the term serious game, since we state that every game has the potential to be a serious game. We are not the first to make this claim. Jantke (2010) also comes to the conclusion that

[1] https://theconversation.com/rough-day-at-work-call-of-duty-can-help-you-recover-26030

[2] http://wiifit.com/training/balance-games.html

"every game is, to some extent, a serious game". He therefore suggests to further analyse the seriousness of a game by asking the following questions: "what is the game useful for", "how good or how much does the game work seriously", "whom does the game serve", "in which condition" and "how". Answering these questions for a game captures well its potential serious use.

However, the term serious game would still be superfluous if a serious purpose can be found for every game. Djaouti et al. (2011) presents a good solution to this issue, by suggesting the use of the term "Serious Gaming". The given definition is in line with our conclusion to consider the actual use of the game rather than the intention of the designer.

> ""Serious Gaming" is a label that refers to any video game used for "serious" purposes, whether the "serious" dimension is or is not designed within the software." (Djaouti et al., 2011, p.121)

Having this broader term for cases where a serious game was not originally developed to serve a serious purpose implies that when separating it from the term "Serious Game", the intention of the designer should be part of the definition. This is the case for Abt's (1970) and Ritterfeld et al.'s (2009) definitions, while the latter restricts Serious Games to computer-based games, which we criticised above. In addition Michael and Chen's (2005) and Barbosa and Silva's (2011) definitions also apply, if we consider the purpose or objective of a game to be determined by the designer.

In conclusion, we use the following definitions for the terms Serious Game and Serious Gaming:

Serious Games = games which are designed to serve a purpose other than pure entertainment

Serious Gaming = the use of games for a purpose other than pure entertainment

2.2.3 Difference between Games and Serious Games

How a serious game differs from an entertainment game can be described via the three perspectives: system, player and real-world (Figure 2.1).

Since a serious game is a game, it needs to include the same elements (player, challenge, rules, outcome) that define a game. This will be discussed in more detail in

Chapter 3. Besides these elements a serious game also needs to provide mechanisms which facilitate its serious purpose, e.g. if a player should improve her balance, balancing needs to be a required activity in the game or if the player should learn a certain information, the information needs to appear somewhere in the game.

Where entertainment game and serious game clearly differ is in their effect on the real-world. For the entertainment game, any impact on the real-world is an unintentional side effect, or as Juul's (2011) calls them "official sanctioned non-optional consequences", like emotional reactions of the player carried over to the real-world, or any effects caused by the player's effort invested in the game, like spending energy and time and possibly even getting injured. For a serious game however, the real-world effect is intended. Szilas and Acosta (2011, p.220) describe this as paradox, as a serious (or educational) game in this point does not meet the definition of a game. If the player is aware of the game's serious purpose and intended real-world impact, which according to Barbosa and Silva's (2011) definition is not necessarily the case, we expect this to have an impact on the player's experience.

2.3 Educational Games

To further differentiate between different types of serious games, we could use the questions suggested by Jantke (2010) and ask how, by whom and for which purpose the game is used. Similarly, Sawyer and Smith (2008) and Djaouti et al. (2011) presented taxonomies for classifying serious games, which include a dimension for purpose and another one for market or scope. Education can be both, either purpose (a game that educates), or scope (a game used in the education sector, e.g. at schools or universities). Since "educational" is an adjective we consider it to describe the purpose of the game. Depending on the scope we could expand the term to further specify whom or which market the game is for, e.g. "Educational Health Game", or "Educational Game for university students", or similar.

2.4 Related Terms

Game based learning is a synonym for educational game, as education and learning are interchangeable terms, as are game and game based. A related term is technology enhanced learning, which is also known as e-learning and does not require a game, but some involvement of technology in the learning process.

Edutainment is the combination of education and entertainment. While addressing any kind of entertainment, it is mainly used in the context of games and appears to be the predecessor term for educational or serious games (Michael and Chen, 2005, p.24).

Simulations are representations of real world environments or processes. They do not require rules which are essential for games. However, they may be games, provided that they include the elements essential to games.

Gamification is according to Deterding et al. (2011) "the use of game design elements in non-game contexts". Enhancing a non-game context with game elements may eventually result in an actual game, but only if the defining elements of a game are integrated.

Figure 2.2 gives a visual overview on how the different terms are related.

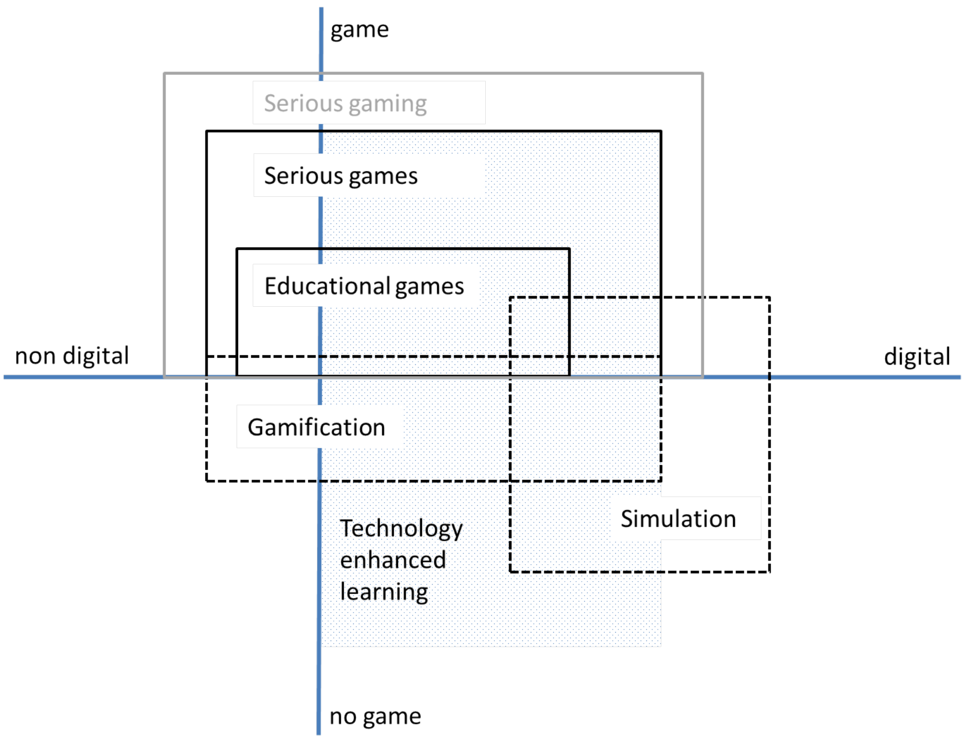


Figure 2.2: Visualisation of relations between educational games and similar terms.

2.5 Digital Educational Games

Games can be based on and mediated by a computer. The computer provides some advantages for the design of games, given its different abilities (Björk, 2013; Salen and Zimmerman, 2004):

- Due to the computational power, game state and appearance can be changed continually, creating responsive environments with immediate feedback and captivating animations.
- Due to its storage, a computer can handle numerous variables, which are required for complex rule sets, accurately simulated worlds and behaviours.
- Due to the programmability, algorithms can be implemented which simulate opponents (making it easier to develop games that can be played alone) and oversee the rules, possibly hiding them from the player.
- Computers can be connected for a fast communication between players who may be far away physically. In addition the internet facilitates the distribution of and access to games.

These advantages make the usage of computers particularly valuable for educational games. Learning content can be simulated with high accuracy in a safe environment. Learners can play alone, at their own pace, and possibly with an individual rule set adapted to their needs. Through the internet they can have access to a variety of games and be connected with others to learn together.

Besides these advantages, there are restrictions when playing a game on a computer. Input and output devices are required to interact with the virtual world. Mouse, keyboard and screen are the dominant input/output devices for personal computers, while game consoles make use of specifically designed controllers. Over the past years, an increasing range of alternative interaction methods have been introduced, like touch-screens, motion control as well as display alternatives like virtual or augmented reality goggles. This is still not equivalent to interactions with and in the real world, like the haptic perception when touching a certain material or sitting next to a person instead of talking to them via a video chat. So even though games played without the use of a computer can usually be transformed into a computer game, they may incur a more or less significant change (e.g. a tennis game played on a computer is quite different from being played in reality, even if a motion controller is used, while "Solitaire" is quite similar being played with virtual or real cards).

Several terms exist to indicate that a game is played on and mediated by a computer, some of the more popular ones being "computer game", "video game", "digital game", or "electronic game". As we only consider games which are administered through computational power, we omit the use of the term "electronic game", since this would include games which in some form make use of electricity, like the game "Hot wire", or any game which uses electrically triggered mechanics. The remaining three terms can be used interchangeably in our opinion. However, the term “video game” may be misunderstood as video is normally non interactive. Similarly the term “computer game” can be problematic, as it is closely related to “personal computers”, which is why for some it does not include consoles, tablets and mobile phones. This leaves the term “digital game” which is predominantly used in this thesis.

Accordingly, we use the established term “digital educational games” (DEGs) for educational games played on a computer, while any combination of the three interchangeable terms and educational would be valid, e.g. “educational digital games” or “educational computer games”.

Chapter 3

Game Elements-Attributes Model: Features of Digital Games

When aiming at finding and comparing alternative DEG options regarding their educational potential, the initial crucial step is to identify game features that support learning. Wilson et al. (2009) present an elaborate list of such features. Since this list is a pure collection of features without any structure relating them to their role in the game system, it is difficult to evaluate if it is comprehensive and base a comparison study on it. Furthermore, Bedwell et al. (2012) identify several of the features as overlapping and address this issue by building categories, adding a structuring layer of abstraction to Wilson et al.'s list. Their results still lack an underlying model that shows how these categories are related, drawing an overall picture of a game and ensuring completeness. To address this issue and find an appropriate model, we conducted a literature search for game models. The models were then evaluated on whether they contain Wilson's game features and a level of abstraction similar to Bedwell et al.'s categories. As none of the identified models satisfied these requirements and modification of an existing model would be problematic, we built a suitable model from scratch called Game Elements-Attributes Model (GEAM). It includes features on two levels of abstractions: elements, similar to Bedwell et al.'s categories, and attributes, which not only cover, but also add detail to Wilson's features and reveal additional ones, possibly allowing further insight on how game features may support learning. Elements and attributes were both derived from literature searches and review. Since the model gives a structured overview on important features of games, it is considered useful not only for DEG comparison, but also for other areas of game research.

3.1 Game Features Supporting Learning

When building a DEG, learning content needs to be included in the game and taught in the process of play, as described in Chapter 5. With learning content and target learner groups identified, an appropriate game type has to be chosen. The features of a game may contribute to the success of a DEG either by enabling the integration of content or by supporting the learning process, i.e. by motivating the learner and thus ensuring that the game is played until the end and all content is perceived. To study these aspects, we lay a foundation by starting from an existing list of game features that are deemed to support learning, derived from the literature, and by means of a game model examine their completeness. Several authors have contributed to generating such a list of features, in the existing literature also called characteristics, or attributes (original terms were kept for the literature review).

Some foundational work was carried out by Malone (1981), who focused on how games can motivate players to learn. Based on his findings he proposed a framework which is built around three categories: challenge, fantasy, and curiosity, each comprising game features with the potential to make learning more interesting. While the motivational aspect of these features was well justified, the discussion on how they could be used for educational purposes falls short. Malone claimed that his framework was more comprehensive than previous theories, but since he enhanced it later with the additional category 'control', as well as a discussion on the three interpersonal motivators: cooperation, competition, and recognition (Malone and Lepper, 1987), other features may be missing as well.

In the following years, more research was conducted in this area, which is summarized in a literature review by Garris et al. (2002). Aiming to unify the findings, they concluded with six characteristics, slightly restructuring but mainly consolidating Malone's list by keeping challenge, fantasy, and control but splitting curiosity into sensory stimuli and mystery and adding rules/goals, which were originally part of challenge. However, information on how the characteristics can be used for learning, besides increasing motivation, interest, and attention, is still limited.

Wilson et al. (2009) built upon the work of Garris et al. (2002). With an updated literature review, also considering game design, they further extended the list of game characteristics, calling them attributes. With twelve additional attributes, this seems to be the most extensive list to date, however its level of completeness is still not clear. An extensive list of 42 references is given with examples of how authors used attributes in their games to teach a certain topic.

Category	Game Features (deemed to support learning)
Action Language	Language/Communication: textual or verbal
Assessment	Assessment: feedback to learn from previous actions Progress: players progress towards the end of the game
Conflict/Challenge	Adaption: adjust difficulty to skill level of player Challenge: progressive, well balanced difficulty + clear goals Conflict: solvable problems Surprise: random element of the game
Control	Control: player's power or influence over elements in game Interaction (Equipment): game responds to player's action
Environment	Location: physical or virtual world the game takes place in
Game Fiction	Fantasy: make-believe, i.e. take on role or simulate process Mystery: sensory or cognitive curiosity to obtain information
Human Interaction	Interaction (Interpersonal): competition, acknowledgement Interaction (Social): activity shared with others
Immersion	Pieces or Players: objects or people included in narrative Representation: perception of game reality, enables focus Sensory Stimuli: temporary acceptance of alternate reality Safety: no consequences other than possibly losing
Rules/Goals	Rules/Goals: criteria of how to win; need to be well-defined

Table 3.1: Wilson et al.'s (2009) game features, sorted by categories identified by Bedwell et al. (2012).

Some of Wilson et al.'s (2009) new attributes were already mentioned by Malone and Lepper (1987) as features subordinate to a category, like adaption (Malone and Lepper: optimal level of difficulty) and assessment (Malone and Lepper: performance feedback), both features of challenge in Malone and Lepper's framework. Others are similar to the existing characteristics, e.g. conflict to challenge, or interaction to control. This suggests that some of Wilson et al.'s attributes are related and can be grouped into categories, similar to Malone and Lepper's (1987) initial work. Bedwell et al. (2012) pursued this objective by using a card sorting technique to capture the mental models of experts (experienced gamers and game designers) on how they felt the attributes should be grouped. Their results suggest 9 categories as shown in Table 3.1, which also shows Wilson et al.'s 19 attributes with short descriptions (the attribute 'progress and surprise' was split by Bedwell et al.).

Due to its extensiveness, Wilson et al.'s (2009) list of game features that support learning is a good starting point for researching the different design options of DEGs.

The primary impediment for its usage is the uncertainty about how comprehensive it is. All main features of games need to be considered for two reasons: (1) They need to be analysed regarding potential further usage for learning and (2) even if a feature does not contribute to learning, the relation between all game features should be studied to gain understanding on how the choice for a certain feature impacts the overall design of a game, i.e. if a certain challenge is chosen, this may impact the features 'interaction' and 'representation' as the challenge needs to be approached and visualized in a certain way. Thus a comprehensive framework capturing all aspects of games and describing games as a whole is needed.

3.2 Existing Game Models

To reiterate, the goal of the following literature search is to find a construct that describes games and serves as a foundation to study game features, their relation and potential use for DEGs. A game model can provide a structural representation of games, where model is defined as "schematic description or representation of something, especially a system or phenomenon that accounts for its properties and is used to study its characteristics" (American Heritage Dictionary).

3.2.1 Literature search

To find existing game models, we conducted a search for this term in three databases: ACM Digital Library, Sage Journals Online, and Science Direct. Due to the limited search base, besides models directly described in the articles, references to models from other sources were also considered, as long as they were published in a journal, a book, or a conference proceeding. Clearly this search is still not exhaustive, since not all models may be found, and constructs with similar qualities to models, such as frameworks, are ignored. However the idea is to get a considerable overview of game models used in the existing literature landscape.

Search strategies

The databases were chosen based on the percentage of search results with a high ranking in Google Scholar. Searching for the term "game model" in combination with "computer game", "video game", or "digital game" (and their plural forms), returned about 1500 hits on Google Scholar. Sorted by relevance, the first 100 hits were examined for their supplying database and the five most frequently found databases identified as ACM Digital Library (12), Science Direct (11), Springer Link eJournals (10), IEEE (10), and Sage Journals Online (5). In a next step each database was searched using the same search term as on Google Scholar, with number of results being (presented in brackets): ACM (73), Science Direct (116), Springer (252), IEEE (177), Sage (22). To keep the scope of the literature search manageable, the selection criterion for databases was the percentage of articles from each database search falling into Google Scholar's top 100. For example, of the 22 articles found on Sage, 5 were listed in the top 100, which is $5/22 \approx 23\%$. As of August 2014, the three databases with the highest percentage were Sage (23%), ACM (16%), and Science Direct (10%).

Filtering search results

The 211 articles from the combined search results of all three databases were screened for the use of the term 'game model' to decide on their relevance. In some articles the search term was not found in the main body, but only mentioned in references or keywords, or the words had different usages (e.g. 'model' was used as a verb instead of a noun). Some articles applied the term 'model' with reference to virtual 3D models, to an early concept/idea for a single game, or a mathematical model in game theory. Others used 'game model' to describe simulations that are based on a mathematical system, like a combat or flight simulation. Even when the term 'model' was used in the sense of the previously given definition, it sometimes described not the game itself, but something in its context or only one aspect, like a cinematic/theatric model, models to analyse aspects in therapeutic games, or technical models like a software model or network model (for details see Table A.1 in appendix).

3.2.2 Analysis of identified game models

37 out of 211 articles use the term 'game model' in the sense of a structural model that describes games. For 14 of the 37 articles no description of or reference to an actual model is given, so the meaning was derived from the context the term was used in (sometimes not with absolute certainty). Of the remaining 23 articles three describe a model, while the remaining 20 reference game models in other sources. Including the referenced models, discarding three from un-reviewed web sources and one unclear reference, a total of ten game models was found. Three of these models

barely include a description of the game itself, as they focus on a different aspect, i.e. the learning process in DEGs (Garris et al., 2002; Kiili, 2007) or the game's relation to reality and meaning (Harteveld, 2011).

From the previous section the following requirements - by which each of the remaining seven models was evaluated - are derived:

- The model includes the features listed by Wilson et al. (2009).
- The model gives a structure with categories similar to Bedwell et al. (2012).

We analysed each model regarding the two requirements. Results for identified inclusions of Wilson et al.'s features (second column) and for categories or other layers of abstractions (third column) are given in Table 3.2. When features used to describe the models could not be matched exactly to those listed by Wilson, the ones identified as being related are listed in round brackets. The numbers in square brackets show how many of Wilson et al.'s features were found to be included in each model. Features which could not be matched with Wilson et al.'s are listed as 'additional' and the ones which rather represent a category as 'too broad'.

3.2.3 Conclusion from search for existing game models

Most of the models include only up to five of Wilson et al.'s (2009) features and do not have any level of abstraction (e.g. categories). The MDA model separates game features into mechanics, describing game components on a data and algorithm level, dynamics, describing the game behaviour at run-time, and aesthetics, describing the player's perception. While mechanics and dynamics describe games as a system, they are only explained on a general level in the model; an elaborate list of features is only given for aesthetics. However, since games may cause different experiences for different players, it is difficult to compare them on the basis of player perception. This issue also applies for some of Wilson et al.'s features, e.g. mystery and surprise, which will be addressed later, in Section 3.3.2. Amory's (2007) model includes the highest number of features and has a hierarchy of multiple levels, but the structure is not well explained and several features seem to be overlapping (e.g. exploration and discovery; narrative spaces, drama, story, and backstory). Multiple theories have been merged in Amory's game object model, which makes it bulky and somewhat incoherent.

Model	Features	Categories
classic game model (Juul, 2003, 2005)	Rules, Outcome (Goals, Assessment), Player Effort (Challenge, Control) [3-5] *additional:* Players attached to Outcome, Negotiable Consequences	No hierarchy
generic gaming and simulation model (Klabbers, 2003)	Actors (Human Interaction), Rules, Resources (Pieces or Players) [3-4]	No hierarchy
MDA: mechanics, dynamics, aesthetics (Hunicke et al., 2004)	Derived from aesthetics: Sensation (Sensory Stimuli), Fantasy, Narrative (Game Fiction), Challenge, Fellowship (Social Interaction), Discovery (Mystery, Location) [6-7] *additional:* Expression, Submission	Separation in mechanics, dynamics, aesthetics
game object model v. II (Amory, 2007)	Communication, Challenges, Problem Space (Conflict), Engagement (Control), Interaction, Exploration (Location), Drama + Story (Fantasy), Social Space (Human Interaction), Authentic + Graphics + Sounds (Sensory Stimuli), Goal [10-11] *additional:* Fun, Technology, Memory, etc.	Up to 4 layers
game design atoms (Brathwaite and Schreiber, 2008)	Players + Avatars + Game bits (Pieces or Players), Challenges, Goals [3] *too broad:* Mechanics, Dynamics *additional:* Theme, Resources, Game State and View	No hierarchy
core elements of the game experience (Calvillo-Gámez et al., 2010)	Interaction, Control, Environment [3] *too broad:* Game-Play *additional:* Ownership, Enjoyment	No hierarchy
active game model (Ruch, 2012)	Rules, Narrative Path (Fantasy), Screen + Speaker + Controller (Control, Interaction), Visual Asset + Music + Sound (Sensory Stimuli) [3-4] *additional:* Physics of Game World	Separation in player, machine, aesthetic, interface

Table 3.2: Game models identified in the literature search, analysed regarding the two requirements.

As it seems difficult to further modify Amory's (2007) model to meet the requirements and since in the MDA model details are missing for the mechanics and dynamics of games, we deemed it necessary to start from scratch when building a game model to address the initial research question. However several insights gained from the analysed models can be taken into account during the process of building the new game model. Juul's (2003; 2005) model seems a good starting point, since it was designed to serve as a definition for games, giving strong arguments for being comprehensive and listing main components shared by all games. Brathwaite and Schreiber's (2008) description of game design atoms is derived from a foundational chapter of their game design book, indicating that similar books on game design might be interesting sources for further input. Some terms have already been identified as being too broad, which shows that the depth and level of detail of the model needs to be decided on, when considering which features to include in the model. Adding more detail to Wilson et al.'s (2009) features can give valuable insight, but keeping the research question in mind, the amount of features should be restricted, to allow an approachable comparison of DEGs. While including an abstract level close to Bedwell et al.'s (2012) categories gives the model a fundamental structure, the aim is to focus on the breadth rather than the depth to achieve comprehensiveness.

3.3 Development of a Game Model

The development of the new game model is divided into two steps, (1) the identification of basic components of a game, which will be called elements and mapped with Bedwell et al.'s (2012) categories and (2) the detection of ways to implement these components in different games, which will be called attributes and should comprise all features listed by Wilson et al. (2009). The model is therefore called Game Elements-Attributes Model (GEAM). Hereafter, to enhance readability the names of elements are capitalized and those of attributes italicised.

3.3.1 Game elements

The first step in building the new game model is to find an underlying structure of elements that define games and serve as the skeleton of a game - the core concepts that all games share, e.g. having 'Goals' as part of each game. Grounding the model on such universal game elements ensures that it is sound and generic. It further facilitates the inclusion of more specific aspects which can be identified by analysing how the elements are implemented in different kinds of games. Game definitions should reveal the crucial components of a game, thus elements are extracted from definitions. The resulting set is then further refined by considering game design literature.

Extracting elements from game definition and design

With the increasing interest in game studies, the issue of finding a solid formal definition for games has been addressed recently. Based on the comparative evaluation of previous work, two oft-cited definitions have been developed by Salen and Zimmerman (2004) and Juul (2011) (see Section 2.1). They agree on the following defining criteria: Rules, Quantifiable Outcome (win or loss), Active Player, and Conflict. Additionally they both point out that games can be considered from three different perspectives: the game as formal system, the player and the game (experiential), and the relation between the game and the rest of the world (Juul, 2011). The game model to be created is supposed to represent a formal system, thus features concerning the other two perspectives are not deliberately considered. However, as Player is an element of this system while also being grounded in reality, the experiential perspective as well as the relation between the game and the rest of the world is automatically included to some degree.

Egenfeldt-Nielsen et al. (2013) note that a formal definition of games is "unconcerned with matters of representation" (p.42) and thus does not need to include aspects such as audiovisual feedback. However, Representation is essential for the design and analysis of games and should therefore be included in the model. Furthermore, Adams (2014) identifies the essential role of Structure in games, acknowledging that the features of a game may change depending on the state or mode of the game (e.g. the player faces a new conflict).

In the first edition of their book about the design of digital games, Rollings and Adams (2003) give a well described set of elements, which are in line with the aforementioned definitions and also cover the aspect of representation. In addition they mention the element "interaction model" which is distinct for digital games, as games "mediated by a computer" (Adams, 2010, p.15). The fact that the elements are still found in later editions of the book and are frequently mentioned in other design books strengthens the assumption of their essentiality. Thus, the core elements of the model are based on Rollings and Adams (2003) with slight modifications: The element "Interaction Model" was renamed to the broader terms "Input/Output"; "Victory/Loss condition", which focuses only on the final outcome of a game, was split into the elements "Goal" and "Rewards/Penalties", as games are often driven by multiple small goals without one final outcome. “Story” is not identified as a core element, but as it usually covers the whole game, it is subsumed by Structure.

To provide a clear and concise representation, the elements of the model as well as their relations are visualized in the form of a diagram, shown in the centre of Figure 3.1. Gameplay as the central part of a game consists of two elements: Actions and Challenges. The gameplay is driven by pre-defined Goals leading to Rewards/Penalty and possibly a victory/loss condition. The interaction between a Player and the game is mediated by Input/Output interfaces. The Setting of a game is the (virtual) space in which it takes place and is shown to the Player through a Perspective (e.g. defined by a camera position). Rules, as the foundational element, are incorporated into the other elements; therefore, they are not analysed independently in this work. The configuration of these elements can change when playing a game; these varying sets of values of the elements are called states or modes, organized through the Structure of a game, which is often implemented as levels and can be supported through a story.

Matching elements with Bedwell et al.’s categories

The categories identified by Bedwell et al. (2012) can be matched to the elements which were chosen as a foundation for the new game model as shown in Table 3.3.

Two of Bedwell et al.’s categories lack a direct representation in the new model since they contribute to multiple elements: game fiction and immersion. Game fiction is included into a game as part of Gameplay and the Game World. The same holds for immersion which is caused by the players’ engagement in the game, thus additionally being attached to Input/Output.

Also, two elements of GEAM have no direct counterpart in Bedwell et al.’s list of categories: Perspective and Structure. Perspective is the link between Game World

Element	Player	Input/ Output	Actions	Challenges	Rules, Goals	Rewards/ Penalties	Setting/ World
Bedwell	Human Interaction	Control	Control, Action Language	Conflict/ Challenge	Rules/ Goals	Assessment	Environment

Table 3.3: Game elements extracted from literature as defining components of games, matched with the categories for game features with potential use for learning as derived by Bedwell et al. (2012).

and Output. Structure enables to track the progress in a game (a feature which Bedwell et al. assigned to the category Assessment), e.g. by providing levels and may also facilitate a story, thus also contributing to Bedwell et al.'s category game fiction.

3.3.2 Game attributes

The second step, leading to the final proposed game model, is to identify attributes that inform how an element can be implemented, e.g. the Setting/World of the game can be two- or three-dimensional. Relevant literature was found based on two search strategies, (1) the identification of research areas concerned with detailed analysis of games and (2) a broad literature search based on a selection of databases and search terms. In addition to this analytical approach, attributes were also derived from results of an empirical study (3). Grouping the attributes under the previously selected foundational game elements generates the second layer of the GEAM.

(1) Analytical approach: specific literature search

Besides the already described game features supporting learning, which were derived from educational games research, two more research areas were found to focus on a detailed understanding of game features: game classification and game design.

Game classification research: Although classifying games by genre is known to have limitations, only a few attempts have been made to develop new classification systems (Section 4.1). Of the existing literature in this field, two papers (Djaouti et al., 2008; Elverdam and Aarseth, 2007) followed the approach of elaborately dissecting games to identify the main differentiating aspects, thereby facilitating the compilation of attributes.

- Djaouti et al. (2008) extract and define so-called 'game bricks', a set of ten core rules to describe the gameplay as central part of video games, by analysing games for the most basic recurrent aspects.
- Elverdam and Aarseth (2007) propose a wider list of attributes by thoroughly comparing games, regarding where and how they differ, generating a collection of so called 'dimensions' and their optional values.

Game design research: Its focus is more practical and often less formal, discussing various aspects of games to give advice on how to create successful games. Several game design books were reviewed (e.g. Koster, 2005; Rogers, 2010; Schell, 2008), spanning a diversity of topics on games, such as technical implementation, the game market, gameplay experience, and storytelling. Some of the books cover game architecture and basic game concepts, but only briefly (e.g. Brathwaite and Schreiber, 2008). Four books (Adams, 2010; Björk and Holopainen, 2005; Fullerton, 2008; Perry and DeMaria, 2009) are found to be rich sources for deriving attributes.

- Adams (2010), as author of the book from which the GEAM elements were mainly derived, also discusses associated game attributes.
- Fullerton (2008) presents a list of elements that overlap to a large extent with those of GEAM and gives several attributes for them.
- Björk and Holopainen (2005) extract design patterns from games, based on a framework with components somewhat related to GEAM's elements. They present more than 200 patterns, many of which are considered to be too detailed for the inclusion in GEAM.
- Perry and DeMaria (2009) provide long lists of examples on design choices. Specifically, relevant attributes are identified from their discussion on 'experience designing', which covers the elements Game Activities (Actions), Challenges, Goals and Rewards. The listed examples are very detailed, but some are grouped as categories, which can be considered as attributes.

In addition to literature from books, the following paper has been found to be another detailed source for game attributes:

- Owen (2004) describes in his paper the anatomy of games in a still experimental stage. Aiming to identify key features of games, he discusses a set of components, overlapping with the elements of GEAM.

(2) Analytical approach: broad literature search

With a similar selection process as for the game model search, the same three databases were chosen for a broader literature search to identify game attributes: ACM DL, Sage, and Science Direct. To keep results at a manageable size, the search was restricted to abstracts-only. While the terms 'elements' and 'attributes' were

Search term	ACM DL	Sage	Science Direct
game element/s	55 (20)	12 (3)	21 (4)
game component/s	4 (0)	8 (0)	7 (0)
game characteristic/s	3 (3)	15 (0)	16 (1)
game attribute/s	1 (0)	5 (2)	6 (3)

Table 3.4: Number of hits for each database and search term; in brackets: number of papers actually listing game features.

chosen to describe features on the two levels of the GEAM model, their meaning may differ in other sources and similar terms could be used. Taking this into consideration, four search terms were used with the aim to find features which contribute to the selection of attributes for GEAM. These terms are listed in Table 3.4 together with the number of search results for each database. It is noticed that this search is not exhaustive, but aims to identify attributes that may have been missed by the specific literature search.

In the first step, articles from mostly unrelated research areas like math or economics were filtered out (21) along with a few articles that were not accessible or in a language other than English (16). Screening the remaining articles with the search terms, game features were extracted and analysed for possible inclusion as attribute in GEAM (Table 3.4).

Duplicates were removed from the obtained list of attributes and results with similar meaning grouped together to identify representatives. These were then matched to the appropriate game element (Table A.2 in appendix).

(3) Empirical approach: Repertory Grid Technique interview

The approach of comparing games in search for similarities and differences which was used in the sources from game classification research motivated a study for further acquisition of attributes.

Aim and participants: Interviews were conducted with twelve participants (7 male; 5 female; aged 20-32; all were regular gamers for at least one decade except one who had limited game experience) using the repertory grid technique (RGT, Fransella et al., 2004). They were recruited through the social networks of the author; their participations were voluntary without any compensation.

Method and Procedure: Each participant was asked to analyse a set of ten games. They were free to choose nine games, instructed to know them well and trying to cover a broad range of types of games they play. The game "Tetris" was the only game suggested to them as it was known to all participants and contrasted with the selection of mainly story centred 3D games, popular to some of the participants. It was the only game included in all the interviews, while other games were also selected multiple times by different participants.

An interview consisted of two stages. In the first one, three games were randomly drawn from the set. The participant was asked to find an aspect which was similar for two of the games but differentiated them from the third game, resulting in a pair of opposite values (e.g., calm vs. anxious). Then, another triad of games was drawn and the procedure was repeated. This elicitation process ended when no more new pair of values was identified. More than one pair could be identified for one triad of games, and overlapping pairs of value across the triads were possible. In the second stage, the participant was asked to rate with a 5-point scale each of the games evaluated against all pairs of values (with duplicates removed) elicited in the first stage. For instance, a scale [calm - - - - - anxious] was presented to a participant, if she thought that the game "Tetris" should be characterised as calm, then she rated it with 2 or 1 (the leftmost "-" is 1).

Evaluation: Often participants found it hard to name an appropriate opposite other than negation (e.g., combat vs. no combat), or they would name a concept to group two of the games and another concept to differentiate them from the third game. While these two concepts might not have any contrasting meaning (e.g., fight vs. explore), they were still useful results. All the derived values were then assigned to the game elements (Table A.3 in appendix).

Results by game element

Player: In game design books (Björk and Holopainen, 2005; Fullerton, 2008) as well as game classification literature (Elverdam and Aarseth, 2007), it is described that players may *play alone* or *with others*. As soon as other players are involved, their relationship can be studied as being *competitive* or *cooperative*. In addition Fullerton (2008) notes that players may still *play on their own* even if others are involved in the game (e.g. in multiplayer online games someone may still play without interacting with others). Further backing these findings, results from RGT and sources from the broad literature search reveal another aspect describing player relation: *communication*. The identified attributes describe Wilson et al.'s (2009) features interpersonal and social interaction.

Input/Output: Adams (2010) gives a list of input/output devices, which was further extended by the first author to cover commonly used devices for playing digital games in combination with respective input devices, resulting in a detailed list of attributes for Wilson et al.'s interaction (equipment) feature.

Actions, Challenges: Most attributes for these elements were found in Owen's (2004) discussion paper, who lists the Actions *build*, *destroy*, *move* (as journeying), *place*, *collect*, *communicate* and the Challenges *opponent*, *obstacle*, *puzzle*, *quiz*, *reaction time*. Their validity is further confirmed as they are in line with Djaouti et al.'s (2008) basic gameplay rules, either matching them (e.g. move, destroy), being more generic (communicate instead of write) or turning out to be even more fundamental than Djaouti et al.'s rules, as combinations allow to describe them (e.g. avoid = move + obstacle or opponent). However, slight modifications were made based on Djaouti et al.'s work, e.g. the rule 'shoot' was considered by extending *destroy* with *fight*, a form of destruction. Similarly features found in other sources (e.g. from the broad literature search) were compared with the existing list. In this process, three more challenges were added: *limited resources*, *search & find*, and *savability*, as well as *freedom of movement* (Perry and DeMaria, 2009) for the Action *move*. The final list covers the following of Wilson et al.'s features: control, language/communication, challenge, and conflict. As an aspect of fantasy Wilson mentions that the player may take on a role, usually achieved by another important attribute related to Action: a *character* controlled by the player, which may be *personalized* and cause *emotional attachment* (Adams, 2010).

Goals: Goals are closely related to Challenges, as it is the goal to overcome a certain challenge (e.g. beat the clock as goal for the Challenge *time pressure*). The aim is to find features that characterise types of goals without duplicating the description of challenges. Often various goals are given in games, called quests or missions (e.g. Elverdam and Aarseth, 2007; Perry and DeMaria, 2009). Björk and Holopainen (2005) add that the player may have a *choice* in which goals to pursue or even *define their own goals*. The attribute *static* (always the same kind of goal) was added as being opposed to *various*. Besides rules/goals Wilson does not list any more features for this element.

Rewards: The broad literature search revealed a range of articles on gamification, using game elements in a non-game context. Commonly used for this purpose is the element Rewards. When identifying attributes described in these sources, the ones with similar meaning were grouped together and one representative chosen for each group:

- *score*: points, point system, achievement points, assessment
- *praise*: leaderboards, medal system, badges, change of the players' virtual status, title, reputation
- *new levels*: gain levels, access to new game space/people/levels, promotions, greater responsibility, leveling up, advancement, progress
- *gain resources*: experience points, virtual currency, gaining new tools, clothes or currency, money and various items
- *power-ups*: power-ups, improving avatar/city/civilization/business to highest level or their own goal

The representative terms were added to the model and Penalties described accordingly, giving more insight on assessment, another feature by Wilson et al..

Setting/World, Perspective: The game world is a space presented by graphics and sounds. Sources from the broad literature search differentiate for sound between *sound effects* and *music* and for graphics between *fantasy* and *realistic*, also derived from the RGT results, in addition to *abstract*. RGT findings also suggest that games have different *graphical detail*. The game space can be *2D* or *3D* as described by most sources and is shown to the player through a Perspective. Elverdam and Aarseth (2007) and Adams (2010) differentiate between omnipresent (overall view) and vagrant (*avatar-based view*), where based on findings from RGT the omnipresent view may be *fixed* or *freely* movable. While the size of a game world may be difficult to determine, more important for the gameplay is *how much the player can explore* (feature extracted from RGT, but also mentioned by game design books like Perry and DeMaria (2009)). Wilson et al.'s feature location describes the world itself, and fantasy as well as representation describe the way it is presented to the player.

Structure: A commonly known structure of a game is its division in levels (e.g. Björk and Holopainen, 2005). As in some games these may instead be separate missions or chapters, the more generic term *separate* parts versus *continuously* progressing was chosen for the corresponding attribute in the game model. Elverdam and Aarseth (2007) state that a game can be *finite* or *infinite*, to which Björk and Holopainen add the aspect of *replayability*. They also note that games can have a narrative structure, but since not all games have a *story* (Perry and DeMaria, 2009), it is of different importance depending on the game. Story contributes to Wilson et al.'s feature fantasy while otherwise structure is related to progress (e.g. which level, how far in the story).

Excluded game features

The selected attributes directly cover 13 of Wilson et al.'s 19 features (Table 3.1). The remaining six features are included indirectly, as there are two reasons as to why a feature was excluded from the model: (1) it is part of another feature already contained in the model, or (2) it does not describe the game as a system, but from another perspective, e.g. player perception. Both reasons are further discussed below, explaining which features were excluded in the process of building the model and illustrating how Wilson et al.'s remaining features are indirectly considered in GEAM.

If features were already represented by one or more elements, then they were excluded:

- Game mechanics and aesthetics are terms that are above the element level, as they comprise elements, e.g. Kosmadoudi et al. (2013) summarize from various definitions that mechanics describe the possible Action and user Interaction with the game based on Rules. As adding an extra layer above the element level has been avoided lest the model become too bulky and abstract, mechanics and aesthetics were excluded.
- Game content or context is rendered in multiple game features such as *story*, *character*, Game World, and Challenges. Through these elements and attributes it is already included in GEAM.
- Features like virtual items, entities, or inventory, are different terms for objects in a game. As all objects in a game have a purpose, they are represented by the existing elements and attributes, e.g. the Challenge *limited resources* requires items that represent these resources or the Game World requires different decorative items. This also comprises Wilson et al.'s feature pieces or players.
- Aiming to find a consistent level of detail for the whole model, features which appeared as too specific and were covered by a more generic term were not included (e.g. "driving speed" = *move*, "applause" = *sound effect*).

If features describe games as formal systems, they are included. Otherwise if terms describe games from a different perspective, then they are excluded.

- Games can be described from the player point of view (Juul, 2011; Salen and Zimmerman, 2004), regarding player experience rather than the formal elements that may cause this experience. All game features describing the players' cognitive or emotional state were thus not included in the GEAM. Examples of these features are: concentration, curiosity, fun. As games are designed for players, observing their perception of the game is crucial. GEAM does not include these terms, as it focuses purely on the architecture of the game, but it allows analysing each feature with regard to the player (e.g. time pressure might enforce concentration). One player perspective which is partially included in GEAM is players' choice (Björk and Holopainen, 2005), which can be found in the attributes *personalise* (a character), *choice of goal* and *of setting*. Four of Wilson et al.'s features are player related: surprise and mystery are caused by Challenges, *story*, and Game World, sensory stimuli by the Output of graphics and *music*, and adaption depends on the level of difficulty perceived by the player.
- Games can also be described in relation to the real world (Juul, 2011; Salen and Zimmerman, 2004). Features such as boundaries were not considered, nor the relation between game time and real world time (Elverdam and Aarseth, 2007). Mixed-reality games are borderline cases. Features like tangible interfaces could be added to GEAM as *input device*, but seem too specific at the time. Wilson et al.'s feature safety is based on the virtual Game World being detached from the real world and allowing actions to be tested and stories to be told without serious impact.
- All attributes are concepts which are above the implementation level (except for Input/Output which describes the technological interface of the game), so all features that describe the game from a software or technology perspective were excluded: GUI, technology, progress bars, clocks and timers, etc. but also random, used to simulate opponents' unpredictability or to keep a puzzle which has been solved before challenging.

3.4 Game Elements-Attributes Model (GEAM)

As described above, we developed a game model which summarizes the main features of games, also including the ones relevant for learning. The resulting Game

Elements-Attributes Model is presented in Figure 3.1. In the centre of the model the game elements, as core building blocks of a game, are depicted as well as their relations. Each element is connected with a list of attributes, displayed in the surrounding area, building the second layer of the model.

In the broad literature search for attributes a definition for the term game element close to the understanding of the author was found, given by Deterding et al. (2011, p.12): "One solution is to treat game elements as a set of building blocks or features shared by games". Based on this, the following definition for game elements is proposed:

Game elements = abstract necessary features of a game which define games and form the fundamental architecture or skeleton of every game

While elements are the features that are essential to all games, attributes are the features in which games may vary.

Game attributes = a set of related concepts for each game element that developers can act on

While most of the elements and attributes are self-explanatory by their names, a short description has been specified for each attribute.

Player: Two aspects are considered: number of players involved (i.e., single-/multi-player) and possible types of interaction between players (*competition*, *cooperation*, *playing alone*). For instance, in MMORPGs a player can either team up with others or battle them or play on her own. *Communication* is included to address the social aspect of gameplay.

Interaction: In order to interact with the game an interface is needed. Depending on the platform a video game is played on (e.g., computers, consoles, mobile devices), different *input devices* such as special controllers (e.g., a wheel for car racing game) and motion sensors can be used.

Actions:

- *Move*: move objects/characters in a game
- *Place/Position*: place objects/characters at a position in a game
  Note: Moving/placing objects/characters can either be restricted to particular

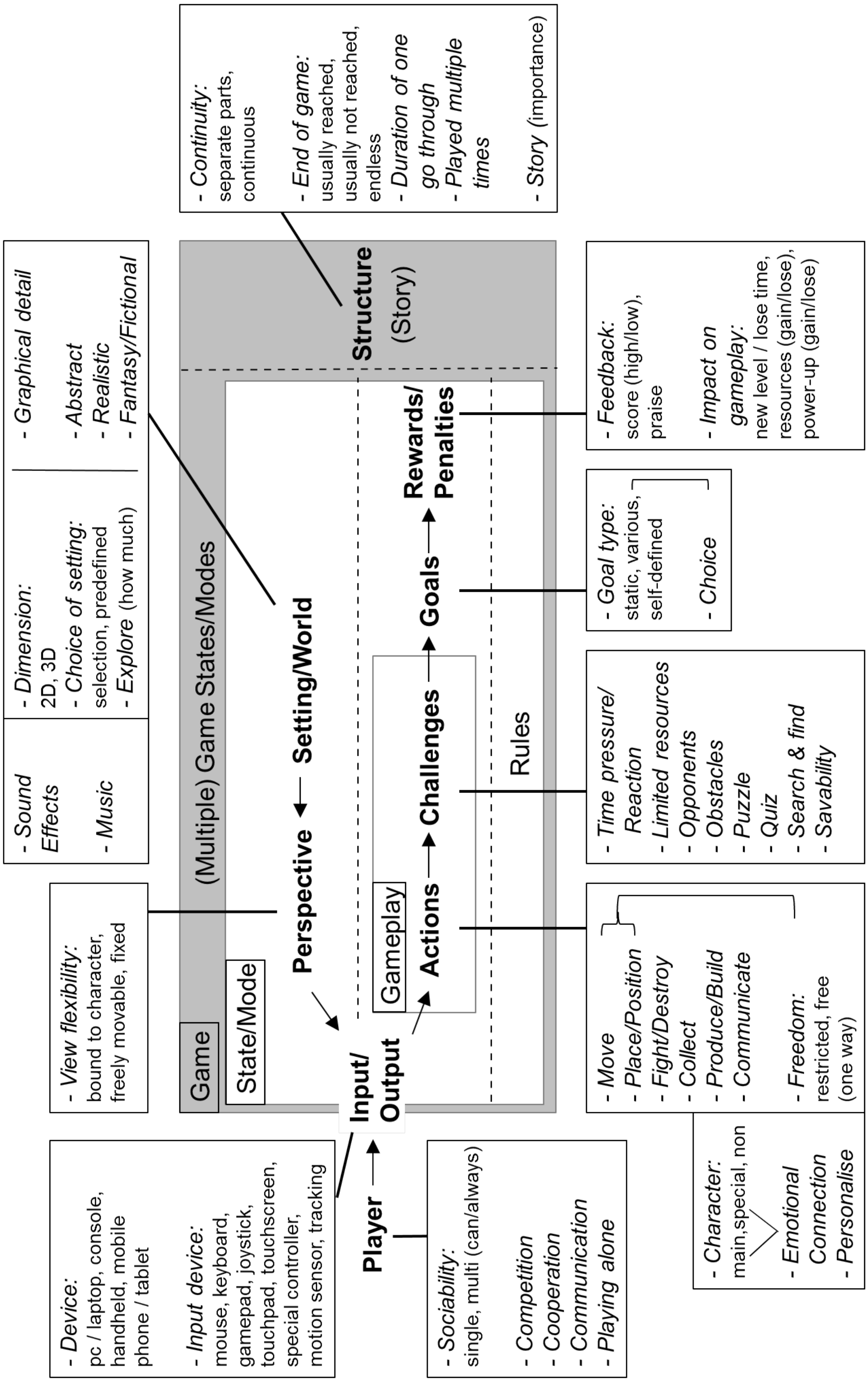


Figure 3.1: Game Elements-Attributes Model (GEAM). Elements in the centre, attributes in surrounding boxes, lines in boxes show dependencies between attributes.

positions or free to any position of the game environment. Movement can also be restricted to a particular direction;

- *Fight/Destroy*: shoot or attack/destroy objects or opponents;
- *Collect*: collect items or resources;
- *Produce/Build*: combine or use resources or items to build new ones;
- *Communicate*: textual or verbal communication;

- *Character*: a token assigned to a player (i.e. an animate/inanimate object), which performs one or more type of action(s) (can be different characters over time);
- *Emotional connection*: link to a specific character;
- *Personalise*: how much the appearance and abilities of a character can be personalised;

Challenges:

- *Time pressure/Reaction*: fast reaction is needed or something has to be achieved in limited time;
- *Limited resources*: limited lives, energy or resources that are needed to perform actions or produce something;
- *Opponents*: objects/characters with (limited) intelligence that attack a player or play against her;
- *Obstacles*: something making it harder to achieve the goal or to reach it in a simple way (e.g. walls or gravity);
- *Puzzle*: skills (e.g. logical thinking, pattern recognition) are required to succeed or move on in a game;
- *Savability*: the ability to save the game status at any time or not at all;

Goals:

- *Static*: the goal remains the same within a quest, task, or mission (e.g., reach the end of a level) or even throughout the whole game;
- *Various*: different goals for different missions, tasks, or quests with a degree of freedom for a player to choose which ones she aims for;
- *Self-Defined*: players set their own goals;

Rewards (or Penalties):

- Access to *new levels* / regions of the map or game world (or lose time);
- *Gain/Lose power-ups* (e.g. buffs), improvements, access to new items (or lose power-ups);
- *Gain/Lose resources* such as money, lives (something a player needs or uses in a game);
- *Praise* through different visual effects, sounds and/or text;

Setting/World: Visual appearances of the objects and environment in a game along the following three dimensions:

- *Abstract*: not a world, but a grid or board
- *Realistic*: similar to the real world
- *Fantasy/Fictional*: not real

Further attributes specifying the game environment:

- *Dimension*: 2D (a player can only interact inside the screen plane) or 3D (interaction towards or away from a player is possible)
- *Choice*: a player can choose the game environment out of a selection of pre-defined maps/levels/environments;
- *Graphical detail*: level of detail of the graphics and animations of a game;

- *Explore*: how much a player can explore in the game environment, depending on how large and rich it is;

- *Sound Effects*: Snippets of sound that are played on certain events, to emphasize visuals or adding non-visual information
- *Music*: mostly melodic accompaniment, but may be influenced by or influencing the gameplay

Perspective: Perspective is the view that the player has on the game. This can be *bound to a character* (from the character's perspective or following it), *freely movable* through the whole game world/environment (which is accessible at the time) or *fixed* (does not move, but might change between different positions).

Structure: The game can be broken down into *separate parts*, where the player finishes a part (level or map) and moves on to the next one or *continuously progressing*. The game can have a *story*, which may be of different importance for the game as it might be a small background story or one that guides the player through the whole game. The game can have a different *length*, possibly even endless if it does not have a specified end. Although every game can be restarted and played again, some are designed to be played a lot of times, enabling a player to improve on performance or to make new experiences, while others are likely to be played only a small number of times.

3.5 Application of GEAM

Although GEAM was developed with the specific objective of further supporting the research on educational potential and design options of DEGs, it may also be useful for application beyond its original purpose.

Gamification: Deterding et al. (2011) define gamification as "the use of game design elements in non-game context". GEAM provides a structured overview on game elements and attributes that may be considered to gamify a non-game context. From the literature search for game attributes, 22 articles originated in the area of gamification of which 10 included a list of features. About 80% of these features were related to the element Reward. Only two articles (Ferro et al., 2013; Villagrasa and Duran, 2013) give a more diverse list of game attributes for gamification, for example

Villagrasa and Duran (2013) suggest spinning a story around the tasks that students have to solve in a class, and utilizing concepts like collaboration and quests (e.g. let students give a collaborative presentation, which is somewhat questionable as giving a presentation seems to be a non-game context, too). GEAM can help improve the understanding of gamification. The elements in GEAM are not unique to games (Huotari and Hamari, 2012). Thus the model can be used to analyse the non-game context and identify elements that are already present. For example, if the non-game context is a computer application, it provides Actions and a representation (similar to a Game World) and if it is a lecture it provides Challenges, Goals, and a Structure. This knowledge is a starting point for choosing which game features to use in the given context. For the computer application (e.g. a project management software), including more Action attributes may clash with or distract from the existing Actions. Considering attributes from other elements seems more promising, especially from the related element chain Challenge, Goal, and Rewards. Hence, not only the selection of attributes but also the structure and relations given by GEAM can support gamification.

Game Design: GEAM illustrates which elements a game consists of and how they work together to generate the game experience. It provides game designers with a general understanding of games, but even more importantly with a structure. Games can be highly complex and exist in a large variety. A fundamental structure helps to organize and keep track of all the details. It may assist experienced designers in sharing their knowledge as well as provide guidance for novice designers in accessing this knowledge. Furthermore, GEAM can support the design process, e.g. by serving as a kind of checklist while giving an overview of the options available for each element. The listed attributes may also inspire designers to think of new attributes or new ways to combine them and thus promote innovation. Finally GEAM provides designers with a vocabulary that facilitates communication within a project by ensuring a shared understanding among the team.

Game Research: GEAM gathers knowledge from multiple sources on what games consist of and presents this information in a coherent model, facilitating the analysis and comparison of games. By breaking down the core concepts of games (elements) into the choices they entail (attributes) the elements are operationalized and can be measured to evaluate games. This enables and fosters all kinds of game research, such as the classification of games and the refinement of game genre or the study of violence in games with all the attributes treated as variables that may influence the emotional state of the player. The model also helps categorize game research by allowing a more exact description of which aspect of a game is studied.

3.6 Summary

The Game Elements-Attributes Model (GEAM) presented is foundational for further research on design options and comparison of DEGs in terms of their educational potential. When designing a DEG, attributes for all game elements need to be chosen. At the same time the learning content needs to be included. As explained by Garris et al.'s (2002) input-process-outcome model, features of games are paired with instructional content, which induces learning. With the listed attributes, the GEAM provides insights into possible options for this pairing. Different combinations of attributes need to be considered to identify how to best include a given learning content into a game to build a DEG. The GEAM supports this identification as well as furthers research on comparing the effectiveness of different options, which may eventually lead to general guidelines on how to approach the pairing of content and game (attributes).

Chapter 4

Game Genre Map: Classification of Digital Games

The Game Elements-Attributes Model (GEAM) presented in Section 3.4 can serve as a basis to identify differences between games. However, games should first be grouped into game types before comparing their educational potential. There are two main reasons for considering the game type. First of all Prensky (2007) stated that "... just having a list of elements does not guarantee you a good game ..." and therefore suggested "... to look at the games and game genre that are out there that work well ..." (Prensky, 2007, p.152). A game genre or type describes a concept which proved to be functioning. So not all attributes from the model may be combinable and the choice of attributes depends (at least to some extend) on the overall choice of a game type. This leads to the second reason for why game types are relevant for this research. If there are common sets of attributes forming types of games, it allows to compare groups of games instead of individual games. The GEAM can serve as basis for the classification and definition of game types.

4.1 Classification of Games

Classification of games is a frequently discussed topic in game research (Aarseth et al., 2003; Apperley, 2006; Wolf, 2001). In this section we present an overview of existing classification systems. First of all, it is necessary to differentiate four related terms: 'game type', 'game genre', 'game class', and 'game classification'.

4.1.1 Clarifying terms

Several attempts have been undertaken with the aim of differentiating the aforementioned terms. Grace (2012) distinguished between 'game genre' and 'game type' by defining the genre as "style of game play" (e.g. Science Fiction) and the type as "the mechanics of gameplay" (e.g. first-person shooter). Jantke (2006) described 'game genre' as artistic design, and 'game type' as architecture and functionality. In addition, Jantke assigned the term 'game class' to the description of player behaviour and experience.

However, in practice the terms 'game type' and 'game genre' seem not to be used as defined, for instance, Grace (2005) pointed out that they are sometimes used interchangeably in industry. To identify the prevailing definition of 'game type' in the literature, we have analysed a selection of papers ('game genre' literature will be discussed in the subsequent section). ACM Digital Library (with the option "Publications from ACM and Affiliated Organizations"), given its broad coverage of publications on a diversity of research topics (including games) from different channels, has been searched for the term 'game type', resulting in 200 hits (as of October 2014).

Half of the papers were regarded as irrelevant for the following reasons: Three papers were not accessible; 21 discussed some unrelated topics, mostly on Game Theory (strategic decision-making); five papers used both words 'game' and 'type', but with no relation; 17 of the papers contained the term 'game-type' instead of 'game type' to refer to something as being a game or game-like; 54 did not provide enough information on the meaning of the term. The remaining papers were analysed for the use of the term 'game type' as well as its relation to similar terms.

The main findings from the remaining 100 papers[1] are:

- Game types define groups of games (only one source McEwan et al. (2012) is contradicting where individual games such as chess and backgammon instead of groups were listed as examples of game types).
- Game types differentiate games based on one or more features (e.g., cooperative vs. competitive games (Rawn and Brodbeck, 2008), or keyboard-controlled vs. gaze-controlled games (Krejtz et al., 2014)).
- The term 'game type' is often used interchangeably with 'game genre' and 'game class'.

[1] Among these papers, only those which are cited as examples are listed in References.

- 'Game type' is sometimes described as if it is subsumed by game genre hierarchically (e.g. "Capture the Flag game type in the first-person shooter" (Dominguez et al., 2011))

Hence, game type is a generic term referring to the division of games into groups based on one or multiple features. Accordingly, the term can be used at a high level, distinguishing between only two groups of games, separated by a single feature, but it can also be a surrogate for 'game genre' or used at a low level as sub-genre. A combination of different game types constitutes a game classification.

4.1.2 Game genre

The most common way of classifying digital games is to divide them into game genres. A variety of genre collections can be found in literature (e.g. Bates, 2004; Lecky-Thompson, 2007; McCann, 2009; Rollings and Adams, 2003). Wolf (2001) even identifies a list of 42 video game genres. A genre is usually defined with a simple description and some sample games. Despite its widespread uses, we have identified several issues with the game genre approach:

- *Genres are not clearly or consistently defined.* One example is Role-play game: While characters and stories are generally considered as important components of this game genre, emphasis is different: Rollings and Adams (2003) described story as more important than character development, and it is vice-versa for Fullerton (2008). Another example is strategy games that require careful thinking and planning (McCann, 2009). However, it can be argued that such requirements are also applicable for puzzle or adventure games.
- *The relation between genres is unknown.* Genres are defined individually and attributes used to describe one genre are not mentioned in the description of another. It is unclear how much genres differ and if they overlap. As sometimes a mixture of genres, e.g. Action-adventure is acknowledged as genre on its own, there is a strong indication that overlaps exist and verifying relations would help to understand why some games seem to fit in multiple genres. Especially when categorised by sellers, games are often placed in multiple genres, e.g. in online-stores, to increase the likelihood of being found by the user.
- *Definitions are based on completely different aspects.* This is another issue arising from individual and unrelated definitions of game genres. For example, educational games are sometimes listed as genre (e.g. Wolf, 2001; Wikipedia;

Amazon) while at the same time basically any type of game designed with educational value can be an educational game. So for this genre the defining attribute is its educational value, while for others it may be the dominant action in the game, or the camera perspective (e.g. First-person shooter).

- *Different sources use different sets of genres.* Seven game genres - Action, Strategy, Role-play, Adventure, Puzzle, Sports, and Simulation - are commonly referenced in different sources (e.g. Fullerton, 2008; McCann, 2009; Rollings and Adams, 2003; Wolf, 2001, Wikipedia; Amazon; metacritic.com - a game rating website), albeit not unanimously (e.g., Puzzle is not included in Fullerton or McCann; Simulation is refined as 'Vehicle simulation' in Rollings and Adams and 'Flight & other simulations' in Fullerton; see Table A.4). There also exist many less common game genres with some being named differently (e.g. Music vs. Rhythm & dance) and some being separated or merged (e.g. Racing & driving; Board, card & casino).

4.1.3 Other game classifications

Several authors have expressed a similar concern that the commonly known game genres lack clarity as well as consistency and therefore proposed alternatives. While some are not well justified (e.g. Na, 2006), three approaches give valuable input.

Aarseth et al. (2003) introduced an 'open' model to differentiate games along 15 dimensions with associated values, which are meta-categorized into five groups (i.e., space, time, player structure, control, and rules). This initial model was revised to include two more dimensions, resulting in altogether 17 dimensions grouped into eight meta-categories (Elverdam and Aarseth, 2007). With minor amendments the model was used by "three game players with a fair amount of diverse game experience" for categorizing 100 games (Dahlskog et al., 2009). Using cluster analysis, four game genres were identified: strategy, first-person shooter, progression & exploration, and perfect information. While basing the classification on a fixed list of differentiating attributes could overcome the issues of the existing classification by game genres, there are still some concerns in the chosen approach. First, the list of attributes was compiled by comparing different games, but it is unclear how comprehensive it is, i.e. if all basic elements of games are considered. Second, the research leading to the identified genre was done by a small team of three experienced players, which could increase the risk of biased results. Thirdly, for the categorisation it was considered that some games have different gameplay modes (e.g. multi- and single-player), but not that the playing style of players can differ (e.g. in “World of Warcraft” one player can be interested in the story, while another player only tries to level up as fast as possible without following the story). A game may be categorised differently depending on which attributes are deemed important by a player, not just by which attributes it contains. Finally the resulting genres are barely defined and thus not convincing. For instance, the first-person shooter is described by controlling an avatar and a vagrant camera position, but this should also be applicable to some progression & exploration games, which are defined as exploring a story, character, or world (e.g. “Skyrim” should fit in both genres). The relations between game genres thus remain unclear.

Lewis et al. (2007) conducted a more conclusive study, where 124 participants were asked to compare pair-wise a set of ten games by rating their similarity. Using multidimensional scaling (MDS) methods, they identified game clusters. While the approach used in Lewis et al. (2007) is methodologically stronger than that in Dahlskog et al. (2009) where participants were bound to some predefined dimensions and values, a major drawback of such flexibility is that variables had to be derived from clusters subsequently, which mostly is a subjective interpretive process. For game comparisons, it is critical to have a robust set of underlying variables for defining a game and to illustrate how it is similar or different from other games. To analyse the influence of a specific game type on learning efficacy of DEGs a well-defined set of variables is deemed necessary.

Another intriguing approach to game classification is to group them based on rules - a key element of a game. By analysing the rules of 588 video games, Djaouti et al. (2007, 2008) extracted a set of ten elementary rules describing the gameplay, which were coined as “game bricks”. Then Djaouti et al. (2011) set up a website to invite people to classify games based on game bricks and some other aspects such as purpose, scope, and market. Although the game-brick-based classification approach is systematic, it is inadequate for game comparisons. An obvious drawback is that it tends to reduce game characteristics to a bare minimum as basic rules, thereby missing other critical elements such as story.

4.1.4 Definition of terms

Besides identifying issues in the current game classification, we propose the following definitions based on our findings:

- *game types = groups of games, differentiated based on one or more game features (game class = less often used synonym for game type)*

- *game classification = set of criteria by which games are differentiated*
- *game genres = a subset of game types; in combination game genres form a popular game classification*

4.2 Revising Game Classification

Given the issues discussed above, we aim to analyse as well as refine the definitions of existing game genres. However, we do not intend to introduce yet another set of terms to aggravate the confusion. If supported by the empirical findings, we aim to retain the names of the genres that are already widely adopted by the game community. The main drawback with the current game genres is that each genre is defined individually. This makes it difficult to compare and identify relations between genres. Being aware of which and how genres overlap is necessary to find a set of genres which represent the full range of different games while achieving maximal separation.

To attain comparability, game types need to be all defined based on the same set of attributes, similar to the approach used in Aarseth et al. (2003). Unlike Aarseth et al., we argued that the attributes must be included in a comprehensive game model to ensure that all the game elements are considered and concomitantly developed our Game Elements-Attributes Model (GEAM) (Figure 3.1).

Analysing the attributes of different games using the GEAM, similar games can then be clustered with the approach as described in Dahlskog et al. (2009). Like Lewis et al. (2007), we planned to collect empirical data with a survey from a large group of gamers rather than a small group of experienced gamers, thereby enabling us to gain a better understanding of how games are played. We also used the multidimensional scaling methods (Lewis et al., 2007), but not on the individual game level, but the clustered games level, to gain further understanding on the relation of game types. Furthermore, we compared how the derived game types match the gamers' understanding of existing game genres.

In summary, we have identified four requirements for a new or refined game classification that underpins the design of our survey. The four requirements are: (i) To provide a comprehensive set of attributes by associating them with individual elements of a game model, which represents the basic architecture of digital games; (ii) To compile consistent definitions of game types, allowing recognition of relations and comparison of games, by building a new classification upon a set of general game attributes; (iii) To examine whether and how the existing game genres can be mapped onto the new ones; (iv) To take into account the assumption that games are normally played and perceived in various ways by different players.

As the ultimate goal of our planned future research study is to compare the learning quality of different game types for learning specific educational content, it is essential to analyse how games differ at the componential level. Hence, we have developed our model GEAM (Chapter 3) of which two critical components are elements and attributes. As a result of examining game definitions, we identified the core elements of games: player, input/output, actions, challenges, goals, rewards/penalties, game world, perspective, structure. Based on an extensive literature review we then compiled a list of attributes for each game element, representing main options on how these elements can be realised in a game (e.g. the challenges in a game can be puzzles, opponents, time pressure etc.). The main advantage of the GEAM, in comparison to Aarseth et al. (2003), is that it depicts the relations between game attributes and elements, contributing to the GEAM's higher comprehensiveness.

4.3 Empirical Study

The main goal of our "Game Classification Survey" (GCS) is to collect data from gamers how they play and perceive video games in order to enhance our understanding of existing game genres. To enhance the readability of the ensuing text, the following conventions are used: A game genre is written in small caps (e.g. ACTION); an element is capitalised (e.g. Actions); an attribute is italicised (e.g. *move*); a game is in quotes (e.g. "WOW", which stands for World of Warcraft).

4.3.1 Gamer survey design

Grounded in the GEAM, we developed the GCS to derive from the survey data clusters of games sharing a similar set of attributes. Each respondent was asked to select a game from the list given or nominate a game which she knew well. Then she was asked to evaluate the game with respect to individual attributes of each of the elements, using a slider with a range of integers from 0 to 100 for 27 of the 29 continuous variables and fill-in-the-blank for the other two, or using a multiple-choice format for categorical variables. Finally, the respondent was required to assign the game to one (or more) of the seven common genres; she could also provide alternative classification terms.

The rationale of providing the respondents with a list of games to choose (though they were still free to nominate one) was to increase the probability that a game would be selected and rated multiple times, enabling us to perform the game clustering. Our critical task was then to identify a list of representative games, which were selected based on sales figures with the intention to include the most popular ones.

We used the figures posted on the website VGChartz (2012) which is generally considered rather accurate for games available in the market. We also included two free and very popular games. For a balanced distribution, five games were selected for each genre (Table 4.1). We were well aware of the issue of "multiple categories" or ambiguity in this website (e.g. "Angry Birds" can be found under ACTION, PUZZLE, and STRATEGY), and this is exactly part of the issue we aimed to investigate.

Genre	**Games**
Action	*GTA, Angry Birds (AB), Tekken (T), Call of Duty (CoD), Super Mario (SM)*
Adventure	*Monkey Island (MI), Myst (M), Tomb Raider (TR), Prof. Layton (PL), Legend of Zelda (LoZ)*
Puzzle	*Tetris (T), Pac-Man (PM), Dr. Kawashima, Solitaire* (S), Portal (P)*
Role-play	*Pokemon (P), Final Fantansy (FF), The Elder Scrolls (ES), World of Warcraft (WoW), Diablo (D)*
Simulation	*Nintendogs, Sims (S), Microsoft Flight Simulator (MFS), FarmVille* (FV), Guitar Hero (GH)*
Sports	*Wii Sports (WS), FIFA Soccer (FS), Tony Hawk's (TH), Just Dance (JD), Gran Turismo (GT)*
Strategy	*Warcraft (Wc), Warzone2001 (Wz), Command & Conquer (C&C), Worms (W), Anno (A)*

Table 4.1: Preselected games based on sales figures and genre allocation from VGChartz (2012); *free, highly popular ones

The questions of GCS were structured and formulated with reference to the 9 elements and 42 attributes (29 continuous; 13 categorical) included in GEAM. To get an overview, an abbreviate version of the questions for measuring the continuous variables is presented in Table 4.2 and for the categorical variables in Table 4.3. The complete GCS can be found in Table A.5.

Player
How would you rate the importance of the following aspects? (*Competition, Cooperation, Communication, Playing alone*)
Actions
How would you rate the importance of the following action(s) in the game? (*Move, Place/Position, Fight/Destroy, Collect, Produce/Build, Communicate*)
Please rate the importance of the following aspects of the main character (*Emotional connection to the character; Personalise the character*)
Challenges
How would you rate the importance of the following challenges in the game? (*Time pressure, Limited resources, Opponents, Obstacles/Forces, Puzzle, Quiz, Search/find, Savability*)
Goals
How *free* are you to *choose* which goals (missions, tasks or quests) you want to complete?
World/Setting
How would you rate the *level of detail of the graphics* of the game?
How much can you *explore* in the game?
How would you describe the visual appearance of the objects and the environment in the game? - *Abstract* (not a world, but a grid or board) - *Realistic* (similar to real world) - *Fantasy/fictional* (not real)
Structure
How important is the *story* for the game?
*How *many times* did you replay the game? [multiple times]
**How long* did it take you to play the game once (pure playtime)? Please state the duration as a range between an approximate minimum and maximum in hours. [duration/range]

Table 4.2: The questions for the continuous variables (in italics) * items are measured with numeric input

Player
When you are playing the game, are there other players in the game with you? *No; can be; always*
Interaction
On which device do you most often or prefer to play the game? *Computer / Laptop; Console; Handheld Console; Mobile Phone, Tablet*
What input device(s) do you usually use to play the game on the [chosen device]? *Mouse; Keyboard; Touchpad; Gamepad; Joystick; Touchscreen; Special Controller; Tracking / no Controller; Motion Sensor; Keys / Buttons*
Actions
How do you move objects or characters through the game? to *restricted positions*; in *one direction*; *freely*
Do you mainly control one character in the game, which performs the actions you chose above? *Yes; No, but one character I have an emotional connection to; No*
Goals
How'd you describe the goal(s) that you try to reach in the game? *Static; Various; Self-defined*
Rewards/Penalties
What rewards do you get for reaching a goal in the game? *Access to new levels; Power-ups; High score; Resources; Praise*
How are mistakes made during the gameplay penalised? *Losing time; Losing power-ups; Lower score; Losing resources; No penalty*
World/Setting
What is the dimension of the game environment? *2D; 3D*
Are you free to choose the environment you want to play in? *Yes, out of a selection of maps / levels / environments; No*
Perspective
What view do you have in the game? *bound to a character; freely movable; fixed*
Structure
How is the game structured? The game is broken down into *separate parts*; The game does not have separate parts, but is *continuously progressing*
Did you ever finish the game? *Yes, I finished at least once; No,* {different reasons to choose}

Table 4.3: The questions for the categorical variables

4.3.2 Participants

The invitation to participate in the survey was disseminated to several mailing lists such as chi-web@acm.org over a period of three months. We have received 560 responses of which 321 were complete and valid for analysis. The age range of the respondents was from 15 to above 51 years old (3%) with the largest group between 21 and 25 (36%); 70% were male. Respondents were from 24 countries worldwide with most of them (73%) residing in Europe. Their game experience varied with 20% being very experienced players who know many different games and play games almost every day; 57% were university students; 35% employees and the rest were high school students, employers or others.

4.4 Results

Individual respondents identified a game they were going to analyse. Some games were repeatedly mentioned; consolidating all the games chosen results in 67 different games with 33 of them from the given list (Table 4.1, except "Dr. Kawashima" and "Nintendogs" which were never chosen). Figure 4.1 shows the frequencies of the games chosen per existing genres and the 'open' one (i.e. for the games nominated by the respondents). The list of games selected in the 'open' category can be found in the Appendix, Section A.3.1.

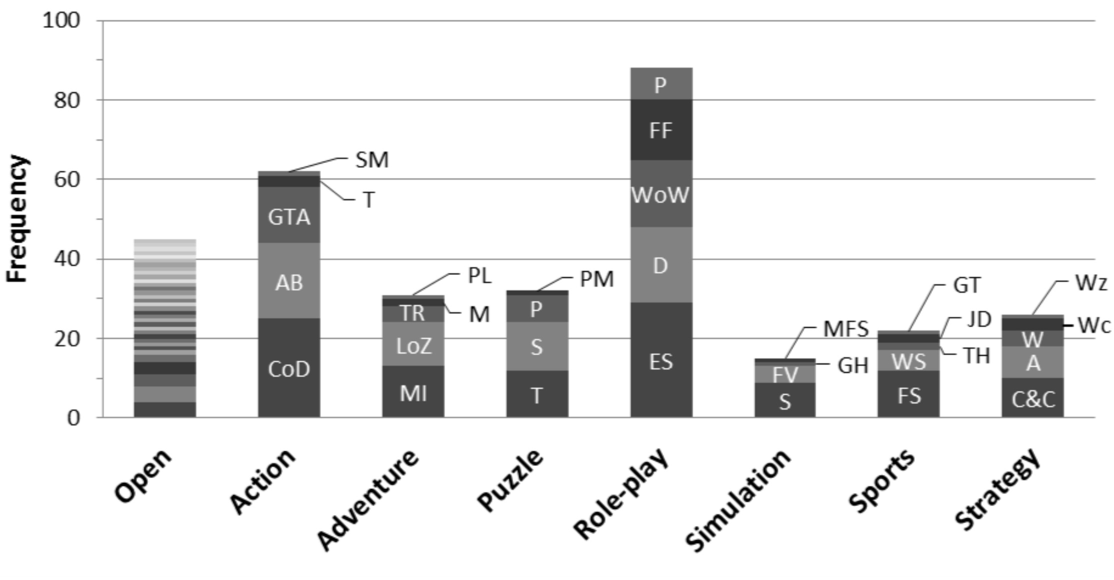


Figure 4.1: Number of games per given genre and the 'open' one (nominated by respondents); abbreviations refer to games in Table 4.1

4.4.1 Hierarchical clustering

To group the games systematically, cluster analysis was employed. Among different established clustering algorithms, we applied hierarchical clustering (HC) for several reasons: HC does not require a predefined number of clusters, the dendrogram (Figure A.1) as part of the analysis outcomes, can visualize well the preliminary grouping of the games, and HC was adopted in the related work (Dahlskog et al., 2009). Nonetheless, it is generally recognized (Mooi and Sarstedt, 2011) that clustering does not give highly precise and consistent results, as different algorithms tend to produce somewhat different clusters. Hence, cluster results need to be validated and interpreted based on an expert's understanding of the topic. As HC that can process mixed data efficiently is yet to be established, only the 29 continuous variables (Table 4.2) were taken into account as they cover the more important aspects such as gameplay. Normalisation of the data was necessary for the two variables with open numeric input.

With agglomerative HC, cases are grouped together based on a similarity matrix with either distance or correlation for each pair of cases. To validate the consistency of the clustering results, different distance measurement methods can be used (Mooi and Sarstedt, 2011). As some methods are (too) sensitive to outliers such as single and complete linkage and some such as Ward's method tend to build equally sized clusters (which is irrelevant to our work), we opted for the more robust methods listed in Table 4.4. The metrics were computed using SPSS v.19.0. The 1st method identified 14 clusters whereas the 2nd and 3rd found 16; the additional ones resulted from the splitting of two of the clusters. The average overlap rate (i.e. a cluster is identified by all three methods) was 61%, which we considered as a reasonable consistency level (Note: no recommended acceptance rate is given in the literature (Mooi and Sarstedt, 2011)). Hence, the 16 clusters were used for subsequent analyses.

ID	**Method**	**Metric**
1st	Average linkage between groups	Euclidean distance
2nd	Average linkage within groups	Euclidean distance
3rd	Centroid	Pearson correlation

Table 4.4: Hierarchical methods used to analyse the game data

4.4.2 Spatial maps with multi-dimensional scaling

To understand how the 16 clusters were related to each other, we created a two-dimensional (2D) map to visualize their similarity. Clusters sharing a similar set of attributes should be placed close to each other. The visual cues of proximity support genre identification. Whereas a genre can comprise more than one cluster, an overlap of clusters can result in a mixed genre. To calculate the distances between the clusters, ANOVA was conducted for the 29 continuous variables (i.e. game attributes). Results indicated that three attributes - *playing alone*, *quiz*, and *savability* - had low significance and were thus discarded. This finding on *savability* is contradicting that of Dahlskog et al. (2009), who argued that it is one of the most important attributes for game classification. Their claim was based on the observation that it is an outlier with low correlation to other game attributes, but we counter-argue that this does not necessarily imply its significance. They measured *savability* as a categorical variable whereas we measured it with a continuous scale, enabling us to use a more powerful method to analyse it.

The remaining 26 game attributes were used to evaluate the relations between the 16 clusters pair-wise. The number of significant differences for individual pairs was used to build a distance matrix. The process was applied for each of the three HC methods (Table 4.4). The distance matrices formed the database for multidimensional scaling (MDS), which is used to map MD data to a 2D representation. Lewis et al. (2007) applied the same method to analyse the similarities between games as directly perceived by their participants. However, such holistic perceptions cannot provide the accurate game classification information which we aimed to identify through our attribute-based approach.

We generated a map from the distance matrix by using the R cmdscale function (cmdscale, 2015). The three maps derived from the respective HC methods were superimposed for comparison. Each number on the map represents a cluster. For identification we assigned increasing numbers to the clusters, consistently across the three methods. Given the high similarities between maps, they were merged into one by using the mean values (Figure 4.2). This resulting map serves as the base for further analysis.

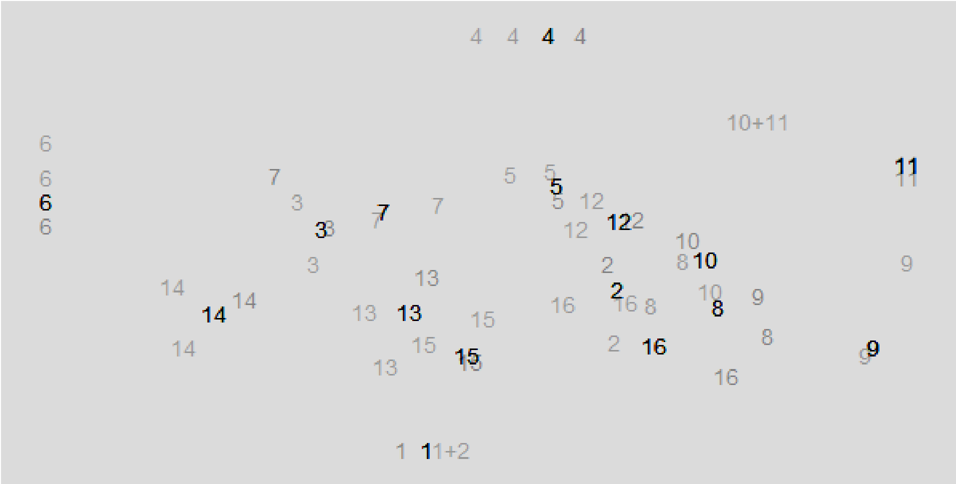


Figure 4.2: Mapping of the differences among the 16 game clusters with MDS; black numbers represent the mean over the respective values of the three methods (in grey)

4.4.3 Mapping to seven game genres

To enhance the comprehensibility of the map, each of the 16 clusters was designated by a leading game with the highest frequency of being selected by the respondents (Figure 4.3) (NB: Cluster 10 and 11 have the same leading game as they were a single cluster as identified by the 1st HC method, Table 4.4). To validate the common seven game genres, we matched them with the mean value map (Figure 4.2). In the final question of the GCS, the respondent was asked to assign the game of interest to one (or more) of the seven genres. These data allowed us to calculate the percentage of the games in each of the 16 clusters that were assigned to a genre. The areas with the percentage above 70%, which was the lowest maximum percentage among the 16 clusters, were identified. Encircling the clusters with the percentages above the threshold leads to an initial "genre map" (Figure 4.3).

The seven genres distribute rather neatly over the map with only one cluster (C) unallocated (C3: "Angry Birds") and one cluster double assigned (C5: "Legend of Zelda"), forming the known combined genre ACTION-ADVENTURE. The genres of the other leading games match well with those they are originally classified in VGChartz (2012). This finding suggests that there is a general shared understanding of which game belongs to which genre. Besides, the clusters falling under the same genre are rather close to each other (except PUZZLE and ACTION), implying that they could work as an overall clustering solution.

To refine the genre definitions and to optimise the genre selection, the distribution of the game attributes over the genre map was also analysed with a mapping technique

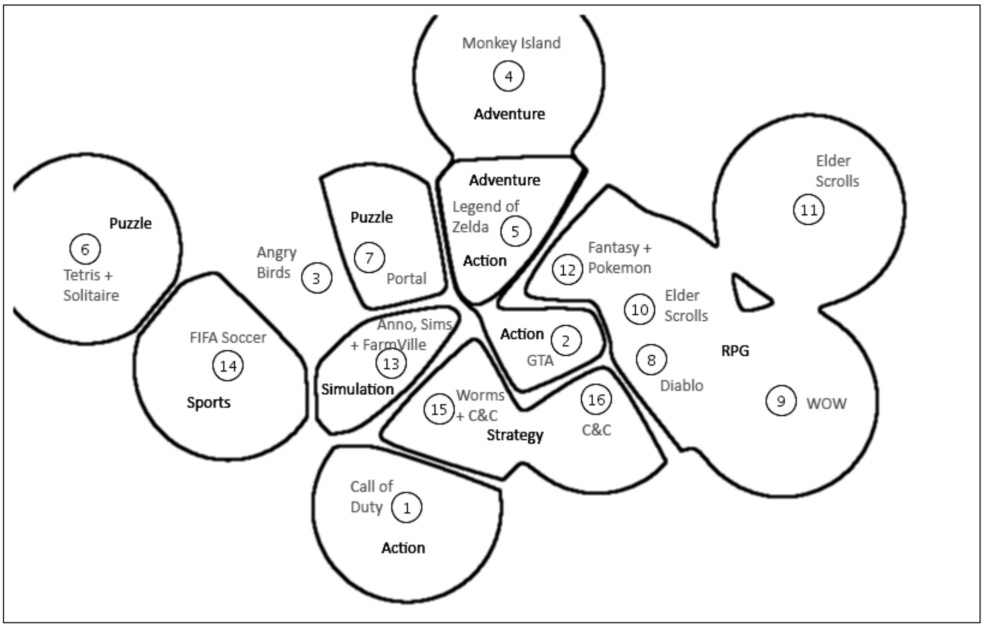


Figure 4.3: Genre map generated according to the percentages of the games in a cluster that were assigned to a known genre

described as follows: To deepen the understanding how the clusters are related in terms of the game attributes, we visualized the relationships by greying the area of a cluster; the stronger an attribute is related to a cluster, the darker the grey colour is (e.g. Figure 4.4). A colouring threshold is determined, using different approaches for the two different types of variable.

4.4.4 Mapping of game attributes as continuous variables

As described earlier, all the continuous variables (except the two with numeric input) are rated with the range of 0-100. For a cluster, if 75% of the ratings for an attribute are above 25 (out of 100, thus also 25% deviation), then the cluster is coloured for that attribute. The allowed deviation 25% is to accommodate the systematic bias of the rating behaviour as some respondents tend to rate higher or lower in general (Wherry and Bartlett, 1982). The Game Genre Map (Figure 4.3) can be seen as divided into five major areas, which are identified by the cluster (C) number and the genre: C6-Puzzle (on the left), C4-Adventure (at the top), C1-Action (at the bottom), C9&C11-Role-play (on the right), and the remaining in between (middle). This division could help refine game genres. We evaluated this idea by identifying the key game attributes of each area. In the following subsections we report the main findings.

Attributes of the Action genre (C1, bottom)

C1 (genre/leading game: Action/"Call of Duty") comprises six attributes of different elements: *fight, opponents, time pressure, realistic world, competition* and *cooperation.*

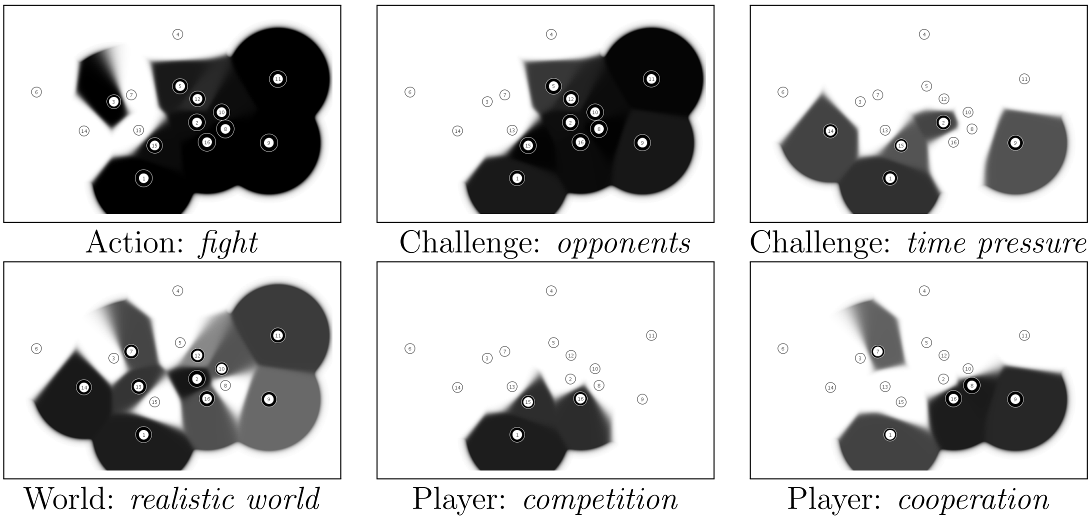
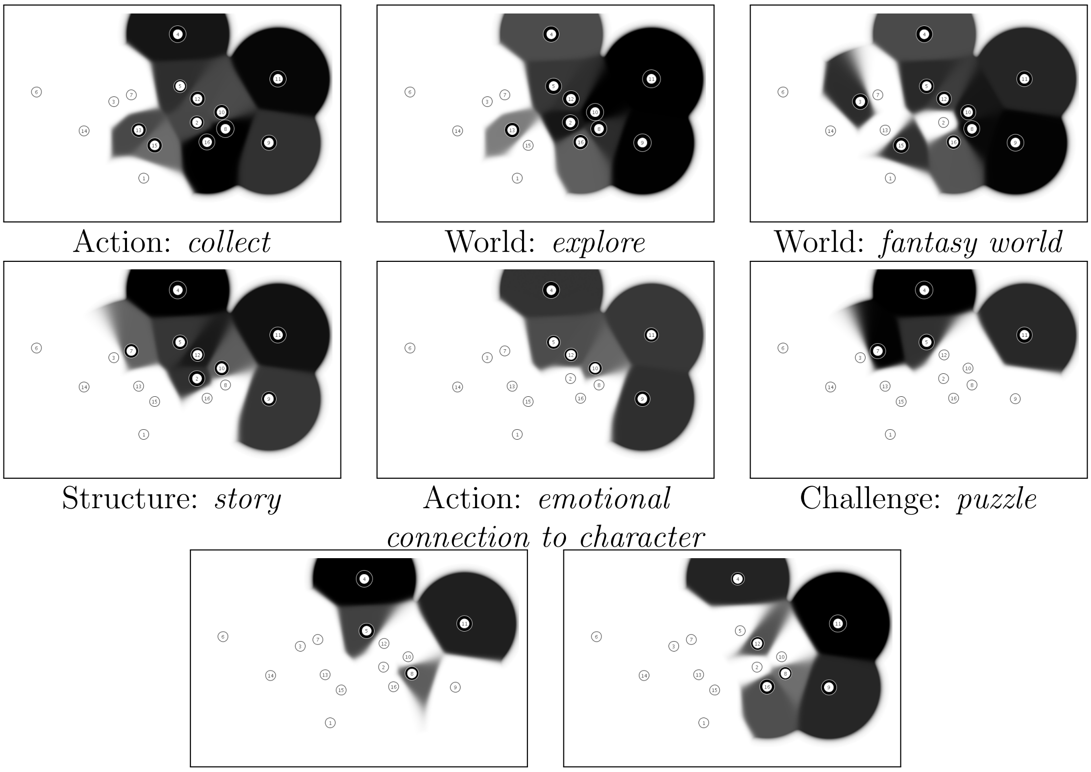


Figure 4.4: The distribution of the six game attributes of the Action genre (bottom area) over the map (NB: the label is formatted as Element: *attribute*)

The distribution of each of the attributes is depicted in Figure 4.4. Clearly, the attributes *fight* and *opponents* are not only highly relevant to the bottom area (Action) but also to the right one (Role-play). However, *opponents* is irrelevant to C3 with the leading game "Angry Birds" where a player attacks other characters that do not fight back and by definition are not opponents. The attributes *time pressure* and *realistic world* are rather oriented towards the left (Puzzle). A realistic representation of the game world is more relevant to the games in C14 (Sports/"FIFA Soccer"), C2 (Action/"GTA"), and C1 (Action/"Call of Duty"). These together with C15 (Strategy/"Worms" + "C&C (Command & Conquer)") and C9 (Role-play /"WOW") are strongly related to *time pressure.* The attributes *competition* and *cooperation* spread to the right.

Attributes of the Adventure genre (C4, top)

For C4 (Adventure/"Monkey Island") eight attributes were identified: *collect, communicate, emotional connection to character, puzzle, search, fantasy world, explore,* and *story* (Figure 4.5). This top area and its right counterpart (Role-play genre) are closely connected. Apart from the obvious top area, the attributes *collect, explore, fantasy world,* and *story* distribute rather densely in the right and middle



Action: *collect* — World: *explore* — World: *fantasy world*

Structure: *story* — Action: *emotional connection to character* — Challenge: *puzzle*

Challenge: *search* — Action: *communicate*

Figure 4.5: The distribution of the eight game attributes of the ADVENTURE genre (top area) over the map

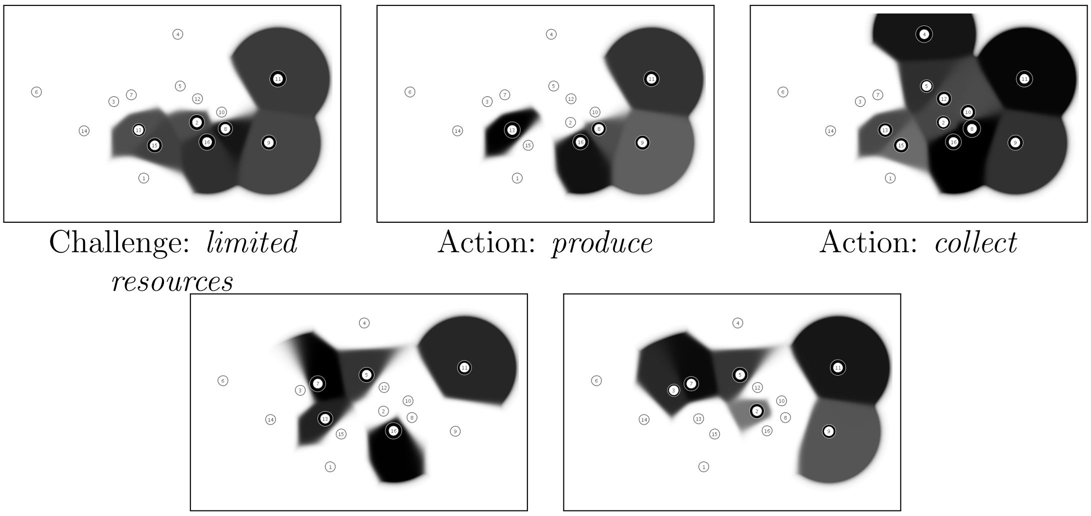
areas of the map. These and the attributes *emotion to character, puzzle,* and *search* cover C5 (ACTION-ADVENTURE/"Legend of Zelda"), which is located right below C4. Among them, *story* and *puzzle* are the only attributes that are related to C7 (PUZZLE/"Portal"), except for attributes not covering C4, like *obstacles.* The attribute *communicate* is located near C16 (STRATEGY/"C&C") and the two groups on the right.

The distribution patterns so identified are intuitive, because an ADVENTURE game typically involves exploring a fantasy world with a story and challenging the player with puzzles on the way of searching objects. The player might also develop emotional connection to the character(s) when undertaking the adventure.

RESOURCE attributes (eleven groups, middle)

Attributes primarily identified in the middle (eleven groups) but not in the other outer groups (except the groups on the right that comprise the majority of the attributes) are: *limited resources, produce, place,* and *obstacles* (Figure 4.6). Of particular interest is the attribute *collect*, which covers most of the middle and, together with *limited resources, produce* and *place*, defines C16 (STRATEGY/"C&C") and

C13 (SIMULATION/“Anno”, “The Sims” + “FarmVille”). The term ‘RESOURCE’ was chose as constituting the prevailing aspect of the area, e.g. collecting resources and producing as well as placing objects. It can well describe not only the more action-oriented game “C&C” where a military base is built and forces are provided with *limited resources* but also the non-violent game “FarmVille” where the resources are plants or animals grown in a farm.



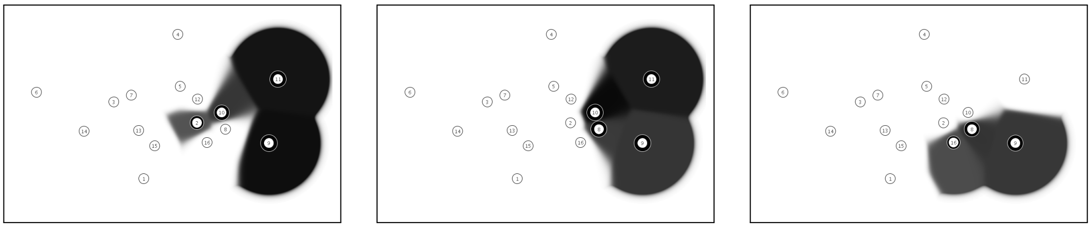
Challenge: *limited resources* — Action: *produce* — Action: *collect*

Action: *place* — Challenge: *obstacles*

Figure 4.6: The distribution of the salient attributes in the middle area consisting of eleven game groups

Attributes of the ROLE-PLAY genre (C9&C11, right)

C9 (ROLE-PLAY/“WOW”) and C11 (ROLE-PLAY/“Elder Scrolls”) are mostly co-defined by the attributes of the bottom, top and middle area (Figure 4.4, Figure 4.5, and Figure 4.6), making it an amalgam of ACTION, ADVENTURE, and RESOURCE games. In addition, the attributes *goal choice*, *personalize character*, and *player communication* are identified in this area (Figure 4.7). This lends further support to the observation that ROLE-PLAY games are inherently rich. These games normally provide players with a broad range of options including *goal choice* - the attribute shared with C2 (ACTION/“GTA”). The attribute *personalise character* of the element Actions can well epitomize ROLE-PLAY of which one of the focuses is configuring a character. *Communicating* with other players is an attribute shared with C16 (ACTION/“C&C”).



Goal: *goal choice* Action: *personalise character* Player: *communicate*

Figure 4.7: The distribution of three specific attributes of the Role-play (RPG) genre (right area) over the map

Attributes of the Puzzle genre (C6, left)

Incidentally, C6 (Puzzle/"Tetris & Solitaire") is barely related to any of the 29 attributes measured as continuous variables, but strongly related to some of those evaluated categorically such as Player: *single player*, Perspective: *fixed*, and Goal: *static*. The attributes - *time pressure* and *realistic* - salient in the bottom area are related to this left area to some extent. So are *obstacles, fantasy,* and *fight* from the top area. Other attributes like *puzzle* contribute to the area, but are not distinctive (75% above 25 rating). Apart from these, the only salient attribute identified for C6 is *abstract world*, indicating that the games thereof are primarily based on a grid or a shaped background rather than a realistic or fantasy world.

Universal attributes

Two attributes - *move* and *detail of graphics* - are related in (nearly) all the 16 clusters. As one defining characteristic of digital games is interactivity, the relevance of *move* to all the games is not surprising. Similarly, most digital games need some form of graphical representation (exceptions are purely text- or audio-based games), though to a different extent of detailedness. For instance, Puzzle games such as "Tetris" are implemented as abstract worlds with less demand for graphical detail. In contrast, Role-play and Action games typically require high graphical details.

4.4.5 Mapping of game attributes as categorical variables

As explained earlier, the categorical variables were not used for cluster analysis. This is not critical to the overall conclusion of our study as the continuous variables address the more important game elements. As the categorical variables can still provide relevant information to the design of educational games, we analysed the data using the same mapping technique as for the continuous variables but a different method to determine the colouring threshold. Specifically, for each group the percentage of the selected option for a categorical attribute was used to estimate the hue; the higher the percentage, the greyer the area is. For instance, for the element Goal, the attribute *goal type* has three possible values: static, various and self-defined. The respondents could select one (or more than one) of these options to define the game of interest.

The following elements comprise categorical variables: Player: *sociability*, Interaction: *I/O device*, Perspective: *view flexibility*, Goal: *goal type*, Rewards: *feedback format*, Structure: *continuity*, Setting/World: *dimensionality*, Action: *character*.

Goal variability

Generally speaking, games with *static* goals are less complex than those with *various* goals. Results show that the distribution of both goal types over the map is consistent with this assumption with less complex PUZZLE games on the left (static goals) and more complex ROLE-PLAY games on the right (various goals); *self-defined* goals are less common.

Rewards/penalties

Gaining rewards and avoiding penalties are typical game mechanics to sustain players' motivation (see Table 4.3 for the related questions). ADVENTURE games (the top area) tend *not to penalize* players but rather support them to progress steadily through a story. A simple but effective reward system is *high score*, which is often used in PUZZLE games in the left area. For ROLE-PLAY in the right area, rewards are more based on developing a game character through *power-ups*.

Structure

Most games are divided into *parts*, which may be linked through with a *storyline* or entirely self-containing. Most ROLE-PLAY games (C9&C11, "Elder Scrolls", right) and some SIMULATION games (e.g. C13, "The Sims", middle-left) offer open, explorative environments and support *continuous* progress. Some PUZZLE games (C6, "Tetris", left) are also continuous as they are too small to be split into parts. Games that most respondents claimed they could play through to the *end* are ADVENTURE games. This may be explained by the motivating nature of the story of these games.

4.5 Discussions and Implications

We summarize the results by creating an overview Game Genre Map, which depicts how the 29 continuous game attributes are related to the 16 clusters (Figure 4.8). Excluding the relatively less critical categorical attributes from the game map is to contain its complexity. The shades of grey on the map indicated how many of the attributes are relevant to individual game clusters; the darker the area, the more attributes it comprises. There is clearly a gradient from light grey on the left to dark grey on the right, corresponding to the richness of the games.

4.5.1 Game Genre Map

The arrows spanning across the map from left to right are three main routes to game enrichment. Here we define the richness of a game in terms of the number of attributes it implements. The first route, which is close to the top area of ADVENTURE games, contains two enriching elements: the two attributes of the element Challenges - *puzzle* and *obstacles* - increase the demand of the cognitive ability or dexterity of players, and the attribute *story* can elaborate the structure of a game by expanding its possible courses. This route may be suitable for educational games that are rather cognitive demanding and of a slower pace. The second route, near the bottom area of ACTION games, also contains two enriching elements: *time pressure* of Challenges and *fight* of Action. Fighting can be quite complex with weapons and opponents. Educational games that are based on training, reaction or physical activity can fit this route. The third route is above the second and more concerned with the RESOURCE games (i.e., SIMULATION and STRATEGY). As an attribute of Challenges, resources imply collecting or managing objects which can be instantiated as educational content. All the three routes can lead to the richest game genre: ROLE-PLAY games, which, by combining the elements of the ADVENTURE, ACTION, and RESOURCE genres, offer broad (i.e. the range of activities) as well as deep (i.e. the granularity of activities) game worlds for players to explore.

When selecting a game genre for a DEG, the designer needs to consider a crucial question: "Which game attributes can be used for delivering specific learning content to specific target groups to attain specific learning outcomes?" Clearly, a host of

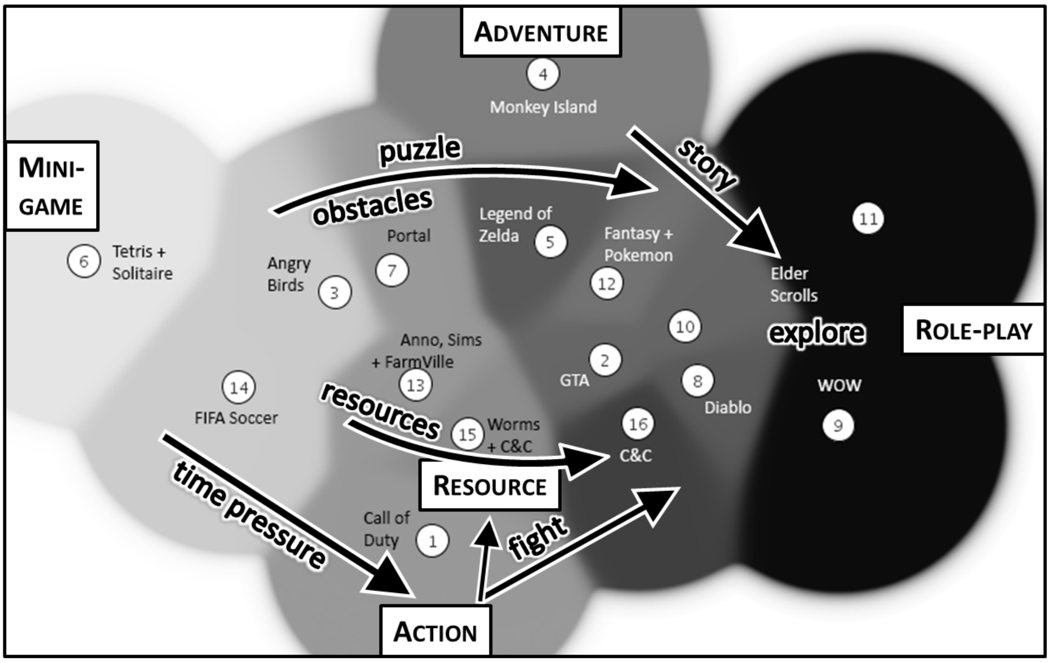


Figure 4.8: An overview of mapping all the continuous game attributes to the sixteen game clusters

pedagogical, technical, social, and organizational issues come into play if we aim to find out not only which attributes to use but also how to use them or even how effective of using them. Addressing all these questions is beyond the scope of the current study whose results enable us to answer the critical 'which' question.

Using games as multifaceted as Role-play for educational purposes increases not only the implementation effort but also the risk of losing focus on a learning topic. If the learning content is integrated in only one of the many attributes of a rich game, the player may be exposed to less content within a short duration. This probably undermines the learning efficacy of the game. The implications drawn from these arguments are to adopt simpler games such as Mini-game (cf. Frazer et al., 2007) as a means for educational ends and to evaluate which attributes can be most suitable for the given educational content by considering the three routes depicted on the Game Genre Map.

In fact, we have renamed the left cluster as Mini-game instead of Puzzle as most games in C6 are short, self-contained, and built on a single principle (Frazer et al., 2007), which is not necessarily to solve puzzle. Frazer et al. (2007), based on their analysis of 30 DEGs, argued that individual Mini-games tend to be too shallow to attain learning goals and proposed building a compendium of MINI-GAMES as a DEG with each of them having different game mechanics. This proposal lends support to our idea of the three routes to game enrichment.

4.5.2 Improving game genres

We have adapted three of the seven common genres (derived earlier from the different sets of game genres found in literature) redefined two and dissolved the other two into the other genres. The two discarded genres are SPORTS and SIMULATION. Both describe the content of a game rather than its game mechanics. SPORTS can be re-categorized as ACTION games with sports content. In the game map, they are located close to the ACTION area. Similarly, SIMULATION describes only one aspect of the game: it simulates a situation or a topic. On the map, it is located in the RESOURCE area. This suits many SIMULATION games that are based on real-life processes such as simulating a person's life (e.g. "The Sims") with resources such as food, furniture and other objects or simulating a farm with growing plants (e.g. "FarmVille"). Other SIMULATION games like driving or flight match better with ACTION games. Nonetheless, given the highly diverse and thus non-discriminative nature of this genre, SIMULATION seems not particularly useful and should be discarded. The genre PUZZLE has been renamed based on the reason already mentioned, but may function as a substitute for MINI-GAMES, as it is still a more commonly known term. The same applies for the STRATEGY genre, which has been renamed RESOURCE, based on the defining attribute in this area.

4.6 Summary

In summary, our results enable us to conclude with five game genres - MINI-GAMES, ACTION, ADVENTURE, ROLE-PLAY, and RESOURCE and their defining attributes (Table 4.5). In the Game Genre Map derived from the survey data (Figure 4.8), the clusters with largest distances (top, bottom, left and right) have the more distinct sets of attributes, and another set was identified in the middle. Mixtures of genre are possible if a game is located between two genres on the Game Genre Map. Its attributes should then be a mixture of two (or possibly more) neighbouring genres. If it was possible to clearly separate game genres, these genres should have emerged as clusters. The results however indicate that there is always an overlap between game genres, which probably matches the experience of game designers and researchers. The advantage of the Game Genre Map is that this overlap is clearly visible and kept to a minimum by choosing the points furthest apart as the centres of the new genres. By locating a game along the four cardinal directions of the map, its complexity and game type can be derived. By identifying its defining attributes, the game nature can be better understood.

Element	Mini-game	Action	Adventure	Role-play	Resource
Player	single player	(multi player), (competition), (cooperation)	single player	(multi player), (cooperation), (communication)	(multi player), (competition), (cooperation), (communication)
I/O	(mobile device)		(PC)	(PC)	PC
Actions	move: (restricted), (fight), (character: none)	fight, move: free, character	collect, move: (free), (communicate), character: emotion	fight, move: free, collect, communicate, (produce), character: emotion + personalise	(produce), collect, (place), move: free, (communicate), (fight), (character: non)
Challenges	(obstacles), (time pressure)	opponents, time pressure	puzzle, search	opponents, (search), limited resources, (obstacles)	limited resources, (opponents), (time pressure)
Goals	static	static, (various)	(static), (various)	choice, various, (self-defined)	(static), (various), (self-defined)
Rewards	score	(score)	(no penalties)	power-ups	(score)
Setting	abstract, (realistic), (fantasy), 2D, (3D), pre-defined	realistic, graphical detail, 3D, (selectable)	explore, fantasy, (2D), (3D), pre-defined	explore, fantasy, realistic, graphical detail, 3D, (selectable)	(explore), (fantasy), (realistic), (2D), (3D), (selectable)
Perspective	fixed	(bound to char)	(fixed), (bound to char)	bound to character	freely moveable
Structure	(game is finished), (divided into parts), (continuous)	(game is finished), (divided into parts), (continuous)	story, game is finished, divided into parts	story, continuous, (endless)	(game is finished), (endless), (divided into parts), (continuous)

Table 4.5: The five recommended genres and their defining attributes with those in brackets being less distinct for the genre

Chapter 5

Educational Game Design

5.1 Introduction

In the previous chapters we have identified main differentiating features in games and looked for common combinations of these features, which led to the identification of five game genres. These genres are not precisely separated but as shown by the Game Genre Map (Figure 4.8) have smooth transitions, allowing the representation of genre mixtures. To pursue our aim of comparing different types of educational games regarding their success in learning outcome and player experience, the next step is to design educational games. The Tripartite Educational Game Model, explained in the following Section 5.1.1, shows the three main components of an educational game, pointing out what to consider for the design.

5.1.1 Tripartite Educational Game Model

An educational game is the combination of gaming and learning (Figure 5.1). The basic components of gaming are the game and the player, who engages in the game to get entertained. Learning requires at least a learning content and a learner who constructs knowledge from it. While a teacher may support this process, he or she is not a core component. When gaming and learning are combined, learner and player become one and game and learning content need to be connected to build the educational game.

This leads to three core components of educational gaming: game, learning content, and learner/player, which are all connected to each other (Figure 5.2). Of particular

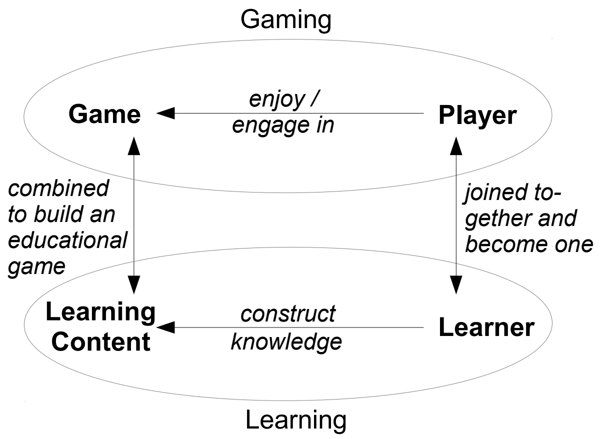


Figure 5.1: Educational gaming as combination of gaming and learning.

interest are the alterations that arise from merging the two worlds gaming and learning. To connect game and learning content it needs to be considered how they can be combined. Combining learner and player means that the needs of both roles have to be taken into account. The learner does not directly access the learning content any more, but only through the game. This affects the connection between the game and learner/player, as the purpose of the game is now not only to entertain, but also to communicate the learning content at the same time. Only if the game and the content can be combined satisfactorily it allows the learner to enjoy playing the educational game and acquire knowledge at the same time.

The Tripartite Educational Game Model (Heintz and Law, 2012) shows that two main components should influence the choices made for the design of the game: the learning content and the learner. In this chapter we discuss both of these components.

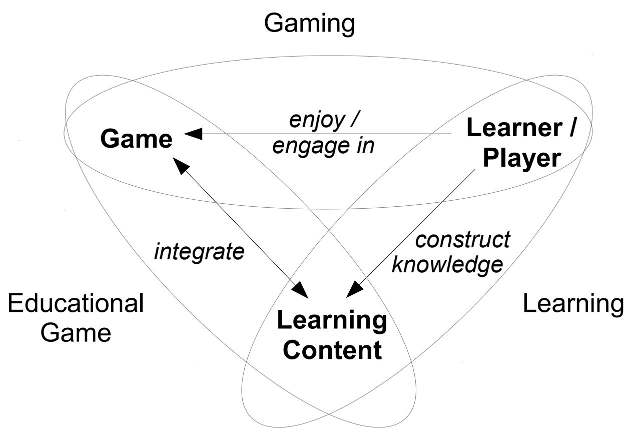


Figure 5.2: Tripartite Educational Game Model: A model of the three basic components of educational games.

5.1.2 Steps for designing an educational game

Moreno-Ger et al. (2008) describe three approaches to the design of an educational game: (1) playability added on top of the content, which may not be considered a proper game due to the dominance of the content; (2) re-purpose an existing game, which may not be possible for every content and may lack focus on the topic due to a dominance of gaming aspects; (3) specifically design a game, with a balance between education and entertainment. For the comparison study we will follow the third approach as we need to control the variables of the game in the experiment.

An educational game is designed to accomplish a particular learning goal. If this is the primary goal of the game, the learning content should be selected first and the game design decisions should be based on how to best support the learning of this content, while still making the game entertaining.

There can be different reasons for why a specific learning topic is chosen, but in our case we first selected the target group (first year university students) and general area (Java programming). To make the educational game most useful, we then conducted a survey to determine which topics students found difficult to learn. In addition we collected information on which part of the process of writing code they struggled most with and what features of a learning tool they would be interested in. The results are presented in Section 5.2.

Once the learning content is decided, the second step is to design the DEG. It is of particular importance to find a way to include the learning content into the game and a strategy on how to teach it to the player, while still being entertaining. In Section 5.3 learning theory as well as theories regarding the game experience will be introduced. To find approaches on how content and game can be combined, existing programming games have been analysed. From the findings and additional literature, DEG design rules are derived for the different elements of a game and as a summary for the different game genres.

5.2 Surveying Target Group

To reiterate, the aim of this thesis is to compare different DEGs to gain an understanding of how design solutions (i.e., choice of game type) influence the learning outcome. In order to conduct a comparison study, a learning topic as well as a target group of learner had to be chosen, leaving the design of the games as the only variable. The selected topic is learning Java programming with first year university students as target group. To be able to build appropriate DEGs, a sufficient knowledge by the author on the learning content had to be guaranteed which is why a topic from Computer Science was an obvious choice. Learning programming is an essential skill for Computer Science students, which makes it particularly relevant for exploring possible didactic improvements by using DEGs. It is commonly observed that a non-trivial number of students in their first year of learning programming struggle with understanding the basic concepts and applying them to solve practical problems. The lack of intrinsic motivation and extrinsic motivators (Kinnunen and Malmi, 2008; Nikula et al., 2011) is identified as a major factor accounting for this problem, which can be improved by appropriately designed DEGs (Barnes et al., 2008). Specifically, fun and other positive emotions derived from playing a DEG can address the motivational issue.

5.2.1 Student survey design

When building a DEG, which has the ability to induce learning but also motivate the player, both aspects should be considered when collecting information from the target group:

1. **Learning requirements:** One goal of the survey was to find out which learning content seems particularly difficult to learn and how a learning tool, or more specifically a DEG, can support the learning.
    - *Which topic should be chosen:* Due to time limitation for the development as well as the evaluation of the DEGs, the learning content needed to be scoped to one specific topic for the study. The target group was therefore questioned to find out which programming topic they found particularly difficult to learn. To further specify the topic, students were also asked which steps in writing a computer program they found most challenging.
    - *Which functionality should be offered:* To gather further requirements we collected information about what students consider helpful in a tool for learning programming.
2. **General motivation**: The second goal of the survey was to monitor the students' attitude towards computer programming and the ease of learning it, to identify if the motivational nature of a DEG could be beneficial for learning programming.

- *How do the attitude towards computer programming and ease of learning it change over time:* If they decrease throughout the semester, using games for learning may be a solution to support students in staying motivated.

5.2.2 Related work

Garner et al. (2005) observed students during multiple programming lab sessions to analyse the problems they encountered. They identified a list of student errors, which were either general issues such as program design and basic structure, or related to specific programming topics such as selections, loops, and arrays. Lahtinen et al. (2005) conducted a questionnaire with students and teachers from six universities in different countries. The students, who had attended one or two programming courses, were asked about general problems with the programming environment and subtasks (e.g. syntax, debugging), as well as different programming topics (e.g. loops, arrays). They were also asked how they learned programming and which learning materials they used. Students' responses were then compared against teachers' opinions on where difficulties lie and which materials they expect to be helpful. An earlier study by Milne and Rowe (2002) also compared responses from students and teachers, but focused on the programming topics. Other researchers (Piteira and Costa, 2012; Tan et al., 2009) reported further results based on the questionnaire items of Lahtinen et al. (2005) and Milne and Rowe (2002).

Based on the above literature, we have developed our own questionnaire that aimed to identify problematic programming topics and options to support the learning.

5.2.3 Method

Participants

The study was conducted in 2012 with first-year university students who attended their first course on learning programming. As we planned to conduct two different DEG comparison studies within one year, two higher education institutes were chosen for the survey: The University of Leicester (UK) and the Furtwangen University of Applied Science (GER). Due to our personal connections to the institutes, the studies were conducted in different countries - one in England and the other in Germany. We have to emphasize that we do not aim to research any cultural background related aspects. The two studies were independent of each other, and we studied both target groups to identify their individual requirements. Both universities teach Java as the first programming language. The British students have further Java courses throughout their study and use Java for individual projects as well as group works. The German students depend less on Java as further courses are optional, but they also need programming skills for projects and as the foundation for other languages (e.g. ActionScript, C++).

Coincidentally, each of these two universities offers three different undergraduate courses with different focuses. The British university runs two computer science courses with a more theoretical (CS-theo) and a more practical (CS-prac) interest. The third course is combined with management modules (CS+Mgt). All courses at the German university are a mixture of computer science, design and management. Two courses focus on computer science, one with a media aspect (CS-media) and one with a web/internet aspect (CS-online). The third course focuses on media design (Design) and is the only course where students graduate as "bachelor of arts" instead of "bachelor of science".

When learning Java from scratch, several fundamental topics need to be covered. There are two common approaches on how to structure the course. Since Java is a strictly object-oriented language, it is possible to start with object orientation and then gradually teach basic topics like variables and methods and procedural programming topics such as loops and conditions, as part of an objects functionality. Another option is to take a more traditional approach by first teaching procedural programming along with a basic understanding of variables and methods and then introducing object orientation as a further developed programming concept (Yau and Joy, 2004). Students attending the three courses of the British university are taught together, starting with object orientation. The German students all start with procedural programming, but because the Design course does not focus on computer science, a more visual approach was chosen, using Processing[1] which is based on Java. Processing provides a simpler programming environment and focuses on visual outcomes. Design and CS-online course are taught together following the same approach, while the CS-media course uses a standard editor to create command line applications. Although the programming course is aimed at beginners, some of the students have some prior knowledge on the subject. In the first questionnaire they were asked to state whether or not they did already have programming experience.

[1]http://www.processing.org/

Qn	semester	questions	
		background+requirements	expectations+perceptions
#1	start	demographic: gender, age math skills prior progr. knowledge	"I'm looking forward to learning programming" "I expect it will be ease for me to learn programming" (1=strongly disagree, 5=strongly agree)
#2	half-way	subtasks involved in progr. learning tool features	"My attitude towards programming" (1=strongly dislike, 5=strongly like) "Overall I find programming" (1=very difficult, 5=very easy)
#3	end	programming topics	

Table 5.1: Overview of conducted questionnaires and types of questions (Qn=Questionnaire)

Instrument

Three questionnaires were administered during an 11-week (15-week in Germany) semester: one in the first session, one half-way through, and one at the end of the course (Table 5.1).

- The first questionnaire (Qn#1) was to collect students' background data and their expectations on learning programming ("looking forward to it", "expected ease of learning") on a 5-point Likert-scale.
- In the second questionnaire (Qn#2), students were asked about their current attitude towards programming and perceived ease of learning it. Furthermore the questionnaire consisted of two sections. The first section was about difficulties with the different subtasks of writing a program - from getting an idea of how to solve a problem, through putting it into code, using the correct syntax, to understanding the program, and being able to get rid of bugs (cf. Table 5.3). The second section was about the features a learning tool could have, including help on memorising, practising and understanding the functionality of commands, visualising how a program works, giving examples, adapting to the student's state of knowledge, and giving feedback on errors (cf. Table 5.4).
- The third questionnaire (Qn#3) asked the same questions on attitude and ease to track changes, if any. Students were also asked to rate how easy they found each topic they had learned during the semester (cf. Table 5.5).

In order to link the data of the three questionnaires, the students were asked to provide their university username each time. Participation was voluntary and students were informed that the data was anonymised before publication.

5.2.4 Results

Preliminary analysis

Number of participants: The number of responses for each questionnaire can be found in Table 5.2. For the first questionnaire Qn#1, which collected demographic data and information about students' programming experience, the number of female students and students with prior knowledge are also reported. The questionnaires were distributed during lectures. Since attendance rates decreased towards the end of the semester, numbers of participations dropped as well. For the evaluation of the attitude towards computer programming and the ease of learning it, a within subject comparison was conducted to compare the answers from the three questionnaires. Table 5.2 shows how many students had filled in all three questionnaires.

	Qn#1			Qn#2	Qn#3	Qn#1-3
university	N	female	preKnow	N	N	N
Leicester	77	10	43	64	41	32
Furtwangen	109	46	45	98	62	44

Table 5.2: Participants from the University of Leicester (UK) and the Furtwangen University (GER) for each questionnaire, as well as number of students who had filled in the full set of all three questionnaires (Qn#1-3).

Outliers: To identify conspicuous data, we examined the boxplot diagrams of all variables and screened each case for repetitive answers. No outliers were found. Particularly high or low ratings were justified by comments or made sense in context with the other answers.

Normality: None of the variables showed a normal distribution. As the median value was often not precise enough to identify differences between the overall responses to different questions, we decided to report the mean value and the standard deviation.

Subtasks involved in programming

Writing a program involves several subtasks, from getting the idea of how to solve a problem to implementing it as lines of code that run without errors. In Qn#2 students were asked to rate how difficult they found each of these subtasks. The results helped us to get a better understanding of where students need more support when learning how to program. Stating the difficulty level of five subtasks involved in programming on a 5-point Likert scale (1=very hard; 5=very easy), students from both university rated the subtasks in the same order, when sorting them by their mean values (see Table 5.3). “Get an idea on how to solve a task” was rated hardest and “Understand how the program works” easiest. However, the mean values are all very close, so no subtask was rated particularly more or less difficult than the others.

	Leicester				Furtwangen			
statement	M	SD	N	dk	M	SD	N	dk
get an idea	3.06	.83	64	0	2.53	.85	95	1
get rid of bugs	3.09	.87	64	0	2.60	.86	98	0
put idea into code	3.13	1.02	63	1	2.87	1.10	97	1
use correct syntax	3.16	1.04	64	0	2.97	.91	96	1
understand the program	3.20	.89	64	0	3.12	.92	98	0

Table 5.3: Results from Qn#2 for rating the difficulty of subtasks involved in programming from very hard (=1) to very easy (=5) ordered by mean values (from most to least difficult). (dk = don’t know)

Desired learning tool features

Presenting seven features of a hypothetical learning tool for programming, the students were asked how desirable they found each feature, rating it on a scale from 1, being very undesirable, to 5, being very desirable. The mean values for all features were above 3.0 and therefore they were all rated as desirable. Both universities highest score was for the feature “Give example programs” and lowest for “Help memorize commands” (Table 5.4). Students thus rather want help with understanding and applying commands than just purely memorizing them. Accordingly “understanding functionality”, “practise”, and “feedback on errors” also received rather high ratings with mean values above 4.0.

	Leicester				Furtwangen			
features that help to/by	M	SD	N	dk	M	SD	N	dk
memorize commands	3.73	.84	64	0	3.48	.89	96	1
visualising how a program works	3.86	1.01	64	0	4.05	.76	97	1
adapting to state of knowledge	3.85	.90	62	2	4.07	.81	97	0
understand functionality of commands	4.13	.81	64	0	4.13	.80	97	1
giving feedback on errors	4.30	.71	64	0	4.13	.71	98	0
practise the use of commands	4.33	.70	63	0	4.07	.89	97	0
giving example programs	4.53	.62	64	0	4.34	.75	96	0

Table 5.4: Results from Qn#2 for rating the desirability of the features for a hypothetical tool to support learning programming from very undesirable (=1) to very desirable (=5) ordered by mean values (from least to most desirable). (dk = don't know)

Programming topics

At the end of the semester students were questioned about each topic they had learned during the course. The list of topics varied with the different courses attended by the students. Since we wanted to be able to use similar games for both studies, only the topics covered by both universities were considered. For each topic, students were asked to what extent they think they understood it (level of understanding: 1=very low; 5=very high) and how easy they found it to learn this topic (1=very hard; 5=very easy). Particularly the ease of learning is of interest, as playful learning with DEGs should enable the students to learn in a pleasurable way. The topics ordered by mean values of ease of learning are shown in Table 5.5.

The topics "variables" and "operands" were in general rated to be best understood and easiest to learn, by students from both universities. "Arrays" were uniformly rated to be the most difficult topic to learn. The level of understanding was also quite low for this topics, with the lowest (Leicester) and second to lowest (Furtwangen) mean scores. Students who started with object orientation rated the corresponding topics "class" and "instance" as slightly easier to learn as compared to other topics, than students who started with procedural programming. Accordingly procedural topics ("selection" and "loop") were ranked higher by the courses starting with these topics.

Leicester					Furtwangen				
M	SD	N	dk	topic	M	SD	N	dk	topic
2.70	.97	40	0	array	2.74	1.05	61	0	array
3.10	.93	40	1	static	2.85	.89	61	0	class
3.13	1.18	40	1	loop	2.88	.98	59	1	static
3.15	.96	41	0	string	2.91	1.01	57	1	instance
3.15	1.08	40	0	instance	3.02	.95	62	0	method
3.17	.97	41	0	class	3.33	1.17	55	3	string
3.46	1.12	41	0	selection	3.40	1.18	50	8	visibility
3.52	.88	40	1	method	3.54	.94	61	1	selection
3.88	.87	41	0	visibility	3.59	.86	61	0	loop
4.00	.89	41	0	variable	3.98	.88	62	0	variable
4.15	.80	41	0	operand	3.98	.84	62	0	operand

Table 5.5: Results from Qn#3 for rating the ease of learning for each programming topic from very hard (=1) to very easy (=5) ordered by mean values (from most to least difficult); N = number of answer from which the mean value (M) was calculated; dk = don't know

Attitude and ease of learning

The attitude towards programming, and the ease of learning it, was measured in each of the three questionnaires distributed over the semester. Since many students did not have any prior experience with learning programming, the first questionnaire was different, as it asked about their expectations. The mean values (Table 5.6) of the tested variables (attitude and perceived ease) decreased for both universities compared to the initial expectations (looking forward and expected ease). Overall students did have a positive attitude towards programming with mean values above 3.4 in Furtwangen and above 4.0 in Leicester. The ease of learning was on average rated as rather difficult than easy, with values below 3.0, except for the expected ease (Qn#1) at the University of Leicester.

To evaluate the changes over time, we used the Friedman test with Wilcoxon signed rank as post-hoc test. The non-parametric tests were chosen, since the data was not normally distributed. No significant differences were found for the (expected/perceived) ease of learning. The attitude however did change significantly over the three measures (Qn#1-3) in Leicester ($\chi^2(2)$=6.859, $p<.05$), as well as Furtwangen ($\chi^2(2)$=22.769, $p<.01$). In Furtwangen the value decreased significantly from a mean rank of 12.18 (Qn#1) to 11.50 (Qn#2) according to Wilcoxon

Questionnaire:		Qn#1				Qn#2				Qn#3			
uni	statement	M	SD	N	dk	M	SD	N	dk	M	SD	N	dk
UK	attitude	4.40	.61	77	0	4.10	.86	63	1	4.00	.93	40	1
	ease	3.06	.85	70	5	2.73	.74	64	0	2.78	.94	41	0
GER	attitude	4.03	.74	109	0	3.64	1.01	97	0	3.46	1.15	59	0
	ease	2.90	.93	87	21	2.45	.84	97	0	2.64	.92	59	0

Table 5.6: Results for students' opinion on programming throughout the three questionnaires Qn#1-3 for both universities in Leicester (UK) and Furtwangen (GER). (dk = don't know; 1 = strongly disagree/dislike, very difficult; 5 = strongly agree/like, very easy)

signed rank (Z=-2.353, $p<.05$) and again from a mean rank of 9.63 (Qn#2) to 8.50 (Qn#3) (Z=-3.258, $p<.01$). In Leicester only the decrease between Qn#1 (mean rank = 7.09) and Qn#2 (mean rank =6.50) was significant (Z=-2.500, $p<.05$). Because of the decrease in attitude, as well as the learning being perceived as rather difficult, we conclude that there is potential for students to benefit from learning with a DEG.

5.2.5 Discussion

Requirements for the design of a digital educational game can be gained from the results on the questions on which subtasks and topics students find difficult and which features they find desirable for a learning tool. "Get an idea of how to solve a programming task" was rated the hardest subtask of programming. The tool should provide training on how to approach a task and build an overall concept on how to solve it. Finding errors in the program derived from this concept along with the correct use of the programming language are further requests, according to the rating.

One way of achieving this is by "Giving examples", which was rated to be a highly desirable feature. Providing students with examples to start from and challenging them with similar problems can serve as the basic concept for a game. It is a common approach of games to start with an easy and understandable scenario and then slowly introduce actions that can be used (e.g. practise the use of commands) to reach a goal. In case of failure the player gets immediate feedback (giving feedback on errors) to understand what was wrong (understand functionality of commands) and the scenario can be replayed (practise) until a solution is found. Therefore the most desired features are compatible with the mechanics of a game.

Topics that students find more difficult to learn can benefit more from the support of a DEG with the potential to motivate and ease the learning. The topic rated most difficult to learn for both universities was “arrays”. Since altering values in “arrays” is done by “loops”, these topics can be combined, although “loops” were rated less difficult to learn. Object orientation is another topic of interest, since “class” and “instance” were also found to be rather difficult to learn for the students, however due to the different learning approaches the topics were found to be less of an issue for the Leicester students, since they learned them early on in the course. We therefore decided to choose “arrays” as topic for our DEGs.

5.3 Design of Digital Educational Games

In this section we discuss the connection between game and learning content in DEGs (Figure 5.2). To gain understanding on different approaches to learning which can be implemented in games, we first give an overview on learning theories. In addition we are interested in how games can foster learning by providing motivation and a productive learning experience. To further investigate options of learning content inclusion in games, especially for topics on learning programming, we analysed DEGs from this area and revealed four general concepts (Section 5.3.2). Finally we link the findings from theory and analysis with our GEAM, by discussing the possible contribution of each element and associated attribute to the learning process. As our redefinition of game genres is based on GEAM, the results of this discussion allow us to gain a basic understanding on how different genres support learning in DEGs.

5.3.1 Theoretical background

Educational games aim to achieve two goals - learning and gaming - whereby the motivating and engaging nature of gaming is expected to be beneficial for the learning. In order to understand how games can support learning, we thus need to consider learning theories as well as theories on motivation and engagement.

Learning theories

A broad range of different theories exist which aim to explain how learning takes place. We discuss the learning theories which are most established and most relevant to DEGs.

Behaviourism is concerned with the change of behaviour in response to external stimuli, but not with the internal processes causing this change. The classical conditioning, as formulated by Pavlov, is the trained association of a stimulus to a response. In games, an example for such trained responses is a player casting magic (response) when hearing the growl of a monster (stimulus), before even seeing it (Siang and Rao, 2003).

Central concepts of behaviourism are repetition and reinforcement. Thorndike found that the strength of an association between stimulus and response increases, the more often it is made, and decreases, when unused (law of exercise) (Hohn, 1995, p.21). However, he emphasized that repetition does not lead to learning without reinforcement. The association is strengthened if it results in satisfaction, but not if it results in punishment (law of effect). The two laws explain what is known as trial-and-error learning. Drill-and-practice games are DEGs which follow this approach (Bruckman, 1999), by heavily relying on repetition of actions, reinforced by rewards. More complex tasks can only be solved if presented in small amounts (Woollard, 2010, p.60).

Skinner's operant conditioning is concerned with how to reinforce desired behaviour. Basic principles are that reinforcement should follow immediately, and consistent, with an appropriate magnitude, contingent on the response (Woollard, 2010, p.47). To avoid extinction (behaviour stops without continuous reinforcement), reinforcement needs to be scheduled. There are different approaches on how to schedule rewards in games, e.g. fixed ratio or random intervals (Siang and Rao, 2003). Rewards/Penalties is one of the central elements in games (Section 3.4), used to learn in-game behaviour (Hense and Mandl, 2014). It serves as feedback for the player, who is expected to change his behaviour accordingly.

Criticism at using behaviourism as learning approach in DEGs is their use of extrinsic motivation, by providing rewards which are otherwise unrelated to the game (Egenfeldt-Nielsen, 2006), like getting points as a reward for completing a level as opposed to receiving an item which can be used later in the game. In terms of the kind of learning drill-and-practice games offer, they are limited to the memorization of information, or training of a skill, but miss to facilitate understanding of a topic.

Cognitivism, in contrast to behaviourism, considers the internal process of learning. It is thus concerned with how information is processed and memorized. Models like the multi-store model by Atkinson and Shiffrin aim to describe how information is stored in memory. Hohn (1995) summarizes the main components of information processing in short as (1) the perception of a stimulus at the sensory register; (2) the processing of the information in the working/short-term memory; (3) the connection of the perceived information with existing knowledge; (4) and finally the storage in long-term memory. The storage of information in the long-term memory needs rehearsal.

Only a fraction of information is passed on from the sensory register to the working memory, while most is discarded to prevent us from information overload. Attention is the deciding factor as to which information gets processed. It is thus important to evoke and guide the learner's attention. This requires that the learning content is structured and organized so that not too much information is presented at once (Siang and Rao, 2003).

The schema theory describes how knowledge is stored and organized in memory. Schemas are units of knowledge or skills which are interrelated, forming a complex network. When information is processed, schemas are updated and new schemas are created (Pritchard, 2014, p.25). Again, the structure of learning content is important, as schemas need to be adapted stepwise, to serve as foundation for further knowledge, which builds upon the previous.

Learning in cognitivism means actively updating and creating new knowledge in memory. Problem-solving in games can foster the process of applying current knowledge and adapting it till a solution is found (Siang and Rao, 2003).

Constructivism considers learners to construct knowledge based on their existing knowledge. This concept of learning is similar to the cognitivist approach. But while cognitivism focuses on internal information processing, constructivism is also concerned with the learning environment and the learners' individual perception of the world. By providing a context in which a problem is presented, the learner is encouraged to explore and try things out (learning by doing). Hence, learners need to be active (Hassan, 2011) and construct their own knowledge.

The engagement of the player is one of the defining elements of games (Section 2.1). Games can provide rich and save learning environments, simulating real world situations or providing context for "self-directed, discovery-, inquiry-, or problem-based learning activities" (Hense and Mandl, 2014). By reflecting upon their experiences in the game, players can adjust their internal representation of knowledge and construct new knowledge (Lainema, 2009).

A special form of constructivism is constructionism, a learning theory which was introduced by Papert (1986). It is based on the idea of learning by constructing or designing things. Papert was particularly interested in using technology to support learning and was involved in the development of the Logo programming language, which was used to teach math to school children.

Situated Learning, also called situated cognition, shares the idea with constructivism that learning takes place in a context. It focuses on the social and cultural setting of learning activities. Instead of learning abstract knowledge, learners are confronted with and encouraged to solve real-world problems in an authentic context with authentic activities (Herrington and Oliver, 1995; Pritchard, 2014). The learner is not isolated but part of a "community of practice" (Lave, 1991). The social aspect of learning is further elaborated in the next paragraph (Social Learning).

If the learning situation is familiar to the learner, who thus can relate to it, the learner is more likely to pay attention and be interested in engaging in the learning activity. Knowledge acquired in a particular situation can however not necessarily be transferred to a different situation.

As mentioned in the paragraph about Constructivism, game environments can simulate real-world situations and provide a context for situated learning. For a DEG which uses the situated learning approach, the environment should be carefully chosen to suit the learning content and evoke the player's interest.

Social Learning is concerned with the interaction of a learner with others, like a teacher, parent, peers or any third person involved in the learning process. The social aspect of learning is discussed in the context of different learning theories, like social constructivism and situated learning.

Some of the central concepts of social learning are (Pritchard, 2014):

- Scaffolding: The learner is supported in understanding something just above her current level of understanding, by someone more knowledgeable (e.g. a teacher). This includes any form of considered help, like a discussion, or provided learning material.
- Apprenticeship: Knowledge is gained by observing and working with an expert, who guides and tutors the learner.
- Peripheral participation: When collaborating with others in a group which hold certain skills or information, this knowledge is passed on, especially to younger members.

As summarised in GEAM (Section 3.4), games can facilitate communication and cooperation between players, thereby providing basic functionality for social learning (Hense and Mandl, 2014). In addition, communities can emerge around games, where players communicate outside of the game, exchange information and support each other, e.g. in forums (Bruckman, 1999; Hense and Mandl, 2014).

Motivation and Engagement

According to the constructivist learning theory, the learner needs to be engaged in an activity, for learning to take place. Similarly, cognitivism assumes that a stimulus is only processed, if the learner pays attention. Games are supposed to be powerful in engaging the player.

The theory behind why people do what they do, e.g. engage in playing a game or a learning activity, is motivation (Hohn, 1995, p.274). This is another property of games, which is related to engagement. As the activity play is mostly detached from the real-world (Section 2.1), there are no external motivations to play games, hence the game itself needs to be motivational in order for people to play it.

There is evidence that motivation, engagement, and learning success are linked (Lepper et al., 2005; Pintrich and De Groot, 1990). For a deeper understanding on the relation between learning and gaming in DEGs, we discuss both concept - motivation and engagement - with an interest on how games can provide them.

Motivation can be explained by humans' basic instincts, needs and desires. Not all of these are equally important for learning. For example the "deficiency needs" at the bottom of Maslow's pyramid of needs (Hohn, 1995, p.276) are mainly prerequisites which can hinder learning if not fulfilled, like being hungry, or not feeling safe.

Reiss (2000, 2004) developed a list of 16 basic desires. Starting out with a list of 328 collected items (after excluding duplicates), which were rated for their importance by 2554 participants and evaluated with a factory study, the resulting 16 desires are: Power (influence others), Independence (be autonomous), Curiosity (acquire knowledge), Acceptance (approval), Order (organization), Saving (collect things), Honor (be loyal, tradition), Idealism (social justice), Social Contact (companionship), Family (raise one's own children), Status (social standing), Vengeance (get even, compete, win), Romance (sex and beauty), Eating (consume food), Physical Activity (exercise of muscles), Tranquility (emotional calm). Relations between these desires and games are discussed in Section 5.3.3.

The self-determination motivation theory distinguishes between intrinsic and extrinsic incentives (Hohn, 1995, p.277). Intrinsic incentives are derived from an interest in the activity itself and depend on individual preferences. Extrinsic incentives are external rewards, like social approval or recognition, or tangible objects.

Ryan and Deci (2000) give a deeper insight on intrinsic and extrinsic motivation, with regards to the psychological needs: competency, autonomy, and relatedness. Intrinsic motivation is part of humans' nature of being curious and active, thus exploring and trying things out. As it depends on the interest of an individual, e.g. if activities "have the appeal of novelty, challenge, or aesthetic value", it is difficult to control this kind of motivation. It can however be facilitated by the "feeling of competence" (e.g. evoked by an optimal challenge, or affirmative feedback) in combination with "a sense of autonomy" (self-determined behaviour). Extrinsic motivations are particularly valuable, when there is no intrinsic motivation. Ryan and Deci (2000) list different forms of extrinsic motivation, from purely external, to internal causes. "External regulation" is what Skinner defined as reinforcement (rewards/punishment) in the operand conditioning. "Introjected regulation" are feelings influenced from the outside that pressure us to do something, like pride or guilt. "Identification" appears if someone accepts the necessity of a behaviour for personal reasons, like learning vocabulary to be able to speak a language. "Integration" is the alignment of an action with desired internal values and is closest to intrinsic motivation, but is still driven by a goal other than the activity itself.

Games offer opportunity for intrinsic motivation (Orr and McGuinness, 2014), e.g. Malone and Lepper (1987) have done extensive research on identifying game features that facilitate intrinsic motivation: challenge, curiosity, control, fantasy (individual motivations), as well as cooperation, competition, recognition (interpersonal motivations).

Engagement is a "necessary part of the learning process" (Brown and Cairns, 2004). Immersion and flow are concepts related to engagement, which have been studied in relation to games. Both are discussed below.

Brown and Cairns's (2004) found that a clear definition for the term immersion was missing. Based on a qualitative study they proposed that immersion occurs in different levels. They defined three levels of increasing depth of immersion: engagement, engrossment, and total immersion. Accordingly they describe immersion as being the highest level of engagement. Engagement requires the investment of time, effort, and attention, as well as the player's interest in the game. Engrossment means that additionally "gamers' emotions are directly affected by the game". Total immersion appears at a high level of attention and depends on atmosphere (e.g. sounds, graphics) and players' empathy for the main character.

Ermi and Mäyrä (2005) identified immersion to be a multidimensional phenomenon, which is individual in each situation, as it depends on the particular game, the particular player, and the context in which it is played. They propose the SCI model to describe immersion, but do not claim that it is comprehensive. The model consists of three dimensions: (1) sensory immersion (audiovisual representation, overpowering stimuli from real world), (2) challenge-based immersion (balance between challenge and motor/mental skills), and (3) imaginative immersion (fantasy and empathy with character).

Jennett et al. (2008) provide a thorough approach on defining immersion in games, based on a detailed literature review and three different studies. Differentiating immersion from the related terms flow, presence, and cognitive absorption, they describe it as experience at a moment in time, which is characterised by "a lack of awareness of time, a loss of awareness of the real world, involvement and a sense of being in the task environment". They agree with Brown and Cairns that the experience is graded and show that it can be determined by subjective and objective measures.

Flow is a concept which is related to immersion. The term was coined by Csikszentmihalyi (1975), who also references to flow as "complete immersion" (Csikszentmihalyi, 1997, p.29); so flow can be seen as the maximal level of immersion. Csikszentmihalyi (1990) describes flow as "a state of concentration so focused that it amounts to absolute absorption in an activity" and lists eight components, frequently mentioned by people who describe a flow experience: (1) completable task, (2) concentration, (3) clear goals, (4) immediate feedback, (5) deep involvement, (6) sense of control, (7) less self-awareness during, but more after the experience, (8) altered sense of duration and time. A close relation can be found between the eight components and learning theory. Cognitivism emphasizes the importance of attention (concentration, involvement) for information to be processed, as well as structured learning content (completable tasks, clear goals) and the ability to actively solve problems (control). The operant conditioning in behaviourism requires immediate reinforcement (feedback).

Jennett et al. (2008) argue against calling flow “simply the extreme end of immersion”, by pointing out that games can provide highly immersive experiences without meeting all the flow criteria (e.g. without clear goals). However, it is not fully clear if according to Csikszentmihalyi’s (1990) flow does indeed require all eight criteria to be met. On one hand he wrote that “the combination of all these elements causes a sense of deep enjoyment” (Csikszentmihalyi, 1990), but on the other hand he said that “to experience flow, it helps to have clear goals” (Csikszentmihalyi, 1997, p.137) and “flow tends to occur when a person faces a clear set of goals” (p.29), which implies that this characteristic is optional.

Nakamura and Csikszentmihalyi (2002) claim that clear goals and immediate feedback on progress are conditions for flow, along with a balance between challenge and skill. Flow is further specified by dividing skill and challenge into areas of low, medium, and high, as the low skill/low challenge combination is despite a balance between them not considered flow, but apathy. Based on results from empirical studies, Nakamura and Csikszentmihalyi believe that the combinations high skill/low challenge (relaxation) and high skill/high challenge (flow) are intrinsically rewarding.

Games are frequently mentioned as example for an activity during which people experience flow. Jones (1998), Sweetser and Wyeth (2005), and Cowley et al. (2008) further investigate the cause of flow in games, based on Csikszentmihalyi’s eight flow components, which they match with game elements. Sweetser and Wyeth however slightly modified the eight components and included social interaction, which is criticised by Cowley et al.. The elaboration from the three authors on how games implement each of the eight flow components is too detailed to describe in full, but references are given in Section 5.3.3.

5.3.2 Analysis of programming games

To gain knowledge about how to design a DEG for learning programming, a search for games was conducted. In particular we were interested in how learning content and game could be merged.

Short description of programming DEGs

Twelve DEGs for learning programming topics have been found and will be described in this section, along with a DEG which has been developed by us in the context of a pre-study for gaining experience in creating and evaluating DEGs.

Barnes et al. (2007) described two games that were designed and created by students during a 10 week summer program in the context of the Game2Learn project. The first game "Saving Sera" is an exploratory game, developed with RPGMaker game engine. Through four quests the player practices nested for loops, do-while loops, print statements and the quicksort algorithm. In the second game "The Catacombs" the user plays a wizard. During the task of casting three spells, the player trains the use of if statements and nested for-loops. This game was created with the Aurora engine for NeverWinter Nights.

Eagle and Barnes (2008) introduced the game "Wu's Castle". It is developed with the RPG Maker XP game engine. Through three different levels the player learns about the use of loops and arrays. In the first level the user can change the values of a for-loop that runs through a one dimensional array to manipulate its entries. Each array position holds the representation of a snowman that can be modified in the loops body by selecting one of different predefined appearances. In five increasingly difficult tasks the player is asked to change the start and end position as well as the step size of the for-loop and choose the setting of the array position to one type of snowman in the body. When the code is executed its effect is shown in an animation. In the second level the player walks through a nested for-loop to experience its process step by step. For each step the progress of the code is shown and the player is asked some multiple-choice questions along the way. The third level is similar to the first, but with a two dimensional array and a nested for loop.

In a later version of the game a level similar to the second was added at the beginning by demonstrating only a single for-loop (Eagle and Barnes, 2009a). The following articles present evaluations of the game: Eagle and Barnes (2008, 2009a,b).

Ibrahim et al. (2011) presented two mini online programming games they have designed. Both games have three levels and are about answering questions related to programming terms. In the first game the answers should be filled into a crossword and in the second game they are attached to ducks in a duck-shooting game.

Hamid and Fung (2007) developed three mobile games, which they described only roughly. In "SpaceOut" the player, represented by a car with a gun, has to answer questions by shooting the right answer that falls among wrong ones from the sky. Shooting a wrong answer turns them into enemies, which will then attack the car. The other two games, "Doggy" and "Snail" are about arranging code lines correctly in limited time. The player, representing either a dog or a snail, has to eat the correct answers for a given question. Because of the short descriptions and the low quality of the provided screenshots, it is not fully clear how the games work.

Muratet et al. (2009) designed a game, based on the Springer engine's mod "Kernel Panic". Through self-written programs students are able to control units, move them around and fight against opponents. A campaign was designed to make the students familiar with the game and enable them to program their own AIs afterwards, to optimize the play. However, Muratet et al. (2009) point out that they still need to be assisted by teachers.

Colobot & Ceebot can both be found online, Colobot under `http://www.ceebot.com/colobot` and its continuation Ceebot under `www.ceebot.com/ceebot`. They aim to teach fundamental programming concepts like variables, loops, conditions, functions, classes and objects. The games are about programming robots to move around, carry things, or shoot enemies to solve given tasks. The robots are programmed by using a special program language called C-BOT, which is similar to languages like C++ or Java. C-BOT provides a pre-defined selection of commands and classes in addition to the basic commands.

Robozzle is a puzzle game which is about problem solving and can be found online (`http://robozzle.com`). The player has to navigate an arrow through a labyrinth collecting stars. To do so a limited set of commands (forward or turn left/right) is provided which can be arranged inside of functions. As the number of commands for each function is limited too, concepts like loops and recursions needs to be used. Through the use of colour it is also possible to create statements which will perform a certain command only if the arrow is upon a field with a specific colour.

Heintz and Law (2012) described a DEG which was designed in a pilot study. It teaches how to process a list using an endless loop which is stopped when the list is empty. In the game, Strings are used as lists of characters. Throughout seven levels, different tasks are set for the player, like counting the appearance of a certain letter. The player is given tiles which can be placed in a pipe through which each letter is passing, when they are processed one after another. The tiles represent lines of code, like an if-statement that checks for a certain letter, the increase of a counter by one, or a flag which can be set to true or false. The player can test a solution by starting an animation which shows how the list gets processed and how the values of variables like the counter or the flag change.

Concepts of learning content integration

The previously listed DEGs for learning programming have been analysed using the Game Elements-Attributes Model (Section 3.4) to identify the game features through which learning content has been included into the game. Results from the analysis can be found in Table A.6 and Table A.7 (in Appendix) and have been summarized by placing the games on the Game Genre Map (Figure 5.3). The placing is based on the features identified from analysing the games, not on an exact measures.

The three areas identified on the map highlight which games share similar characteristics:

- A: Very small games, based around one task. These games are used mainly to learn certain programming terms.
- B: Either multiple small games combined by a story or one with levels of different difficulty. These games are used mainly to learn the application of one programming concept.
- C: Games that involve tasks that build upon each other with increasing difficulty and involve collecting and making use of resources (e.g. command an

army, collect/learn new abilities). These games are mainly for learning the fundamental concepts of a programming language.

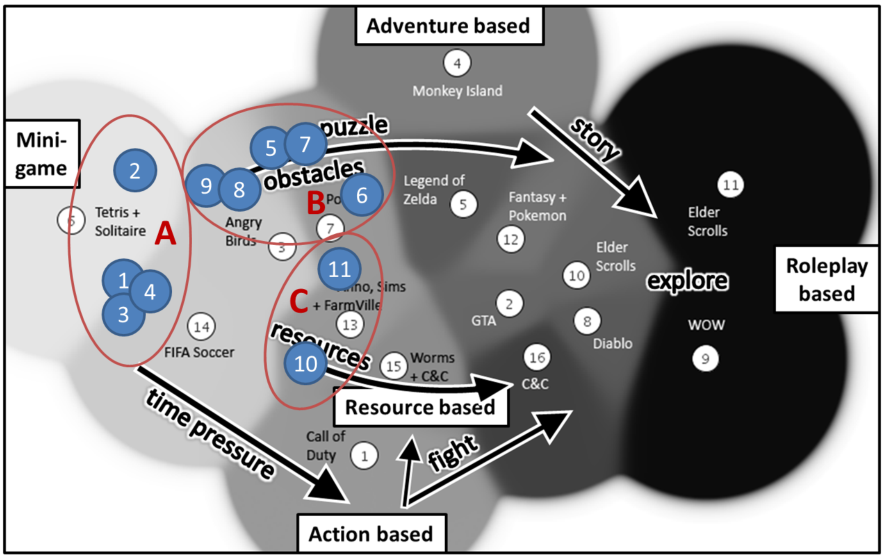


Figure 5.3: Location of analysed programming DEGs on the Game Genre Map. 1: Duck-shooting (Ibrahim et al., 2011), 2: Crossword (Ibrahim et al., 2011), 3: SpaceOut (Hamid and Fung, 2007), 4: Snail/Doggy (Hamid and Fung, 2007), 5: Saving Sera (Barnes et al., 2007), 6: The Catacombs (Barnes et al., 2007), 7: Wu's Castle (Eagle and Barnes, 2008), 8: Robozzle (robozzle.com), 9: List processing (Heintz and Law, 2012), 10: Kernel Panic (Muratet et al., 2009), 11: Colobot, Ceebot (ceebot.com)

Inspecting the games from each area, four general concepts of content integration have been found; two from the games in Area B and one each from the games in Area A and Area C.

- *Concept 1: Content mediated through quizzes (Area A)*
  Games in this group have the learning content included in form of a quiz (an attribute of Challenge), e.g. question/answer, multiple choice, or bringing something in order. The main task in these games is to answer questions. Examples are "Duck-shooting" (multiple choice answers attached to ducks flying across the screen to be shot by the player) and "Crossword" (fill answers into a crossword) (Ibrahim et al., 2011) as well as "Space out" (fill in blanks by shooting correct text falling from the sky), "Doggy" and "Snail" (bring code-lines in a correct order) (Hamid and Fung, 2007). They mainly train the memorisation and identification of programming terms. Some help is given for answering the questions. The Actions are limited to the answer selection process (e.g. through shooting). Besides *quiz*, the other Challenge is *time pressure* (e.g. timer in duck game), apart from possible small extra Challenges

(e.g. avoid being hit by a wrong answer which turns into a bomb in "Space out"). The Goals are *static* (reaching next level) and rewarded via a *score*. The game Worlds are not very elaborate (*2D*, rather low *graphical detail*), the Perspective is *fixed* and the Structure is *multiple levels* (possibly also only one).

These types of games can be used for various learning topics, since they are based on a question and answer concept which is applicable to any kind of learning content. However, they tend not to exploit the full potential of a game by mediating the topic only through one type of Challenges (*quiz*). Frazer et al. (2007) also make aware of a possible distraction from the game by a too heavy use of quiz like questions. A game that is specifically designed for a certain topic with a higher level of content integration might be more effective (see Concept 2 and Concept 4).

- *Concept 2: Content mediated via visualisation of code (Area B)*
  In this group more game elements are used to mediate the content. Instead of *quiz* like for Concept 1 games, the main Challenge is now *puzzle*, so instead of giving or choosing answers, the player is now supposed to use the code to solve puzzle. Another major difference to Concept 1 games is that visualisation is used to explain the content (e.g. a field of eggs representing 2D arrays). "Robozzle" is a Concept 2 game where the player navigates a token through a grid to collect stars using visualised commands. Like a Concept 1 game it has a *fixed* view and is structured in *levels*.

  For Concept 2 games the content is more integrated into the game (especially through visualisation) and cannot be exchanged as in a Concept 1 game.

- *Concept 3: Compendium of Concept 1+2 games linked by a story (Area B)*
  This group includes several Concept 1+2 games, integrating them as quests in a larger game context. A *story* is used to connect these quests. The quests may build up upon each other or are simply a collection of different topics. In the games analysed, the *story* is not used to impart knowledge. Examples are "Saving Sera" (rescuing a princess by solving problems along the way), "The Catacombs" (as wizard saving children from catacombs while learning spells to do so) (Barnes et al., 2007) and "Wu's Castle" (escape from a mirror universe) (Eagle and Barnes, 2008). The games have a *character* which is *moved* around to follow a *story*, building the Structure of the game and guiding the player through different tasks. For each task the current game mode is changed, resembling a Concept 1 or 2 MINI-GAME.

Concept 3 games have similar qualities as those of Concept 1+2, but offer the option to cover a broader learning topic, by braking it down into separate parts which may build upon each other. Although the story parts were not used for learning in the analysed games, this could be a good addition. Otherwise the story helps to engage the player and to link the different tasks.

- *Concept 4: Executed programs coded by players (Area C)*
  For the first three Concepts, the player gets a lot of guidance and help by given code templates (fill in blanks/change parameters) or a selection of answers. Games in Concept 4 require writing small programs autonomously. The player is supported either by instructional material or by a lecturer. Actions are performed by writing a piece of code, offering a large range of options for Actions and Challenges in the game. In "Kernel Panic" (Muratet et al., 2009) the player commands an army and has to *move* units to *find* things, solve *puzzle*, and *fight opponents*. Another example is "Colobot/Ceebot" where the player commands a robot. Similar to Concept 2, visualisation is used to illustrate how the program performs (e.g. robot executing commands) but without focusing on single programming concepts (like visualising loops or arrays).

  Concept 4 games can be used to teach learning content of a much broader scope. However, they are inevitably more effort-demanding.

5.3.3 Design rules for Educational Games

In this subsection we will elaborate some design rules for educational games. After describing the general approach to the design of DEGs we look at each game element and describe possible usages of the game attributes to facilitate learning. Considering results from the previous sections, we then aim to identify how different game genres may be used to incorporate learning content. We want to emphasize that besides any design rules presented in this subsection, there is always creativity involved in designing games, e.g. finding a suitable and captivating story, or an attractive way to visualise the learning content.

General approach

Although it may be possible for certain learning topics to use existing entertainment games which can be repurposed to be used for education (Moreno-Ger et al., 2008), in most cases DEGs will have to be specifically designed.

As described in the Tripartite Educational Game Model (Section 5.1.1), a DEG consists of two parts, the game and the learning content. For the game part, all design rules for entertainment games apply. However, the big difference between entertainment and educational games is the "what" question. For entertainment games, choices about basic idea, setting, story and genre is at the discretion of designers, who can exercise their freedom possibly restricted by budget or customer requests. Whereas for an educational game it is determined by the learning content and learning goal, which are usually identified first (Section 5.1.2). It is important to define a clear learning goal and appropriate sub-goals which allow the learner to achieve the overall goal step by step (Linehan et al., 2011).

How much the design decisions for the game are influenced by the choice of learning content depends on how closely they should be linked. In general game and content can be separate or merged, which is called exogenous or endogenous (Bergervoet et al., 2011; Halverson, 2005; Winn, 2009).

- *exogenous:* The game is only used to entertain between learning sections, or to get rewarded for finishing a learning section.
- *endogenous:* The learning content is integrated into the game, which means that the content is presented to the player in the game and not when the game is paused.

As mentioned in Chapter 1, the research question of how the game type influences the learning outcome and player experience would not be very relevant for exogenous DEGs. The game would only influence the experience, but not the learning outcome, except by motivating the learner to continue playing. We therefore only consider endogenous DEGs, which are also believed to be the more promising approach (Linehan et al., 2011), considering that they make actual use of the different game features and that games inherently support learning (Isbister et al., 2010). Research by Habgood (2007) indicates, that endogenous DEGs are more motivating than exogenous ones.

As identified in Section 5.3.2 the learning content integration can be of different extent, depending on how many elements are utilized for presenting and teaching this content. In principle, learning content can be attached to any game element. The question is how important it is to achieve the learning goal. If it is optional, with the educational purpose being only an additional value of a game, there is much freedom of how to include the content. For example if the optional learning goal is to recognize the faces of historic people, pictures of them could be placed on posters in the game environment. There is a good chance that the player would not pay attention to them and not achieve the learning goal. The highest extent of integration is when all defining game elements (Section 3.3.1) are involved. In the next subsection the game elements are analysed for how each may contribute to the learning in an educational game.

Design choices by game elements

Every game element can contribute to the learning in a DEG. If learning is the main purpose of the game, the learning content should be included in the central parts of the game (Actions+Challenges) to make sure it will be perceived by the player. Incorporating it into the Goals, Rewards, and Structure of the game can assure that progress in the game is only possible if learning goals are reached. In this section we discuss each game attribute of each game element, based on the previously summarized theory on learning, motivation and engagement, as well as the analysis of the programming games.

Player: Design decisions for the game element Player depend on the teaching approach of whether DEG players should learn on their own, or in communication and collaboration with others (Social learning, Section 5.3.1).

single/multi: In a single player game, the designer has more control over the learning process. The interaction with other players, and the learning input resulting from this interaction, is less predictable (Harteveld and Bekebrede, 2011). It depends on how knowledgeable the other players are about the content, whether concepts like scaffolding or peripheral participation are applicable. The apprenticeship concept even requires experts, which means that either teachers or tutors would have to be involved, or players need to spend enough time in the game to become experts and be willing to guide others. On the other hand, fellow players in multi-player games can give intelligent responses and advice, which games with current AIs are not able to provide in such a form.

compete/cooperate/communicate/play alone: Communication is a key requirement for the exchange of knowledge. Players who cooperate are encouraged to exchange information, if it helps them to overcome challenges and win the game. Players in competition with each other are not expected to share their knowledge. In this case, learning can only take place by observing the opponent. Cooperation in combination with communication are the most promising features, if the social learning approach is chosen for a DEG.

Cooperation and competition are according to Malone and Lepper (1987) interpersonal motivations which can enhance “the appeal of the activity”. This is supported by Reiss’s (2000) findings, which defined Social contact and Vengeance as basic desires, as well as Power gained from influencing others, which is possible when collaborating with other players, e.g. by becoming the leader of a group of players.

Input/Output: The choice of device for the Input and Output generally needs to be based on what is accessible to the player and on how to optimize usability. It is important for the player to be able to master the interaction with the game, in order to get a sense of control, which is a criterion for feeling immersed (Jones, 1998; Sweetser and Wyeth, 2005).

Furthermore input and output devices can be essential tools for learning. Particularly motion controllers are relevant for learning physical skills, like “sensorimotor and eye-hand coordination skills”, which can be trained with the behaviourist learning approach (Siang and Rao, 2003). Another use of motion controllers is to facilitate learning by utilising the kinaesthetic memory (Edge et al., 2013). For example in the game “Word Out!”, children learn the shape of letters by forming them with their bodies (Paul et al., 2015; Yap et al., 2015). In addition, according to Reiss (2000), physical activity is one of humans’ basic desires. The use of motion controllers draws on this motivation.

Augmented reality (AR) devices are also relevant for learning. They superimpose computer generated information over the real world and can be used to create interesting learning environments for DEGs (Fotouhi-Ghazvini et al., 2009; Schmitz et al., 2012). AR is particularly interesting for situated learning, as it takes place in an authentic situation (Yusoff et al., 2010), which avoids the problematic transfer of knowledge from a virtually simulated to a real-world situation.

Actions: If a certain behaviour should be learned, it makes sense to incorporate it as Action. However, the Actions always have to match the Challenges in the game as they need to provide the tools to overcome them. This close dependency thus needs to be considered. Actions are responsible for the player’s sense of control and self-determination in the game, which motivates the player (Malone and Lepper, 1987) and facilitates intrinsic motivations (Ryan and Deci, 2000).

move: is a basic Action which is found in nearly every game. It is mainly used to navigate through the game world. If the movement is restricted by rules, the player will have to learn these rules. This can be utilized, like in the game "Wu's castle" (Section 5.3.2), where the player is restricted to walking in circles to understand the functionality of a for-loop in Java.

place/position: is used in combination with the *puzzle* or *quiz* Challenge, to express an answer. Similar to *move*, the rules of where to *place* something on a board or a map can be taught by this action. For example the location of cities can be taught, or a mathematical algorithm on how to determine where to place something.

fight/destroy: can also be used to select an answer, like in the Duck-shooting game with answers being attached to the ducks (Section 5.3.2). A more relevant use of this Action is to train fast reaction, physical skills (in combination with motion controllers as input device), or strategies, if the outcome of the fight depends on the right application of knowledge. Training a skill with a repetitive fight action conforms with the behaviourist approach, while applying existing knowledge to solve a problem follows the constructivist learning theory.

collect: works well with knowledge that persists of single entities. According to the cognitivist learning approach it is important to break knowledge down in small entities and present them one at a time. In this way they can be incorporated in the schema structure of the long term memory. For example in a memory game, the learner can be asked to find and collect pairs of vocabulary with the same word in the native and the foreign language. The Action *collect* is closely related to the Challenge *limited resources*, but also to the Action *produce/build.* Collecting things is described by Reiss (2000) as the basic desire "Saving", which motivates players to make use of this Action.

produce/build: is particularly relevant for the constructionist learning approach. By building or designing something, the learner is actively involved in finding a solution. Any learning content that involves the understanding of how something is built or what it consists of can make use of this game attribute. An example is the game "Crazy Machines"[2], which is an entertainment game that can be used in a serious context, as introduction to physics (Romero and Barma, 2015). "Crazy Machines" is about constructing machines, with provided parts, to solve given tasks.

communicate: is an Action in the game, which is slightly different from the communication with other players, as it resembles communication with non-player characters

[2]http://www.crazy-machines.com/

(NPCs). Artificial intelligence is not yet able to fully simulate a real conversation. The communication with NPCs in the game is thus scripted. It can deal with the individual needs of learners by offering different choices of answers and different paths in the story. An advantage of scripted conversations is that it is carefully thought through, to offer an optimal and efficient learning experience. Communication is in general an important tool for knowledge transfer and for providing reinforcing feedback. Some concepts of social learning can be applied, like scaffolding or apprenticeship, if an NPC is given the role of a tutor or an expert, whom the player can ask for help or further information.

character: gives the player a representation in the virtual world. By observing the character, the player can adapt the behaviour. Voulgari and Komis (2013) describe how characters in Role-play games can support learning. Games which offer multiple characters with different roles, enable the player to explore the game from different angles. Players are encouraged to play different characters with different skills, to support other characters of their own or cooperating players, facilitating social learning. Furthermore, being able to choose and personalise a character, according to individual preferences, motivates the player. Empathy with the character can lead to high immersion (Brown and Cairns, 2004; Ermi and Mäyrä, 2005).

Challenges: For the cognitive and constructivist learning approach, Challenges are a key element, as they provide the problems to be solved by the learner. To support intrinsic motivation and aim for a flow experience, the challenge needs to be balanced (not too easy and not too difficult), so that the player feels competent (Ermi and Mäyrä, 2005; Malone and Lepper, 1987; Ryan and Deci, 2000).

time pressure: can be used if learners need to be able to do a task in limited time (Linehan et al., 2011). Time pressure is less suited for the constructivist learning approach, as learners need to have time to explore. Similarly, in social learning, too much time pressure hinders communication between collaborating players. It is more suited as Challenge for the behaviourist learning approach, where the learner's behaviour is trained by constantly repeating small entities of learning. At the same time, time pressure can prevent a repetitive Action from getting boring, as can be assumed from the result of an experiment by Jennett et al. (2008). Findings by Cox et al. (2012) suggest, that the Action still needs to be cognitively demanding, in order to cause immersion.

limited resources: is a Challenge related to the Action *collect* and *produce/build*. When combined, the constructionist learning approach can be implemented. Limiting the resources helps to shape the problem, which needs to be solved by the player. For example in the game "Crazy Machines"[3], only a selection of parts is provided to build a machine, providing a clear task for the player. In a memory game, only two cards can be revealed at a time, forcing the player to memorize their content. Limiting resources can also be used to teach the value of a resource, by making it harder accessible, or in general how to manage resources.

opponents: are competing players, which are controlled by the computer. We therefore reference to the paragraph on Player *competition*. However, the opponents can also be of low intelligence and serve as reinforcement when defeating them. As mentioned before, power and vengeance are basic desires and motivators (Reiss, 2000). Hence, *opponents* motivate the player to train a repetitive, or a strategic and cognitive demanding *fight* Action.

obstacles: work well in combination with the *move* Action. Overcoming an *obstacle* serves as goal for training a particular *move* Action, as it requires effort and dexterity (e.g. jump from one rock to another without falling). Hence, physical skills can be trained by this Challenge. If the user needs cognitive abilities to pass an *obstacle*, we regard this as *puzzle*, rather than an *obstacle* (see below).

puzzles: are particularly suited for the cognitive and constructivist learning approach (Bruckman, 1999). Learners apply their current knowledge on a presented problem and acquire new knowledge during the process of solving it. Puzzles are thus important to train the learner's cognitive skills. To be motivating, they need to meet the player's abilities, which means that either hints have to be provided or the difficulty of the puzzles need to build up slowly (Sweetser and Wyeth, 2005). Puzzles were identified as main Challenge in games using the identified Concept 2 for learning content integration (Section 5.3.2).

quiz: is a universal approach to learning, as questions can be designed for any kind of content. It may however not be the preferable approach, unless being able to answer questions is the learning goal (Linehan et al., 2011). According to Egenfeldt-Nielsen (2006), question and answer serves the behaviourist learning approach, as linking the correct answer with a question multiple times fosters learning this link. Quiz is also the main Challenge for learning content integration, when following the identified Concept 1 (Section 5.3.2).

search & find: encourages the player to *explore* the environment. Further information on how exploring facilitates learning is given in the description of the game

[3]http://www.crazy-machines.com/

element World/Setting. In the game "Catacombs", *search & find* was used to learn by spotting the right solution between wrong ones, i.e. the player has to search through multiple roles with spells to find the correct spell, with the correct code fragment for a computer program (Table A.6).

savability: is important for assessment of the learning goal. If a process or routine has to be done correctly from start to finish, or if it is the learning goal to answer a certain number of questions correctly in a row, it should not be possible to save the game state, go back and correct mistakes.

Goals: If the learning tasks have been integrated into the Challenges of the game, the Goals of the game (overcome the Challenge) will also be the learning goals. The importance of clear goals is emphasized by the flow theory (Csikszentmihalyi, 1990).

static/various/self-defined: For the behaviourist learning approach, a repetitive action needs to be carried out and new knowledge is introduced in small entities. Similarly the cognitivism theory requires the learning content to be structured in small tasks, as only a limited amount of information can be processed at a time. Both approaches can typically be structured in form of levels with increasing difficulty (see game element Structure), for which *static* Goals are a straight forward choice (i.e. reach the end of similarly designed levels). Constructivism and situated learning are more concerned with an exploring learner, which requires multiple different tasks, and therefore *various* Goals (e.g. different kinds of quests or missions). Giving the player a *choice* of which Goals they want to aim for, or let them *self-define* their Goals facilitates intrinsic motivation as it supports self-determined behaviour and autonomy (Ryan and Deci, 2000). However, it also makes it more difficult to guide the learning, as all possible learner choices need to be considered when designing the learning progress.

Rewards/Penalties: This game element provides reinforcement, which gives the player feedback on his behaviour. Rewards can foster flow experience, as they clarify if a goal was reached (Jones, 1998) and support the player's skill development (Sweetser and Wyeth, 2005). They are also extrinsic motivators. Rewards like praise and score, which do not have an impact on the game, are external regulators, while the rewards or penalties that do have an impact on the game are more integrated (Ryan and Deci, 2000). Egenfeldt-Nielsen (2006) claims that extrinsic motivators bare the risk for DEG players to only focus on the game part, but not the learning.

praise/score: serve as a more or less exact measure of how well the learner performed on the task (Linehan et al., 2011) and are well suited for the behaviourist learning approach (Hense and Mandl, 2014). They can motivate the player, as they satisfy the desire for acceptance, by giving approval, and for status, as highscores shared with others impact a player's social standing or recognition (Malone and Lepper, 1987; Reiss, 2000).

new level/resources/power-up: have an impact on the game and work well with structuring the introduction of new knowledge. Considering the cognitive learning theory, when a player has learned one piece of information, something new can be introduced, which builds upon the existing knowledge. The player gains access to a new level with a more difficult task, or gains resources or power-ups, which are needed to solve the next task.

World/Setting: Particularly situated learning emphasizes the importance of the learning environment. Simulating an authentic task in an authentic environment provides a concrete learning context. Visualisation can be used to illustrate and explain the learning content; a method which we found to be used in DEGs, which follow Concept 2 of content integration (Section 5.3.2).

sound effects/music: are relevant for any audio related learning, like language learning, or learning the sound of different instruments. Sound is also important for sensory immersion (Ermi and Mäyrä, 2005).

2D/3D: depends on how the learning content is presented best. 3D allows for a more natural representation and needs to be chosen if an authentic world is required. de Freitas and Neumann (2009) claim, that 3D is an important element to make the exploration of a world more engaging. On the contrary, 2D reduces the complexity and can provide a clearer presentation.

explore: is central to the constructivist and situated learning approach. The environment needs to include the learning information and present it in a form that allows learners to actively find and try things out. Making use of curiosity as basic human desire (Reiss, 2000), the World needs to be interesting and inviting for the player to interact with.

graphical detail: Objects, which are part of the learning content, need to be shown with enough precision. In combination with sound, graphics support immersion (Ermi and Mäyrä, 2005) by generating an atmosphere (Brown and Cairns, 2004).

abstract/realistic/fantasy: needs to be chosen in accordance with the learning content. For situated learning, an authentic setting requires realistic graphics. Abstract graphics can be used for a schematic representation of information that focuses the learner's attention on the relevant information, removing irrelevant details. Also, if the learning content is about abstract concepts, an abstract visualisation is the obvious choice. On the other hand, fantasy can "provide appropriate metaphors and analogies" (Malone and Lepper, 1987) which can be used to explain abstract concepts. In addition, the use of fantasy or realistic graphics supports the imaginative immersion of the player (Ermi and Mäyrä, 2005) and enables the player to feel a sense of presence (Jennett et al., 2008).

Perspective: The Perspective is chosen with the aim to achieve the highest possible usability (e.g. camera follows automatically the character). It can also be used to guide players' attention, by showing only the relevant information at the time and thus reducing the cognitive load, ensuring that the important information is processed. The virtual environment allows players to explore things from a perspective which cannot be provided in real world, e.g. giving insight into the circulatory system of the human body (Barbosa and Silva, 2011). Restricting the view can induce curiosity, and motivate the player to explore the game world.

Structure: The Structure of the game makes it possible to guide the player through the learning process, depending on the chosen learning approach. Cognitivism requires a clearly structured learning content, behaviourism requires repetition, and constructivism requires self-directed learning with freedom to explore. In order to support a flow experience, the skills of the player should increase during play and challenges should be presented at an appropriate pace (Sweetser and Wyeth, 2005).

parts/continuous: One approach to learning is to break down the tasks into multiple *parts*, which build upon each other and slowly increase in difficulty (Linehan et al., 2011). The division of a game in *parts* leads to a clear separation between different tasks, requiring that one has to be finished to gain access to the next one, which follows the idea of the cognitive learning approach. If the player should be free to explore new knowledge, or if several steps can be learned in parallel, a *continuous* design needs to be chosen, which suits the constructivist or situated learning approach. Typically *parts* are presented in form of levels. Clearly structured and completable tasks can induce flow (Jones, 1998).

usually finished/endless/multiple times: If the learning content is spread throughout the whole game, the player needs to finish it to perceive all knowledge. An endless game thus only makes sense for teaching a skill, which can be further improved through more training, e.g. a game to learn typing. Following the behaviourist theory of exercise, skills can also be training by playing a game multiple times (Egenfeldt-Nielsen, 2006).

duration: depends on the amount of knowledge and the speed in which it is taught. Small entities of learning content can be combined in one game, following the identified Concept 3 of content integration (Section 5.3.2). An advantage of the virtual reality in games is, that time spans are independent from reality (Jones, 1998). This makes it possible to speed up the presentation of an experiment, which would take much longer if presented in reality.

story: can be utilized to provide additional information to the learner, e.g. to support the cognitive or behaviourist approach of learning (Anagnostou and Pappa, 2011; Hense and Mandl, 2014). The story can be a powerful tool to motivate and immerse the player, as due to the narrative, players identify more with the character and feel more present in the game world (Schneider et al., 2004). According to Bopp (2008), a story can be used "as a frame that adds sense to an activity", which motivates the learner for this activity. Another use of the story, which he points out, is to make it easier to memorize information, as humans are good at remembering stories.

Design choices by genres

In Chapter 4 we identified five game genres and their defining game attributes (Figure 4.8, Table 4.5). Based on the preceding discussion on how the different attributes can support the learning process in DEGs, we conclude design rules for each game genre.

Mini-games are using only a small set of game attributes. This game genre is suitable to teach a particular learning topic with a narrow scope. Frazer et al. (2007) criticised that Mini-games may be too small to immerse the player, lacking conversation options, and being unable to provide a context for the learning topic. On the other hand, their restrictiveness makes it possible to focus on a specific task, and features like conversation or context may not always be required, unless the situated or social learning approach is chosen.

In general, Mini-games are well suited for the behaviourist learning approach. Simple Actions like *move* or *fight* (without a complex combat system), which are characteristic for Mini-games (Table 4.5), can be used repetitively to train single learning entities. To set a Challenge and possibly also assess the learning success, *time pressure* can be used (repeat Action correctly in limited time) as well as simple *opponents* (defeat a certain number of opponents by correctly applying the Action), or for training physical skills *obstacles* (overcome a certain combination of obstacles). Furthermore, the *quiz* Challenge can be used for fact learning (Anagnostou and Pappa, 2011). Extrinsic rewards like praise, or points are typical for the behaviourist learning approach. A highscore can add additional motivation. To progress to the next learning entity, after successfully learning the previous one, a level structure is suitable.

Tending more towards the slower paced adjoined Adventure game genre, another learning approach for Mini-games is cognitivism. The Challenge *puzzles* provides problems which encourage the learner to apply prior knowledge and gain new knowledge while solving them. Levels help to structure the learning content, as required by cognitivism.

Learning content is integrated in Mini-games either as simple Actions, which are the behaviour to be trained, or through the Challenges *quiz* or *puzzles*. Moreover the characteristic *2D* interface of Mini-games, mainly with *abstract graphics* and a *fixed view*, provide a clear and schematic representation for visualising abstract concepts. A small background *story* can provide a metaphor which helps the player to make sense of the Actions or Challenges she encounters in the game.

Adventure games have a strong narrative component, which can be used to provide information (Hense and Mandl, 2014). In the course of the story, different paths can be offered to the player, to try out different solutions, but also to be send on a detour, when giving a wrong answer, to receive further information and explanations. In addition, if the story is captivating it motivates the player to finish the game, which is characteristic for this game genre. With the Challenges *puzzle* and *search & find*, Adventure games confront the player constantly with problems which need to be solve, which is why they are particularly suited for the cognitive learning approach (Hense and Mandl, 2014; Siang and Rao, 2003), as well as the related constructivist approach. Especially Adventure games which are less linear and closer to the Role-play genre on the genre map, give the player freedom to *explore*, which is an essential concept of constructivism.

Further defining attributes of Adventure games are the *emotional connection to the character*, facilitating players' immersion, and *no penalties*, which means that this genre focuses on positive reinforcement.

Sommeregger and Kellner (2012) introduced guidelines on how to design an educational Adventure game. First the setting should be chosen, considering the learning content and its appeal to the target group. The setting lays ground for decisions on story, character, world, and puzzles. While we agree that the appeal to the target group is important, as it increases the chance of players being intrinsically motivated (Ryan and Deci, 2000), we think that considering the *puzzles* first may be a better approach, as problem-solving is the central concept for the learning process. The setting needs then to be chosen accordingly.

Action games are similar in their Action and Challenges as the fast paced Minigames, since they also use *fight* and *move* as Actions and *time pressure* and *opponents* as Challenges. The distinction between the two genres is that the more elaborate Action games are often played in rich *3D* environments with *realistic* and *detailed graphics* and with *multiple players*.

In literature we found that DEGs which are based on Action games (as well as sports, racing and fighting games, which are all summarized under Action in our game map), make use of the behaviourist approach of learning, as they mainly teach physical skills, like eye-hand coordination or fast reaction (Hense and Mandl, 2014; Siang and Rao, 2003). We argue however, that they are also suited for situated learning, due to their authentic environments, as well as social learning, as players cooperate in teams. A drawback here are the Actions, since for example shooting is a very narrow activity which does not leave room to integrate different types of behaviour to be learned. *Action-Adventure* as mixed genre can therefore be a more promising choice for non-physical learning in DEGs, by adding *puzzles* as Challenge (Anagnostou and Pappa, 2011).

Resource games provide with *limited resources* in combination with *collect* and *produce*, a foundation for learning by construction. E.g. Hense and Mandl (2014) suggests that strategy games (a classic game genre related to our redefined Resource genre) are suited for constructivist learning, from which constructionism is derived. The player is encouraged to *explore* the environment in order to find the required resources, or use them in a *fight* against *opponents*. According to

Siang and Rao (2003) strategy games use the cognitive learning approach, which is however related to constructivism.

Learning content can be integrated in form of the resources available to the player and the way they are used to produce something. Tasks like building machines or producing food by following a particular process, can teach the player how certain things work. For example in the entertainment game "The Settlers Online", the player has to build a production chain for making bread[4], including a farmer who grows corn, a mill which grinds corn to flour, a well which supplies water, and a bakery that uses flour and water to produce bread.

Roleplay games combine characteristics of the Action, Adventure and Resource game genre. They thus offer a lot of potential for the integration of learning content, but contain the risk of being less effective, if the learning content is only integrated in some parts of the game.

Roleplay games are particularly well suited for constructivist and social learning, as *exploring* and *communication* with other players and non-play characters are two main characteristics of this genre. Players collaborate in teams, where different players usually have different roles and need to be able to coordinate their actions in order to be successful (Hense and Mandl, 2014). Furthermore players are able to *personalise* their characters and tend to get *emotionally attached* to them, which has a positive impact on the players' motivation and immersion. The player takes on a role through the *character* and acts it out in a *fantasy* or *realistic* environment. Exploring the environment with the character while solving various tasks, follows the constructivist or situated learning approach, which is further supported by providing the player with *choices* on which goals they want to aim for.

5.4 Digital Educational Games for Comparison Study

For the comparison studies of the impact of game type on the quality of learning in educational games, three games have been developed for Study 1 and two games for Study 2, where one is a modification of a game from the first study.

[4] http://thesettlersonline.wikia.com/wiki/Bread_Production_Chain

5.4.1 Games for Study 1: differing by one attribute

Mini-games are particularly well suited to compare single game attributes, as they consist of a lower number of attributes. Thus there are less interrelations between them and a single attribute can be changed without making too many further adjustments to the game. For Study 1, three Mini-games have been developed, which can be assembled in two pairs that both differ by one attribute only (Game 1 and Game 2 differ by an attribute of the Challenge element: *puzzle* versus *time pressure*; Game 2 and Game 3 differ by an attribute of the Setting element: *realistic* versus *abstract* graphics):

1. Game 1 - PUZZLE: The attribute to be measured by this game is the Challenge *puzzle*. It is a Mini-game that leans towards Adventure by including *puzzle* as one of the typical Challenges in Adventure games. However, it does not include a *story*, which would be characteristic for Adventures. There is only some introductory information to motivate the aim of the game and its Setting. The graphics are *realistic*, as they represent real world objects, but are drawn in a simplistic style, with a *low level of graphical detail*.

2. Game 2 - ACTION: This game differs from the PUZZLE game in the element Challenge, as it leans towards the Action genre with *time pressure* as the main Challenge.

3. Game 3 - ABSTRACT: This game differs from the PUZZLE game in the graphical style of the Setting, as it is *abstract* instead of *realistic*. The objects in the PUZZLE game are replaced by symbols like dots and letters.

Although all three games are Mini-games, their slight differences makes them lean towards different more advanced game genres. As mentioned above, the PUZZLE game is somewhat close to the Adventure genre, while the ABSTRACT game with its abstract graphics is further away from it. The ACTION game is close to the Action genre, thus heading into a different direction. This can be visualised by positioning the three games on the Game Genre Map (Figure 5.4).

The three Mini-games PUZZLE, ACTION and ABSTRACT enable the following comparison of game attributes:

- Challenge: *puzzle* versus *time pressure* (comparing the PUZZLE and the ACTION game)

- Setting: *realistic* versus *abstract* graphics (comparing the PUZZLE and the ABSTRACT game)

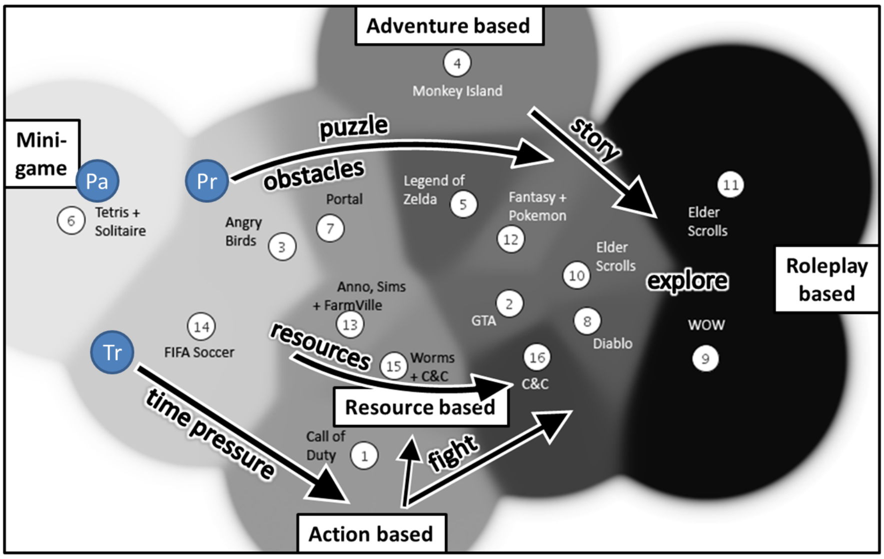


Figure 5.4: Location of the DEGs designed for Study 1 on the Game Genre Map. Pa = *puzzle* Challenge + *abstract* graphics (ABSTRACT); Pr = *puzzle* Challenge + *realistic* graphics (PUZZLE); Tr = *time pressure* Challenge + *realistic* graphics (ACTION).

Design of DEGs

As described in Section 5.1.2 we first chose the target group and learning content. As described above, we decided to use the genre MINI-GAME for the DEGs in Study 1 and identified different design options for this genre on how to teach the selected topic in the context of the game.

The aim of all three DEGs is to explain the basic functionality of accessing a 2D array in Java, from a single cell up to multiple cells via loops. The topic was chosen based on the results of a survey in which computer science students, as our target group, stated that 2D arrays was a topic which is difficult to learn (Section 5.2.5).

For the design of MINI-GAMES, two basic options have been identified (Section 5.3.3): Either the behaviourist learning approach with *time pressure* as main Challenge, or the cognitive or related constructivist approach, with *puzzle* as main Challenge. We decided to design games for both options. With the choice for MINI-GAMES, we can follow their characteristic set of attributes (Table 4.5) in deciding how to implement the different game elements. MINI-GAMES are usually *single-player* games with *restricted movement* as Action and *no characters* which resemble the player.

The Goals are *static* (e.g. by using a level structure) and Rewards mainly given in form of *scores*. The Setting is *pre-defined* with a *fixed* Perspective, which often shows the whole small *2D* Game World at once. The graphics are often *abstract*, but can also be *realistic* or *fantasy*. Due to this choice we decided to make the type of graphics another attribute which can be modified for comparison - for Study 1 we chose to compare *abstract* and *realistic* graphics.

For Mini-games the identified Concepts 1 or 2 are suitable to integrate the learning content into the game (Section 5.3.2). Since Concept 2 is more endogenous, as the content is merged closer with the different game elements, we chose to follow this approach. In Concept 2, visualisation is used to explain how programming code is processed and programming tasks are included as *puzzles*. The ACTION game which uses *time pressure* thus also has *puzzle* as a challenge. However, *time pressure* is the main Challenge, which consequently requires the *puzzles* to be solved in limited time and thus to be much easier than in the PUZZLE and ABSTRACT game, where *puzzle* is the main Challenge.

Below we describe the design of the three DEGs and in particular the integration of learning content. We first show how the learning content was structured, how it was merged with the game aiming not to compromise the game experience, and how the visualisation of code was implemented. These general design decisions were the same for all three DEGs. Subsequently their differences in Gameplay are presented, with a brief description of each game.

Stepwise introduction of content: When designing a DEG it is essential to define clear learning goals, to be able to assess the learning outcome, but also to divide the content into smaller entities as needed for the behaviourist, cognitivist, or constructivist learning approach. Using multiple levels, the player is guided through the learning content while the game gradually builds up in difficulty. The learner is given the time to learn the content step by step, which is how a good level of difficulty can be achieved. For the player to feel immersed and possibly experience flow, the Challenges should not be too difficult but also not too easy, to avoid frustration or boredom.

The following figures show how the learning topic was broken down into four different blocks with increasingly complex code, which in the games with realistic graphics were represented by four different types of planes:

 accesses one **field**, using the code:

```
dropAt(field[1][3]);
```

 accesses a **column**, using the code:

```
for(int r=0; r<4; r+=2){
    dropAt(field[r][1]);
}
```

 accesses a **row**, using the code:

```
for(int c=1; c<4; c+=1){
    dropAt(field[2][c]);
}
```

 accesses an **area**, using the code:

```
for(int r=0; r<4; r+=3){
    for(int c=0; c<5; c+=2){
        dropAt(field[r][c]);
} }
```

For each plane except the first one, the variables are introduced over time, to limit the amount of new knowledge and options given to the player. For each new variable or plane type, there are multiple tasks to practise and apply the newly learned content. This is done by the division into levels (Figure 5.5).

Figure 5.5: The 12 levels of the DEGs for Study 1 (ABSTRACT game has letters instead of planes).

The levels are alternating introducing new concepts or variables and practising all previously learned concepts. Each game has the following structure:

level 1: introduces concept of accessing a **field**
(variable for index of row and column)

level 2: **practices** concept introduced in the previous level

level 3: introduces concept of accessing **whole row or column**
(for-loop plus variable for row or column)

level 4: **practices** concepts introduced in the previous levels

level 5: introduces concept of accessing **part of a row or column** (variables for start and end condition)
level 6: **practices** concepts introduced in the previous levels
level 7: introduces concept of accessing fields in a **row or column with a given step size** (variable for incrementation)
level 8: **practices** concepts introduced in the previous levels
level 9: introduces concept of accessing a **defined area** of the array (second for-loop plus variables for start and end conditions)
level 10: **practices** concepts introduced in the previous levels
level 11: introduces concept of accessing an **area with given step size** for row and column (variables for incrementation)
level 12: **practices** concepts introduced in the previous levels

Each new topic is introduced by a screenshot of the following level, with overlaid descriptions explaining the new variables and how they control which fields in the grid (see visualisation) are accessed. An example for this information screen is shown in Figure 5.6. All previous information screens can be retrieved at any time during the game via the help button.

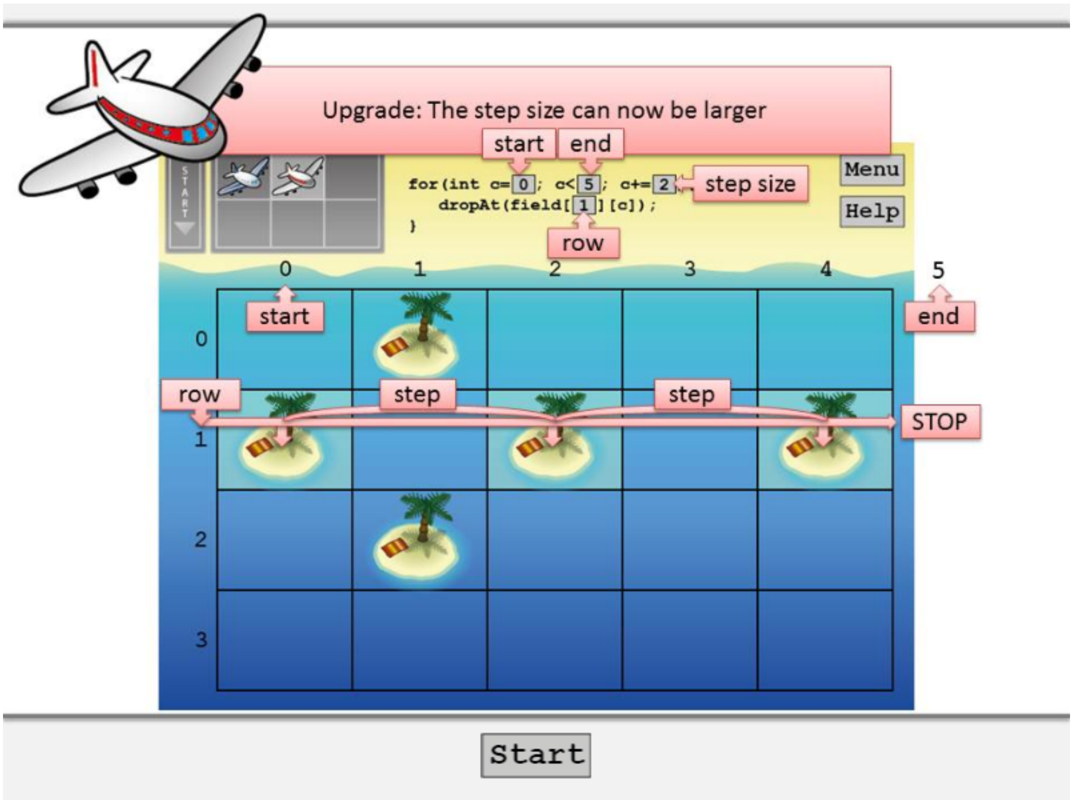


Figure 5.6: An info screen explaining how the variables control the access to fields of the 2D array.

Non-disruptive content presentation As an educational game is first and foremost a game, the presentation of the learning content should not be disruptive, i.e. there should not be interruptions with long text explaining the content. This was achieved by the following design decisions:

- Only the essential code is displayed to the user, which is one line specifying the access of a cell in the array and up to two more lines with *for*-loops. This draws players' attention to the code of interest, by not distracting the learner with the code of a whole program, and leaves more space to visualise the array and the process of how it is accessed.
- Most parts of the code are already given and cannot be edited. Only the values of the variables that define which cells of the array are accessed can be modified. This avoids, for example, syntax errors, where knowing the correct syntax is not the defined learning goal. It makes the interaction with the code much easier, thereby allowing more time for players to explore the effects of changing the values of the variables. For the ACTION game, the time pressure requires even simpler interaction, which is why the code cannot be edited in this game, but the accessed cells are selected instead (see below).

Visualisation All three games are based on the idea of visualising the 2D array and the access to its cells. As they are Mini-games there is a *fixed* view on the Game World, which is *two dimensional* and dominated by a grid, representing the 2D array (e.g. Figure 5.7). To use more game related terms, we call the grid a board, and the cells fields on the board. To determine the index of each field, the board is labelled with numbers on two sides. The access of a field in the array is visualised by animation. In case multiple fields are accessed, the animation also shows the order in which the fields are accessed.

PUZZLE Mini-game

The challenge in this game is to use multiple planes to match a certain pattern on the fields of the grid, representing the array. With increasing functionality (new planes and revealed variables) the patterns can get more and more complex. It is marked which fields should be accessed and how many times. Some fields need to be accessed more than once. An example level is shown in Figure 5.7.

The graphics of the game are designed in accordance with a background story, explaining why to solve the pattern. Each field that should be accessed includes

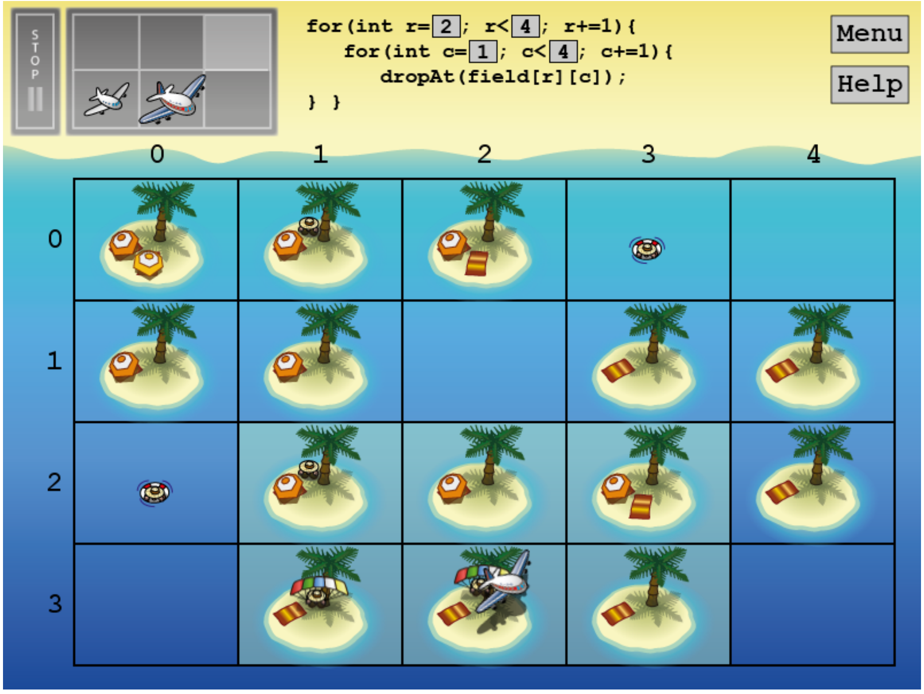


Figure 5.7: PUZZLE Mini-game for Study 1, level 10.

an island on the ocean, while the other fields are empty and thus only show water. The code is used to guide planes to these islands by dropping holiday makers at each field accessed by the code. The number of towels on an island indicates how many planes should drop a holiday maker at the particular field. Dropping too little holiday makers on an island means that not all places are taken, dropping too many means that they will not all find a place. If they are dropped at a field without an island, they will end up in the water, secured by a lifebelt. In all three cases the player has to revise her answer.

A main characteristic of the PUZZLE game is its slow pace. Players can explore the effect of the variables with no *time pressure*. Fields that are accessed by the code are highlighted, so students immediately see the effect when changing the value of a variable. The solution can be tested at any time. If it is not correct, the player can just continue working on it, without being penalised.

ACTION Mini-game

In this game the player competes against the clock. Directly clicking on the large fields of the board, representing the 2D array and taking up most space in the Game World, is faster than changing the variable values in the code. Therefore in the ACTION game, instead of adjusting the code, the player needs to read the code and mark the fields that will be accessed by it. This is done for one plane

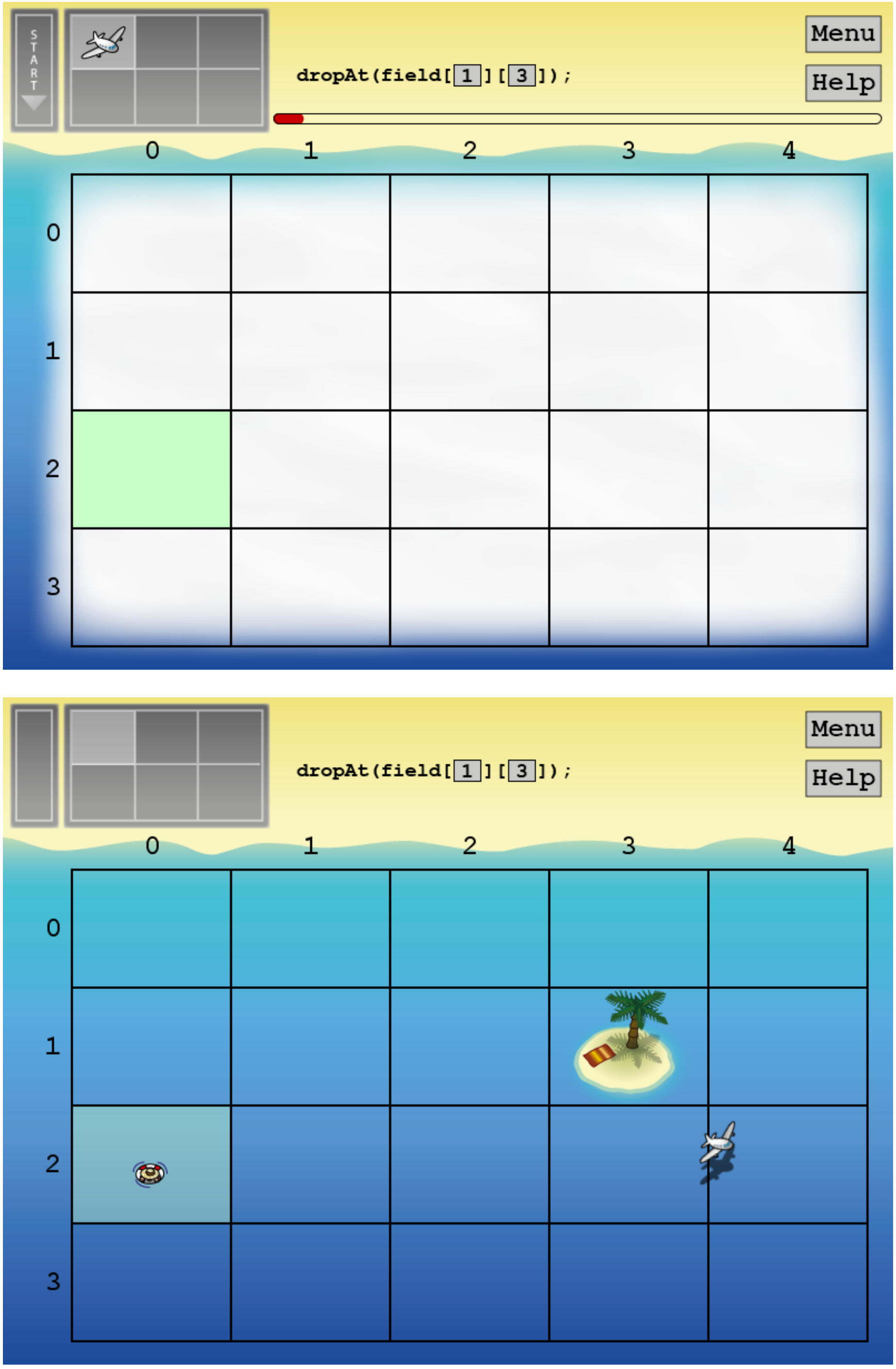


Figure 5.8: ACTION MINI-GAME for Study 1, level 1.

at the time. To hide the solution (islands on the fields which are accessed), fog is placed on top of the ocean and only when the player starts the plane, or when it starts automatically as soon as the time limit is up, the fog disappears. The islands become visible and the holiday makers are dropped at the fields selected by the player, verifying if the solution was correct. For the more difficult planes, more time is granted. Most levels have more than one plane. As soon as a plane is solved, the fog reappears, a new plane with a new code example is given and the timer starts to run. If the player makes a mistake on one of the planes, the level restarts. The two states of selecting fields and revealing the solution can be seen in Figure 5.8.

ABSTRACT Mini-game

The ABSTRACT Mini-game is exactly the same game as the Puzzle Mini-game, but with abstract graphics. Instead of planes, letters are used to indicate if the associated code is to access just one field (F), one row (R), one column (C), or an area (A). The islands are replaced by circles, indicating the fields which should be access by setting the variables in the Java code. Multiple circles indicate that a field should be accessed multiple times, similar to the number of towels on an island in the PUZZLE game. If solved correctly, a green dot appears in each circle. Red dots on a field indicate that a mistake was made. Figure 5.9 shows an example level of the ABSTRACT Mini-game.

E-learning TOOL for control group

Using the GUI of the ABSTRACT Mini-game, but removing the game elements Challenges, Goals, and Structure, an e-learning tool was built which served as DEG substitute for the control group in Study 1. The info screens explaining how to use the application were shown at the beginning and could be accessed via the help button later on. In the TOOL, all functionality is accessible from the start by providing the four different letters with associated code for accessing a field, row, column, or area (see ABSTRACT Mini-game) and by making all variables editable without the step by step introduction of new content like in the games. Also carried over from the ABSTRACT Mini-game was the animation option, to visualise the order in which fields are accessed. A screenshot of the TOOL is shown in Figure 5.10.

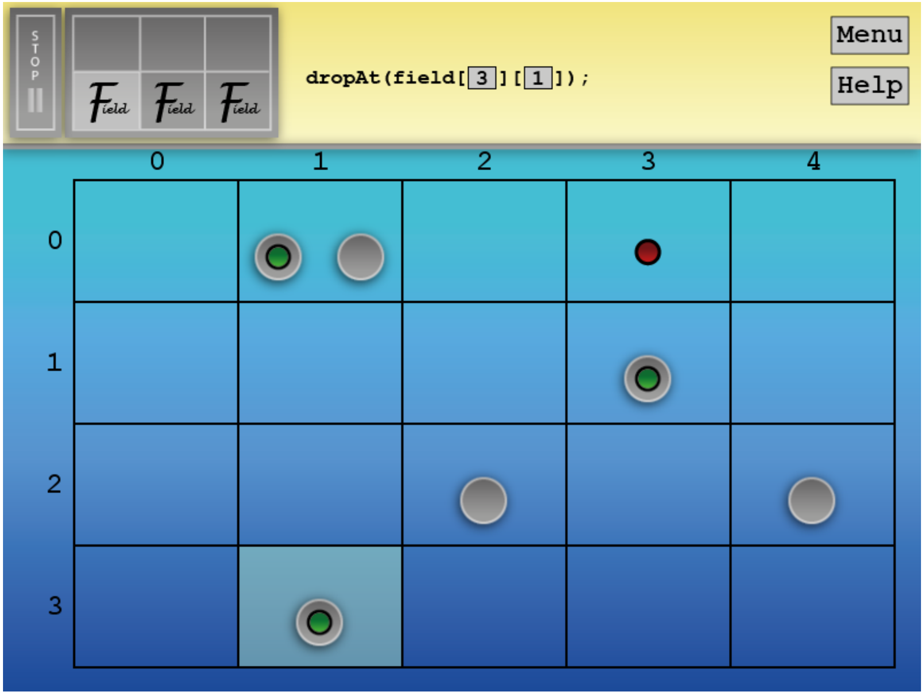


Figure 5.9: ABSTRACT Mini-game for Study 1, level 2. Circles mark the fields which should be accessed, while the field with two circles would be accessed twice. The screenshot shows how feedback is given, by revealing the results for each piece of code one after another. So far, there has been one error, indicated by the red dot.

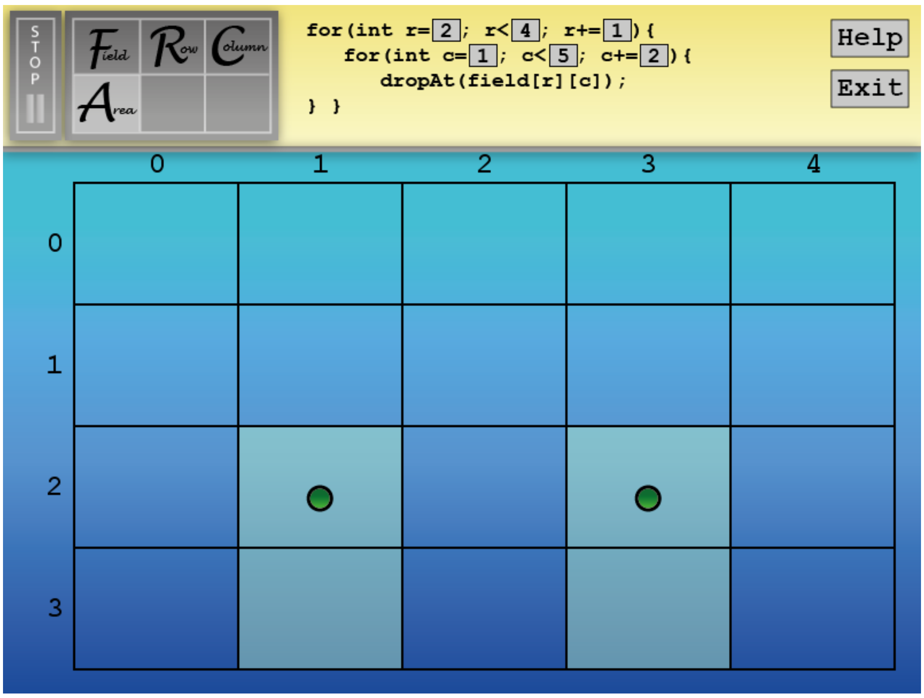


Figure 5.10: E-learning tool for Study 1, similar to the ABSTRACT Mini-game, visualising the area access of the 2D array with an ongoing animation showing the order in which fields are accessed by placing green dots on the fields.

Comparability of games

The levels of all three games are exactly the same, to ensure that the participants in the study will all come across the same pieces of code. While in the PUZZLE game, the planes in a level collectively build the target pattern, in the ACTION game the same set of planes will be solved individually, one after another. This is not fully ideal for the ACTION game, since an error is penalised by a level restart and if the code attached to the planes does not change randomly, the player can memorize the revealed solution. However, since there are often five or six planes in a level it may take a while to remember all solutions, to make it to the end of the level.

Time restriction during study

Due to the limited time during the study, hints were given in the PUZZLE and ABSTRACT game, where normally the player would have been given as much time as needed. After a fixed amount of time, planes which had already been solved correctly were revealed and fixed, so that they could not be altered to a false solution any more. If none of the planes was solved correctly, a solution for a plane would be revealed instead. Without the time restriction, hints would only have been offered on a voluntary basis to the player.

As already mentioned, the solutions in the ACTION game may be memorized by the player when making a mistake. Thus it was possible to finish the game in a limited time frame.

For both games the described limitations were not optimal, but necessary to ensure that participants would finish the study in time.

5.4.2 Games for Study 2: differing by genre

In Study 1, single game attributes are compared, which is why the games tested against each other are from the same game genre. In Study 2 the aim is to compare two games of different genres, to see how much their difference influences the learning outcome and player experience. On the Game Genre Map the two games are therefore placed further apart (Figure 5.11). One game is a modification of the ACTION Mini-game from Study 1, which is leaning towards the Action genre. The other game is an Adventure game which shares the Challenge *puzzle* with the

PUZZLE game from Study 1 but has additionally a story (not just an introductory background story). Both games will be described in the subsequent sections.

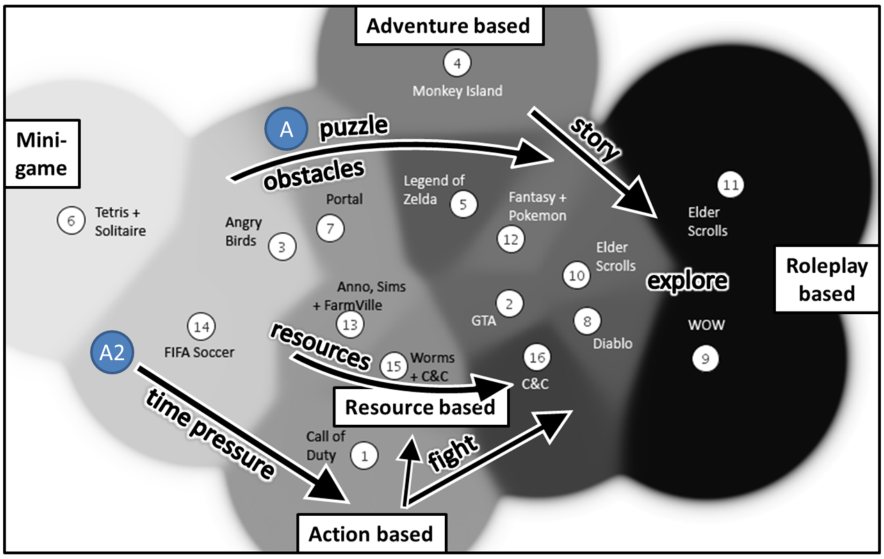


Figure 5.11: Location of games designed for Study 2 on the Game Genre Map. A = ADVENTURE; A2 = ACTION 2.

ACTION 2 Mini-game

Some changes have been made to the ACTION Mini-game produced for Study 1 (Section 5.4.1). As it is no longer compared to the PUZZLE game and differs in multiple game attributes from the ADVENTURE game, the comparability restrictions (Section 5.4.1) are no longer necessary. Thus the level design was revised by increasing the number of planes in the introductory level, practising only the newly learned concept and reducing the overall number of levels. So called bonus levels were added, for rehearsing previously learned concepts.

The ACTION 2 Mini-game has the following nine levels:

level 1: introduces concept of accessing a **field**
(variable for index of row and column)
level 2: introduces concept of accessing **whole row or column**
(for-loop plus variable for row or column)
level 3: **bonus level** for further practise of concepts learned so far

level 4: introduces concept of accessing **part of a row or column**
(variables for start and end condition)

level 5: introduces concept of accessing fields in a **row or column with a given step size**
(variable for incrementation)

level 6: **bonus level** for further practise of concepts learned so far

level 7: introduces concept of accessing a **defined area** of the array
(second for-loop plus variables for start and end conditions)

level 8: introduces concept of accessing an **area with given step size** for row and column
(variables for incrementation)

level 9: **bonus level** for further practise of concepts learned so far

As Study 2 has a within-subject design, time was even more an issue than in Study 1, since participants had to play two games instead of one. Restarting the game as penalty for a wrong solution would have taken too much time. Also, the code assigned to the planes could still not be randomised when restarting a level, as to ensure comparability, participants all needed to play the exact same game. Thus the level continues after mistakes, but a point system was used to reward correct solutions. The levels introducing new concepts can be finished early. They each have six planes and players who manage to guide three planes correctly, automatically skip any remaining planes (Figure 5.12). This prevents players from having to practise a concept longer than necessary. Depending on the number of correct and incorrect guided planes and the time it took the player to finish the level, a score is calculated and added to the total score. Bonus levels will not finish early, as their purpose is to rehearse the different concepts. With each correct solution the player earns a fixed number of points (Figure 5.12).

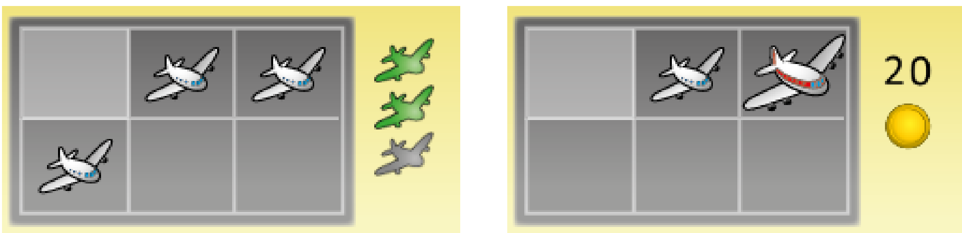


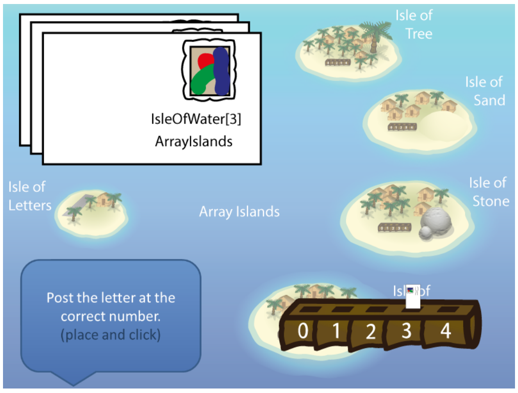
Figure 5.12: ACTION 2 MINI-GAME feedback for normal level (left), with green planes indicating the number of correct solutions so far, and bonus level (right), with points for every correct solution.

ADVENTURE

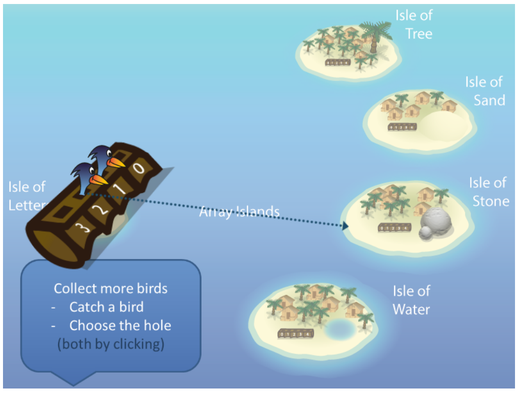
The main characteristic of an ADVENTURE game is that it is based on a story, with the player being resembled by a character in the story. Since the ADVENTURE for our study was designed to teach a rather small learning topic, it has a rather simple story line and a small Game World. The ADVENTURE genre shares some defining game attributes with the MINI-GAME genre, especially if it is of low complexity. It is also usually a *single player* game with a *fixed* Perspective on a *2D* World (Table 4.5). The learning content can be integrated following the identified Concept 3, with multiple *puzzles* being linked by a *story* (Section 5.3.2). Like the PUZZLE game in Study 1, the ADVENTURE is well suited for the constructivist learning approach. Besides the *puzzles*, learning content can also be included through the *story*. In order to develop suitable *puzzles* and an associated *story*, we thought of a metaphor for the access of a 2D array and decided to resemble arrays by postboxes with consecutive numbers as the index of the cells in the array. 2D arrays are multiple one-dimensional arrays linked by an additional array. In the game this was realised as post system for multiple islands, which each hold a one dimensional array (a post box with numbered holes for each resident of the island). Figure 5.13 gives an overview of the visualisations used in the game.

The access of a 2D array is explained in four steps through small puzzles:

- puzzle 1: explains how the access of a one dimensional array works via a given index
  (delivering a letter to a given address, while up to this stage different arrays have different names)
- puzzle 2: introduces the concept of adding a second dimension, by using one array only for linking to other one dimensional arrays
  (a bird needs to be caught for each island and placed in an additional post box to link from this box to the already existing ones)
- puzzle 3: practises the use of the two indices in 2D arrays
  (instead of an island name as in puzzle 1 an index for the bird in the linking post box is given plus the already known index for the box on the island which is accessed via the bird)
- puzzle 4: explains how the index for the access of the 2D array can be modified by one or two for loops
  (addresses need to be written on multiple letters, following given instructions on which numbers to start with, how to change them, and when to stop)

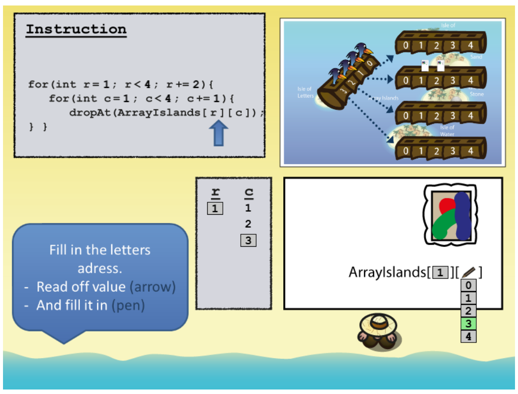


puzzle 1



puzzle 2

puzzle 3



puzzle 4

Figure 5.13: ADVENTURE MINI-GAME for Study 2; story parts alternating with puzzles.

The puzzles are enclosed by story parts which motivate the tasks in each puzzle and give a first explanation of what to do. They are also slightly interactive as the main character is talking to another character in the game and the player can select answers. For simplicity and to ensure comparability, possible answers will always lead back to the same story path. However, they allow to express opinions or ideas to solve a given problem (although very limited) and get appropriate feedback. E.g. the player may express that they understand the idea of an additional array for the second dimension, or not. They can also give feedback on how they liked a task.

The overall story is that the main character is stranded on an island, rescued by the postman, whom he then helps to deliver mail and improve the mail system between islands. In order to leave the island, the main character takes over the work of the post plane pilot, writing addresses on letters, who in exchange offers him a flight back.

5.4.3 Additional game ideas

Besides the DEGs which were built for the two different studies, there were additional ideas of how the topic 2D arrays could be taught. Especially the Resource game approach has been left out in the study and will be briefly introduced to give an overview of the variety of possible DEG designs.

Resource/Mini-game

The background story of this game is a battle on sea, where the player has to fight pirate ships. The main Challenge besides the *opponents* is *limited resources*. The resources needed to access cells in a 2D array are the coordinate values (or variable values in the for-loop). Different code frames are given, represented by different types of ship, which can shoot a canon at one field, a row, a column or a whole area. The ships can be placed around a 2D array grid, which is the area on sea where the battle takes place. Pirate ships are slowly sailing over the grid and can be attacked if a ship is built in time with code that allows to shoot at the fields the pirates currently sail on. New resources (values to complete the code for a ship) may be received over time, or by successfully attacking a pirate. This idea could be used for further comparison studies in the future.

5.5 Summary

The steps to designing an educational game are as follows:

- decide on target group and general area of a topic
- decide on the learning content (e.g. by surveying the target group)
- decide on the learning goal and how this can be reached, usually by breaking it down into smaller goals
- identify game features which may allow the integration of the learning content (possibly multiple solutions)
- identify a game genre which provides the necessary features
- decide on how the features which are not needed for including the content can maximise the player experience

Chapter 6

Comparison of Digital Educational Games with Different Game Types

6.1 Introduction

This chapter approaches the overall research question pursued in this thesis: The impact of game type for digital educational games on their success in terms of learning outcome and player experience. As it is prohibitively time consuming to explore every game feature individually, we were looking for classification systems, which group games based on typical combinations of game features. Having revised the existing game genres, the commonly used classification system for games, five redefined genres were identified: MINI-GAME, ADVENTURE, ACTION, RESOURCE and ROLE-PLAY.

In Section 4.1.1 we clarified that games of different types differ at least in one game feature. To investigate the impact of game types, we therefore need to consider games which differ by one, but also games which differ by multiple features. The revised game genres help us structure the comparison. Instead of testing random sets of game features, which may not add up to a functioning game (Prensky, 2007, p.152), we rely on the genres that have proven useful for the creation of successful games.

We conducted two studies to compare game types:

- Study 1 compares games which are all of the same genre (MINI-GAMES) and differ by only one game attribute. A genre still allows for some flexibility in the choice of game attributes. Keeping the core defining attributes, single adjustable attributes (Challenge: *time pressure* or *puzzle*, Setting: *abstract* or *realistic* graphics) were varied to create three games of different types (Section 5.4.1).

- Study 2 compares games from two of the five genres, a Mini-game which leans towards the Action genre and an Adventure game (Section 5.4.2). These games vary by multiple features (e.g. Challenge: *time pressure* vs. *puzzle*, Goals: *static* vs. *various*, Actions: *no character* vs. *character*).

We particularly chose the two studies, with one comparing games within and the other one between genres, to identify how large the impact of game type can be on the success of different DEGs in terms of learning outcome and player experience. Results should give some indication about how crucial the decision of game type is when designing a DEG, and accordingly, whether it is important to further research the effects of game type, to be able to give recommendations to DEG designers. We certainly are not able to fully answer the question of which type of game to choose for a certain target group and topic, but lay the foundation for how to approach this question.

6.1.1 Aim of the studies

The question about the impact of game type on the success of digital educational games is divided in two sub-questions. Success is measured in terms of positive learning outcome as well as positive player experience, so the two general questions to be answered are:

What differences can be found between DEGs of different types in terms of:

- Learning outcome
- Player experience

While our main aim was to answer these two general questions, learning outcome and player experience measured in the experiments are likely to be influenced not only by the type of game, but also by extraneous or confounding variables. Effects of these extraneous variables need to be considered when evaluating the results. In addition to basic analysis we thus chose to also conduct multivariate analysis, taking participants individual differences into account.

Hence, we also aim to analyse if there is a relation between the success of a DEG and the following individual characteristics of learners/players, which we derived from literature as described in the subsequent Section 6.1.2:

- *Learning preconditions:* Learners preconditions for learning a certain topic can differ and may influence the learning outcome.
- *Learning style:* A learner's style of learning may be more or less suited for learning with a certain type of DEG and therefore result in a different learning outcome.
- *Play preconditions:* Players are expected to have different play preconditions based on their gaming backgrounds, which may influence their play experience.
- *Personality:* Players' personalities are presumed to play a role in their liking towards a game and thus the experience they have when playing this type of game.

6.1.2 Possible impact of individual differences

In Chapter 5 we presented the theoretical background based on which we predicted the impact of game type on learning outcome and player experience. Drawing on learning theory as well as motivation, immersion and flow theory, we discussed in detail how each game attribute can contribute to the learning and experience of the player. According to our game genre definitions we then summarized the characteristics for each genre based on their defining game attributes (Section 5.3.3).

Possible impact of individual differences between learners/players on the success of a DEG has not been discussed yet and is therefore presented in this section.

Learning preconditions

As explained in Section 5.2 the learning topic chosen for the two studies was about programming, in particular how to access cells of a 2D array in Java. There are several variables described in the literature, which were identified as predictors for students' success in learning programming (Kinnunen et al., 2007; Rountree et al., 2002; Woszczynski et al., 2005). We expect these variables to have an impact on how well students learn, when playing our DEGs. As described below, we considered the

type of precondition	variable	B	R	W	K	our model
demographic	age		x		x	n/a
	gender	x	x	x	x	
educational	math skills	x	x	x	x	x
	prior knowledge	x	x	x	x	x
	learning style	x		x	x	
perceptional	attitude		x	x	x	x
	ease of learning		x			x

Table 6.1: B:Byrne and Lyons (2001); R:Rountree et al. (2002); W:Woszczynski et al. (2005); K:Kinnunen et al. (2007)

variables “ease of learning”, “attitude to programming”, “math skills”, and “prior programming knowledge” (Table 6.1).

While often considered, the demographic variables age and gender are typically not found to be significant predictors for success in learning programming (Bennedsen and Caspersen, 2005; Byrne and Lyons, 2001; Kinnunen et al., 2007). There are some contradicting results for age (Rountree et al., 2002), however, we were not able to include this variable, because participants were students from the first semester at university and hence all of a similar age.

Educational preconditions frequently researched for their impact on success in learning programming are math skills, prior programming knowledge, and student’s learning style. Math skills were confirmed as a predictor for success, given a positive correlation (Byrne and Lyons, 2001; Kinnunen et al., 2007; Woszczynski et al., 2005). Furthermore, significantly higher success rates in programming courses were found for students who already had programming knowledge when starting the course (Byrne and Lyons, 2001; Rountree et al., 2002).

Learning style was considered by Byrne and Lyons (2001), but no significant results were found. We thus do not take learning style into account as a variable which could influence the learning precondition. While it is not added as sub-variable for the covariate learning precondition, learning style is added as an independent covariate, since there are considerations that learning style has a direct impact on the learning outcome for different types of DEGs, which will be discussed subsequently.

Rountree et al. (2002) showed that students’ expectation (e.g. keenness, expected difficulty) was a strong indicator for how well they would do in a computer science course. Woszczynski et al. (2005) and Kinnunen et al. (2007) support the findings for attitude having an influence on success in computer science, by referencing to relevant literature.

Learning style

As mentioned in the Introduction (Section 1.3), a research study was conducted by Hwang et al. (2013) on how different types of games can support different learning styles. They compared two DEGs, each of which supported one of the two contrasting styles “global” and “sequential”, as defined by Felder and Spurlin (2005). According to their findings, players whose learning style was met by the DEG showed a better learning outcome. Following Hwang et al.’s (2013) example, we also decided to measure and consider the participant’s learning style by Felder’s scale, which in total has four dimensions. In this section, we will give an overview on different measures for learning style and further justify our choice for Felder’s scale.

The theory of learning styles is built upon the assumption that not everyone learns the same way and individual’s strengths lie in different approaches to learning. A very extensive review on 71 models of learning style is given by Coffield et al. (2004a,b). We only give a brief description of some of the most popular ones, especially Felder’s learning styles (Felder and Spurlin, 2005), which we selected for measuring the learning style of the participants in our studies.

Kolb: In 1984 Kolb introduced his experiential learning theory, which describes the process of experiential learning in a cycle with four consecutive stages (Kolb and Kolb, 2005, p.3):

- *Concrete Experience (CE):* In the CE stage the learner approaches a task and obtains concrete experiences.
- *Reflective Observation (RO):* Following the CE stage, the learner observes the experiences made during this stage and reflects upon them by formulating the insights gained during the active experimentation (AE, see below).
- *Abstract Conceptualisation (AC):* The learner aims to derive more general concepts from the observations made in the previous RO stage. This may be supported by prior knowledge and additional material like text books.
- *Active Experimentation (AE):* The concepts from the precedent AC stage can then be tested in a new situation to explore their validity. Closing the circle, the testing will evoke new experiences which leads back to the CE stage.

Kolb's learning theory assumes learners' active involvement as knowledge is gained by making experiences and drawing upon them. Performing the cycle, the learner can derive new abstract concepts, adjust and refine them.

Based on the experiential learning theory, Kolb further developed the learning style inventory (Kolb and Kolb, 2005), a questionnaire with twelve items, to investigate learners individual strength in certain stages of the circle. Concrete Experience and Abstract Conceptualisation are seen as divergent and build one axes of measurement, while similarly Reflective Observation and Active Experimentation are considered opposites and build the second axes. The questionnaire identifies where learners are positioned on both axes, which means that they fall into one quadrant between the two axes and thus between two stages. This results in Kolb's identification of the following four learning types (Kolb and Kolb, 2005, p.5):

- *Diverging style:* This learning style represents people who are located between CE and RO. They like to analyse situations and gather information but are also thought to be good in working with others and generating ideas.
- *Assimilating style:* Located between RO and AC, people with this learning style are good at theorizing from information and are less interested in the practical aspects.
- *Converging style:* Practical application is what people with this learning style are interested in, who are located between AC and AE. Applying the theoretical findings, their strength lies in actively finding solutions to problems.
- *Accommodating style:* Located between AE and CE, people with this learning style are furthest away from abstract concepts and rely more on feelings or intuition.

Myers-Briggs: Based on Jung's theory of psychological types (Jung, 1921), Myer and Briggs constructed the Myer-Briggs type indicator (MBTI) to identify individuals' perception preferences. After gradual development, starting in the 1940s, the first manual accompanying the indicator was published in 1962 (Myers, 1962) and with continued research till today, the current standard form M is a 93 item questionnaire (Schaubhut et al., 2009).

The MBTI identifies individuals four preferences out of opposing pairs, resulting in 16 possibly types. They are equally valuable, as one is not more preferable than the other and results allow respondents to be aware of their personal preferences, possibly supporting decisions, e.g. regarding career choices. The four pairs are as described by Quenk (2009):

- *Extraversion (E) - Introversion (I):* Describing where individuals draw their energy from, which can be more through contact with the outer world, activities, and social interaction, helping them process their thoughts (E), but also from the inner world, spending time reflecting upon thoughts before expressing or acting upon them (I).
- *Sensing (S) - Intuition (N):* This pair is concerned with the input of information. Senses are the input channels used for every perception. Individuals may either be more interested in current reality, easily memorizing facts and details (S), or be more taken to connect information, reveal patterns and theories allowing individuals to imply future possibilities (N).
- *Thinking (T) - Feeling (F):* Concerning the judgement of individuals based on which decisions are made. These can be more objective and logical with effects on themselves and others being secondary (T), whereas others prioritise harmonizing situations, considering people's feelings and personal values (F).
- *Judging (J) - Perceiving (P):* This pair expresses how individuals handle tasks from the outer world. Either with the need of structure, being punctual and disliking distractions (J) or more flexible, able to work under pressure, while likely to put off tasks in favour for diversion and more input (P).

While the MBTI describes different types of personal preferences and therefore has been compared with other personality measures like the Big Five (described below), its focus lies on the perception of information and the way individuals process and act upon it. In this it is also valuable for capturing different styles of learning and has been used in this regard (e.g. Rosati et al., 1988). It is also listed in Coffield et al.'s (2004a) review on learning styles as one of the main models.

Learning modalities: Several authors consider the influence of sensory modalities and perceptual preferences on the learning style (e.g. Barbe and Milone Jr, 1981; Dunn and Dunn, 1979; Fleming and Mills, 1992). Their claim is that individuals have different preference or strength for a certain form of perception, which can either be one, or a mixture of modalities. Most often the three modalities visual (V), auditory (A) and kinaesthetic (K) are mentioned, but also tactual (T) (Dunn and Dunn, 1979) and read/write (R) (Fleming and Mills, 1992).

- *Visual:* Learners with strength for the visual modality benefit from graphical representations of learning material, like summarizing diagrams, images and demonstrations.
- *Auditory:* Auditory learners prefer to perceive information by listening, e.g. to explanations from a presenter or via discussions with others.
- *Kinaesthetic:* Physical experiences are more relevant to learners with a kinaesthetic preference. This requires practical experimentation to be able to touch and feel things.
- *Tactual:* While sometimes included in the kinaesthetic modality, tactual can be used to separate fine motor skills (touching things) from experiences made by body movement.
- *Read/Write:* Reading and writing requires the visual channel, but like the auditory modality is very language oriented. For this reason Fleming and Mills (1992) proposed an additional modality for the written word.

Felder: Felder and Silverman (1988) developed a learning style model which as they indicated is "neither original nor comprehensive". From different other models and theories they tried to identify the dimensions which they considered most relevant for studying learning preferences of engineering students. The model consists of the following four dimensions (Felder and Spurlin, 2005):

- *Active/Reflective:* This dimension is analogous to Kolb and also related to Myer-Briggs extrovert/introvert. Active learners prefer to try things out and work in groups, while reflective learners need the opportunity to think things through.
- *Sensing/Intuitive:* Adopted from Myer-Briggs model and similar to Kolb's concrete/abstract axis, this dimension differentiates between sensing learners who prefer facts and are more practical oriented and intuitive learners who are more interested in theory and abstract concepts.
- *Visual/Verbal:* The previously described learning modalities were also included as one dimension. Ignoring the less relevant kinaesthetic modality, the visual preference indicates that learners benefit from visual representation, whereas verbal is a combination of the auditory and read/write modality and indicates learners' preference for language as a transmitter for information.

- *Sequential/Global:* This dimension is based on multiple sources (Felder and Spurlin, 2005) and describes whether a learner prefers a sequential approach to learning, which means that learning content is best perceived in small sequential steps or a global approach, where after perceiving multiple chunks of information, everything falls into place.

Felder and Silverman's (1988) approach on learning style was chosen to be used in the studies presented in this chapter for the following reasons: (1) It was developed for engineering students, an area which is highly related to computer science; (2) it is a mixture of several models, where supposedly the dimensions most relevant for engineering students were chosen.

Criticism: Even though various different models for learning styles exist, of which some are even highly popular and have influenced teaching methodologies at educational institutions, they are also criticised. While some analyse the credibility of instruments used to measure styles for the individual models (Coffield et al., 2004b), others claim that credible evidence for the existence of learning style is missing (Pashler et al., 2009). It may indeed be difficult to capture the complex process of learning in a defined number of different learning styles. Nonetheless some instruments show validity, consistency, and reliability (Coffield et al., 2004b, p.57) and thus they enable the identification of certain dimensions relevant to learning.

Play preconditions

We found two variables - game preferences and prior game experience - which we expect to have an influence on player experience.

Skalski et al. (2010) found that "prior use of games from the same genre" was a "consistently positive predictor of enjoyment". To research on the naturalness of controllers, they conducted two studies, each with a different game, and included player's game preferences as a factor for explaining their reported enjoyment.

Besides the preference for a game genre, the prior experience with a genre, or with playing games in general, is also presumed to have an impact on the player experience. Skalski et al. (2010) reported that prior game use was a predictor for the enjoyment "that approached significance" in one of the two games they studied. Ermi and Mäyrä (2005) stated that prior experience with a certain game genre is a factor of "the contexts of a gameplay experience". Cowley et al. (2008) claimed that full control in a game is dependent on players' knowledge about genre conventions, while the sense of control facilitates the game play experience.

More convincing evidence was given by Cox et al. (2012), who showed through a study the influence of expertise on player experience. Playing the game Tetris, expert players were bored when playing low levels, but more immersed when playing the more challenging higher levels, whereas novice players were more immersed in lower levels.

Personality

Difference in personality may be a reason why player's game experience differs. Johnson et al. (2012) found correlations between personality traits and player experience. They also searched for relations between preferred game genre and personality traits, but found no significant results.

Zammitto (2010) reported a study which also aimed to link personality traits to game genre preferences. Using multiple regression analysis, she analysed how well the Big Five personality trait dimensions predicted the preference for a certain genre. Out of twelve genres (definitions based on Rollings and Adams (2003)), eight showed significant results, however the predictive power was quite low (between 4.7 and 7.3%).

Bateman et al. (2011) pointed out that the psychometric models might have limitations for the research on games and argued for the development of a more specific player typology. However, they also stated that a robust player model is only just emerging, so we decided to use the established Big Five personality trait theory in our studies, which is further explained below.

Big Five personality traits: John et al. (2008) gives a detailed overview on the history of the Big Five personality traits. Based on the theory that a language's vocabulary should allow one to describe the most relevant aspects of a person's personality, Baumgarten (1933) and Allport and Odbert (1936) extracted a very extensive list of attributes from a dictionary. Following their work the main aim was to shorten this list and find some key distinguishing traits. Applying factor analysis, multiple sources report the finding of five factors (Costa Jr and McCrae, 1995; Goldberg, 1990; John and Srivastava, 1999). With the so-called Big Five, John et al. (2008, p.116) stated that "the field has now achieved an initial consensus on a general taxonomy of personality traits".

The following descriptions of the five factors are extracted from the John and Srivastava's (1999) BFI questionnaire and a list of items which were assigned with high agreement to the Big Five categories by judges (John and Srivastava, 1999, p. 61):

- Extraversion: People with high scores are characterised by attributes such as talkative, energetic, enthusiastic, assertive, outgoing, sociable and with low scores by attributes such as reserved, quiet, shy, withdrawn.
- Agreeableness: People with high scores are characterised by attributes such as helpful, unselfish, forgiving, trusting, kind, cooperative and with low scores by attributes such as fault-finding, quarrelsome, cold, unfriendly.
- Conscientiousness: People with high scores are characterised by attributes such as thorough, reliable, efficient, planful, organized and with low scores by attributes such as careless, disorderly, forgetful.
- Neuroticism: People with high scores are characterised by attributes such as tense, worrying, moody, nervous and with low scores by attributes such as emotionally stable, calm.
- Openness: People with high scores are characterised by attributes such as original, curious, ingenious, imaginative, inventive, artistic, sophisticated and with low scores by attributes such as commonplace, narrow interests.

6.1.3 Conceptual model

In our two studies we compared games of different types regarding their learning outcome and player experience, while also considering the impact of participant's individual differences as shown in the model in Figure 6.1. It shows the two dependent variables - learning outcome and player experience - with game type as independent variable and the several covariates we wanted to control. We found that the player experience may be influenced by play preconditions like the player's game experience and preferences for game genres, as well as the player's personality traits. For learning outcome, we identified learning preconditions like math skills and prior knowledge to be relevant, in addition to the player's learning style. We also expect to find a relation between player experience and learning outcome as further explained in Section 6.2.1.

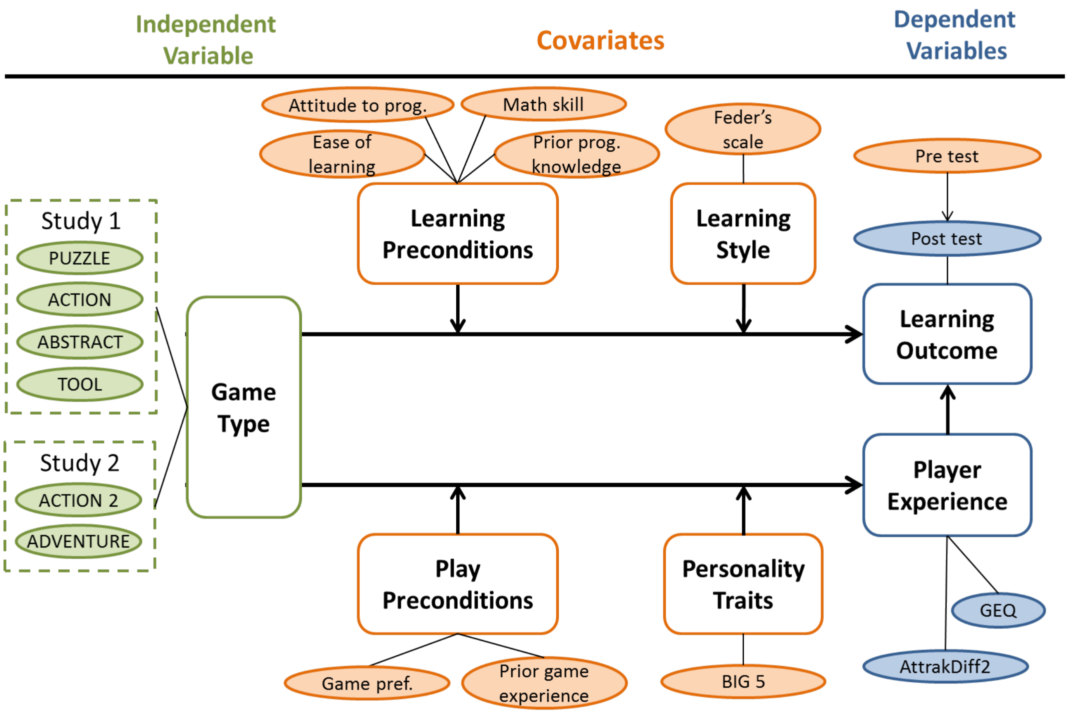


Figure 6.1: Model for impact of game type on learning outcome and player experience under consideration of covariates.

6.2 Study 1: Comparison of Game Types within Genre

6.2.1 Introduction and hypotheses

Study 1 compares three DEGs, which differ only by one game feature. According to our definition, this is the minimal requirement for games to be of different type. The three games all belong to the same genre - MINI-GAMES - with two orienting to the neighbouring genres ADVENTURE- and ACTION-game (cf. the Game Genre Map, Figure 4.8). Thus the three games are subtypes of the MINI-GAME genre. In addition, a non-game-based application was developed to be employed as a control.

Specifically, the following three games and the application were compared:

- PUZZLE: A MINI-GAME with slight orientation towards ADVENTURE, which uses *puzzle solving* as the main challenge and has *realistic graphics* (islands, planes, etc.).
- ACTION: A MINI-GAME with slight orientation towards ACTION, using *time pressure* as the main challenge.

- ABSTRACT: The same game as the PUZZLE game, but with *abstract graphics* (dots, circles, etc.).
- TOOL: An e-learning application with the same GUI and functionality as the ABSTRACT game, but without any gameplay or structures (no Goals, Challenges, or *level*), only the visualisation component.

Names assigned to the cases are referred to in the subsequent explanation and analysis of the experiment. To make them more visible they are capitalised. The games are described in detail in the previous chapter, Section 5.4.1.

The three games were designed to compare two different pairs of GEAM attributes belonging to the GEAM elements - Challenge and Setting:

- For the element Challenge the PUZZLE and ACTION game were compared to identify differences caused by using the attribute *puzzle* or *time pressure*.
- For the element Setting the PUZZLE and ABSTRACT game were compared to identify differences caused by using the attribute *realistic graphics* or *abstract graphics*.

The difference in Challenge between the PUZZLE and ACTION game led to a slightly different approach in teaching the topic. While the PUZZLE game trains writing code, the ACTION game trains reading code. The reason for this difference is that for writing code variables need to be set, which works well with the explorative *puzzle* approach, but not with *time pressure*, as further explained in Section 5.4.1.

The TOOL case was added as a control group who did not play any of the games, but had the opportunity to increase their knowledge by interacting with the TOOL.

Learning outcome and player experience

Effect of game type on learning outcome: There is no obvious prediction on which type of game is particularly suited for learning. As described in Section 5.3.3, different types of games facilitate different forms of learning, which are not necessarily better or worse. The ACTION game follows the behaviourist learning approach, which is expected to train behaviour rather than cause deep understanding. We assume this approach to be particularly suited for answering questions which are very similar to those in the game. The PUZZLE and ABSTRACT game, which follow the constructivist learning approach, are in contrast to the behaviourist approach expected to evoke better learning outcome for more advanced questions, which are beyond the scope of the tasks in the games and require the learner to grasp the underlying concepts, rather than training to answer one type of questions.

Hypothesis 1a *The ACTION game achieves a better learning outcome for knowledge questions which are very similar to the ones in this game than the PUZZLE and ABSTRACT game.*

Hypothesis 1b *The PUZZLE and ABSTRACT game achieve a better learning outcome than the ACTION game for knowledge questions which are beyond the scope of questions trained in the games.*

Another potential impact of the game type on learning outcome is derived from Malone and Lepper's (1987) statement, that fantasy can "provide appropriate metaphors or analogies". The planes in the PUZZLE game, which fly holidaymakers to their island destinations, are such a metaphor and can support students in understanding the concept, by providing a more realistic context than the ABSTRACT game. On the other hand, abstract graphics may be better suited to explain an abstract topic like 2D arrays in Java, with a more plain representation. Thus no hypothesis can be made regarding a difference in learning outcome between the PUZZLE and the ABSTRACT DEG.

Effect of game type on player experience: Literature on relations between game type and game experience is sparse. Johnson et al. (2012) investigated the relationship between game genre and player experience, using the Game Experience Questionnaire (GEQ) as measure for game experience. They identified significant differences for the GEQ dimensions Flow and Immersion, however definitions for the selected genres are not provided and rather narrow genres like "Shooting" are compared against broad ones like the combined genre "Strategy and Role-Playing". Since the studied genres are apparently different to ours, we cannot draw any conclusions from Johnson et al.'s results, other than an indication for a relation between game genres (or more broadly game types) and player experience.

In general, all DEGs designed for this study have potential to immerse the player. The ACTION game mainly through *time pressure*, and the PUZZLE and ABSTRACT game mainly through *puzzle* solving. Cox et al. (2012) reported a study, which shows an impact of the Challenge *time pressure* on game experience. Participants gave higher ratings on the level of immersion and challenge they perceived when playing the game version with *time pressure* as opposed to the one without.

In their study, *time pressure* was used to achieve an increased cognitive challenge, which should however also be achievable by the *puzzle* Challenge. Cox et al. also pointed out that their research is novel. This suggests that there has been a lack of studies, at least till 2012, on the impact of the *puzzle* Challenge in comparison to *time pressure*, which is why we cannot derive any hypothesis for a difference in player experience between the PUZZLE and the ACTION game.

Besides the impact of Challenge, we also evaluate a difference in graphical representation. In contrast to the ABSTRACT game, the PUZZLE game has more potential to immerse the player through realistic instead of abstract graphics. Although Jennett et al. (2008) argue that an abstract puzzle game can still be immersive, we derive the following hypothesis, according to Ermi and Mäyrä's (2005) imaginative immersion, requiring imaginary worlds:

Hypothesis 2a *The PUZZLE game is more immersive than the ABSTRACT game.*

Effect of player experience on learning outcome: Player experience is a variable with multiple dimensions like how much a player feels immersed, challenged, or enjoys themselves. We decided to use the Game Experience Questionnaire (GEQ) for measuring player experience, which divides the experience in seven factors: Immersion, Flow, Competence, Tension, Challenge, Positive Affect, and Negative Affect, as further described in Section 6.2.2.

The relation between the two dependent variables - learning outcome and player experience - has already been indicated in Section 5.3.1 for four of the seven GEQ factors. The feeling of Competence, which depends on an optimal balanced Challenge, facilitates intrinsic motivation (Ryan and Deci, 2000). Immersion and Flow describe how deeply the player engages in the game. Motivation and engagement are both important for learning, as motivation is what drives people to do something, like participating in a DEG, and engagement is necessary for active learning as well as attention, which according to cognitivism is required to process information.

The remaining GEQ factors are Tension, Positive and Negative Affect, which are related to the player's emotional state. Not until recently (Rienties and Rivers, 2014), the effect of emotions on learning has received much less attention in the past than cognitive and motivational processes (Astleitner, 2000). Studying the impact of attitudinal factors on learning, Giannakos (2013) found a positive relation between enjoyment (Positive Affect) and learners' performance, however not between happiness and performance. This can be due to the fact that negative emotions also have an important role in games. Hense and Mandl (2014) differentiate between activating and deactivating negative emotions. While deactivating emotions like boredom should be avoided, a certain amount of frustration can encourage the player to try again after a failure. Boredom and disinterest are characteristics of the Negative Affect GEQ factor, while frustration and restlessness are characteristics of the Tension GEQ factor.

Summarizing, we formulate the following hypotheses for the relation between player experience and learning outcome:

Hypothesis 3a *The GEQ factors Immersion, Flow, Competence, Tension, Challenge, and Positive Affect are positively related to the learning outcome (Tension only up to a medium level).*

Hypothesis 3b *The GEQ factor Negative Affect is negatively related to the learning outcome.*

Individual differences

Learning preconditions: From the literature we derive the following hypothesis for the learning precondition as reported in the introduction to this chapter (Section 6.1.2):

Hypothesis 4a *The learning preconditions ease of learning programming, attitude to programming, math skills, and prior programming knowledge are positively correlated with the learning outcome, when learning a programming topic with a DEG.*

Learning style: As the DEGs designed for our study follow different learning approaches (Section 5.4), we expect to find an impact on learning outcome for Felder and Spurlin's (2005) four learning style dimension: active/reflective, sensing/intuitive, visual/verbal, and sequential/global.

Games should be particularly good for active learners, who prefer to try things out. Especially the ACTION DEG, which uses *time pressure* as Challenge and follows the behaviourist learning approach, is more suitable for active learners, while reflective learners are not given enough time to think in this type of game.

Hypothesis 5a *Active learners have a more positive learning outcome than reflective learners for DEGs from the* Action *genre (in our case* Mini-games *with a close tendency to the* Action *genre).*

Since we chose to include the learning content by visualising how 2D arrays are accessed, visual learners should benefit more from the DEGs designed for this study (Section 5.4).

Hypothesis 5b *Visual learners have a more positive learning outcome than auditive learners for DEGs which make use of visualisation to explain the learning content.*

In addition, we decided to use levels, which stepwise introduce new knowledge, as the structure of the DEGs - an approach which should be beneficial for sequential learners. The TOOL does not provide such a guided learning and should differ from the games by being more beneficial for global learners.

Hypothesis 5c *Sequential learners have a more positive learning outcome than global learners for DEGs which structure the learning content by levels.*

For sensing as opposed to intuitive learners, we are not able to make a prediction, as programming requires both, understanding theories and being able to apply them.

Play preconditions: As discussed in Section 6.1.2 we assume the play preconditions to have an impact on the player experience, which is why we propose the following hypothesis:

Hypothesis 6a *The player's game preferences for a certain type of game are positively related with their player experience when playing this type of game.*

Hypothesis 6b *Player's prior game experience is positively related to player experience.*

Personality: Johnson et al. (2012); Zammitto (2010) did research on the relation between game type, personality, and player experience.

Johnson et al. (2012) found evidence for an impact of personality traits on player experience. For measuring the personality traits, Johnson et al. first used a short ten-item measure, but later on decided to change to the Big Five Inventory (BFI) by John et al. (1991, 2008), as a more reliable measure. For determining the player experience, they used the Game Experience Questionnaire (GEQ). Positive correlations between BFI and GEQ dimensions were found for: extraversion and challenge; agreeableness and competence, enjoyment, and annoyance; conscientiousness and challenge, competence, annoyance, and flow; openness and immersion.

Zammitto (2010) found a link between personality traits and eight out of twelve genres they researched. Genres from the study which seem to fit our DEGs best are: for the ACTION game "Action no Shooting", which showed positive relation with neuroticism, extraversion, and consciousness, and negative relations with agreeableness; for the PUZZLE and ABSTRACT game "Puzzle", which showed a positive relation with openness and consciousness.

Since Johnson et al. (2012) did not find any relations between preferred game genre and personality traits, and Zammitto's (2010) results showed only low predictive power, we keep our hypothesis for the impact of personality generic:

Hypothesis 7a *The BFI personality traits have an impact on player experience.*

It is difficult to make more precise predictions on how the impact of personality varies between the different game types. The best indications which we identified are Zammitto's (2010) results, which we aim to validate.

6.2.2 Method

Participants

The study was conducted in two consecutive years with Computer Science students learning Java programming during their first semester at the University of Leicester, UK. As reported in Section 5.2, a pre-survey had been run beforehand, to identify requirements for the DEGs and for choosing the learning topic. To refer to the different groups of students from different years of study, we assign the following code, consisting of country and year:

- academic year 2012/13: UK2012 (Pre-survey)
- academic year 2013/14: UK2013 (Study 1)
- academic year 2014/15: UK2014 (Study 1)

In the first year of the study (UK2013) 89 students participated in the experiment and 104 in the second year (UK2014). Nine and eight students dropped out from the experiment in year 2013 and 2014, respectively, and not all of them answered the prior- and post-experiment questionnaires gathering further information about learner characteristics. Table 6.2 shows how many participants answered which questionnaire. The questionnaires are all described in the subsequent "Material" section. Note that the ProgQ 1-3 was not distributed to the UK2014 cohort, as it was a short-term decision during the writing up stage of this thesis, to repeat the UK2013 study for a larger data base.

	Experiment				Learner Characteristics				
Year	KnowQ 1	App	KnowQ 2	ExpQ	ProgQ 1	ProgQ 2	ProgQ 3	CharQ	Interv
2013	89	83	80	80	88	76	62	41	4
2014	104	102	96	96	n/a	n/a	n/a	73	0

Table 6.2: Number of students participating in the experiment, starting with the knowledge pre-test (KnowQ 1), followed by interaction with the application (App) and the knowledge post-test (KnowQ 2) together with the game experience questionnaire (ExpQ); numbers of participants who filled in the three questionnaires about learning programming (ProgQ 1-3), one comprising three questionnaires - game experience, learning style and personality traits (CharQ) and attended an interview (Interv).

Comparison to results from Pre-survey: The students who participated in the DEG comparison study (UK2013/14) were different from those who completed the Pre-survey (UK2012). Results of the Pre-survey were used to identify which programming topics the students found easy or difficult and which requirements they had for a learning application. To evaluate whether there would be significant differences in the perceived difficulty of the topic "arrays" and the gathered requirements based on which the DEGs were designed, the three ProgQ questionnaires, which were used in the Pre-survey, were also distributed to UK2013.

The data (UK2013) was analysed with the same procedure as described for the Pre-survey (Section 5.2.4) and results were compared using the non-parametric Mann-Whitney U test. Table 6.3 shows the numbers of participants for UK2013. The numbers are higher than in the preceding Table 6.2, since this considered only participants which had also taken part in the experiment. Overall students from the different academic years gave very similar responses. The topic "arrays" was not perceived as more easy to learn by the new cohort of students (UK2013), who used the DEGs in the study to learn this topic, compared to the students from the Pre-survey.

	Qn#1			Qn#2	Qn#3	Qn#1-3
university	N	female	preKnow	N	N	N
Leicester	123	12	77	101	73	61

Table 6.3: Participants from the University of Leicester (UK2013) for each ProgQ questionnaire, as well as number of students who had filled in the full set of all three questionnaires (Qn#1-3).

Material

Applications: A detailed description of the three games ACTION, PUZZLE, and ABSTRACT as well as the TOOL can be found in Section 5.4.1. The DEGs and the TOOL have been developed as HTML5/JavaScript applications and can be played in any browser. To ensure consistency of representation, the DEGs were tested and compared in the most commonly used browsers (Google Chrome, Internet Explorer, Firefox) which were available to the participants. Text was included as images, since fonts differed slightly between browsers. The participants were free to use the browser of their choice. The PCs provided a dual boot OS, so students could also choose between Linux Ubuntu 12.04 (default system) and Windows 7.

The technical requirements of the applications are not demanding in terms of CPU and graphical power and all computers provided in the laboratory were suited for the DEGs to run smoothly. Although screen sizes differed slightly, the games' and tool's dimensions (800 x 600 pixel) were smaller than the screens. The applications were displayed on a black background, which filled the remaining screen space. Play conditions were thus highly similar for all participants.

Overview of Questionnaires: To collect the required data, the participants took part in an experiment, where they played the DEGs of different game types and filled in several questionnaires. An additional set of questionnaires was distributed during the semester for further background data. Table 6.4 lists all questionnaires which were used in the study and shows which variables they measured. In order to be able to assign the data from different questionnaires correctly to individual participants, students had to provide their university user names. This reason was explained and they were assured that the data would be handled confidentially and that any publication of results would be anonymised. All participation was voluntary. No bonuses were provided for taking part, nor penalties for not doing so.

instrument	abbreviation	measured variables
Questionnaires distributed during experiment		
Knowledge questionnaire Experience questionnaire	KnowQ ExpQ	learning outcome player experience
Questionnaires distributed outside of experiment		
Pre-study programming questionnaires Learner characteristics questionnaire	ProgQ 1-3 CharQ	learning preconditions play preconditions personality traits learning style

Table 6.4: Overview of questionnaires used in Study 1.

Knowledge questionnaire (KnowQ): Directly before and after playing the DEGs or using the tool, participants' knowledge was tested with the same[1] domain-specific questionnaire created by ourselves to assess the learning outcome.

Knowledge questions To evaluate the learning outcome, we developed a knowledge test for measuring students' knowledge before and after using the applications (DEGs and e-learning tool) in the experiment. The learning goal was to be able to read and write code for accessing a cell, column, row, or area in a 2D array. These are the four fundamental ways of accessing a 2D array and we created an example for each of them to test students' knowledge. To see if students were able to transfer gained knowledge and use it to solve more complex questions, we extracted questions on 2D arrays from former exams.

The test was divided in three sets of questions. The first set assesses the ability to read code and the second set to write code for accessing a cell, column, row, or area in a 2D array. The third set is to assess the ability to transfer gained knowledge.

- *Question set one (read) with four items:* Tests the ability to read code, with questions asking which fields in the 2D array are accessed by a given code. This is specifically trained by the DEG using *time pressure* (ACTION). An example for this set of questions is given in Figure 6.2. The term "field" instead of "cell" was used in the context of the games, where the 2D array is displayed as a board with fields, and was for consistency also adopted in the knowledge questionnaire.

[1] with one exception where letters to be identified in the task were changed to avoid memorisation

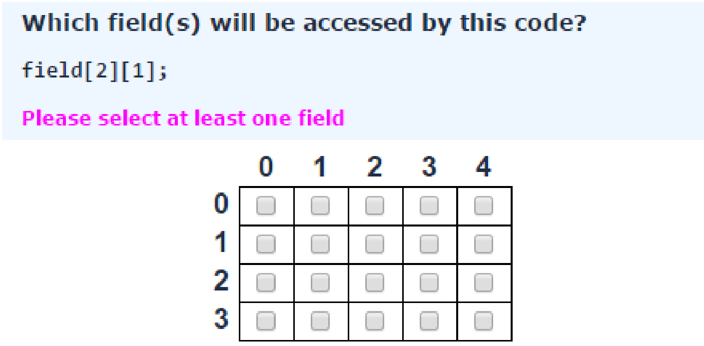


Figure 6.2: Knowledge question example for reading code (Q1.1).

- *Question set two (write) with four items:* Tests the ability to write, or more precisely complete code with questions asking to fill in the values for the variables so that a given selection of fields is accessed. The DEGs using *puzzle* solving with no *time pressure*, as well as the e-learning tool are focusing on this aspect (PUZZLE, ABSTRACT, and TOOL). Figure 6.3 shows an example for this set of questions.

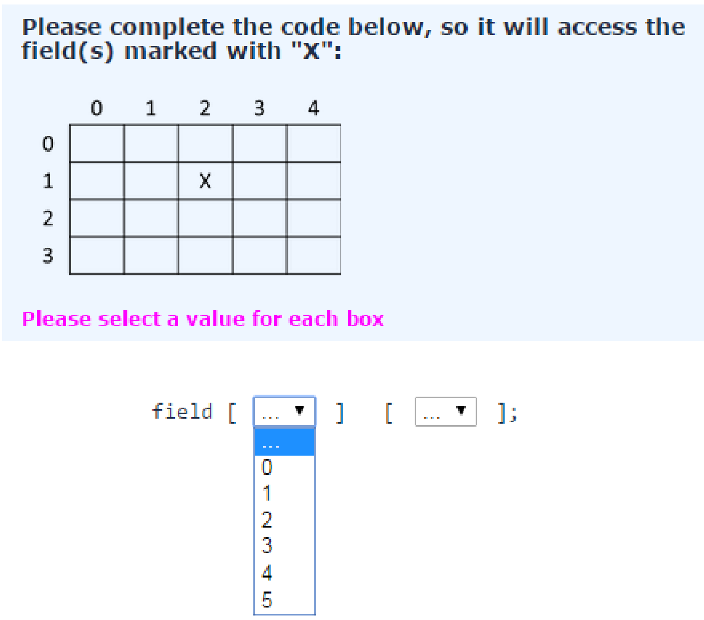


Figure 6.3: Knowledge question example for writing code (Q2.1).

- *Question set three (transfer) with three items:* Includes advanced knowledge questions to identify if knowledge can be transferred to answering more difficult questions, which are beyond the tasks in the applications. An example for these advanced questions is given in Figure 6.8.

Set one and set two each contain four questions, which is consistent with the structure of the DEGs, where four blocks of knowledge are introduced step by step (Section 5.4.1).

- *Question one:* Tests knowledge about accessing one cell in the 2D array.

Which field(s) will be accessed by this code?

```
field[2][1];
```

Figure 6.4: Knowledge question for cell access (Q1.1).

- *Question two:* Tests knowledge about accessing cells in part of one column of the 2D array.

Which field(s) will be accessed by this code?

```
for(int m = 0; m < 4; m += 3){
   field[m][1];
}
```

Figure 6.5: Knowledge question for column access (Q1.2).

- *Question three:* Tests knowledge about accessing cells in part of one row of the 2D array.

Which field(s) will be accessed by this code?

```
for(int n = 2; n < 4; n += 1){
   field[2][n];
}
```

Figure 6.6: Knowledge question for row access (Q1.3).

- *Question four:* Tests knowledge about accessing cells in a 2D area of the 2D array.

Which field(s) will be accessed by this code?

```
for(int m = 1; m < 3; m += 1){
  for(int n = 0; n < 4; n += 2){
      field[m][n];
   }
}
```

Figure 6.7: Knowledge question for area access (Q1.4).

Set three contains three advanced questions. These were derived from the questions of a former exam for the Java course of the University of Leicester (UK), where Study 1 was conducted.

- *Advanced question one:* This question is very similar to question four of the first question set (Q1.4), but combined with a print statement, to slightly increase the difficulty.
- *Advanced question two:* This question is again similar to question four of the first question set (Q1.4), with the exception that the column index is dependent on the current row index.
- *Advanced question three:* This is the most advanced question, with the least relation to the previous questions. For this question, the cell which is accessed by the current index, has to be filled with the sum of values from neighbouring cells. The question is shown in Figure 6.8.

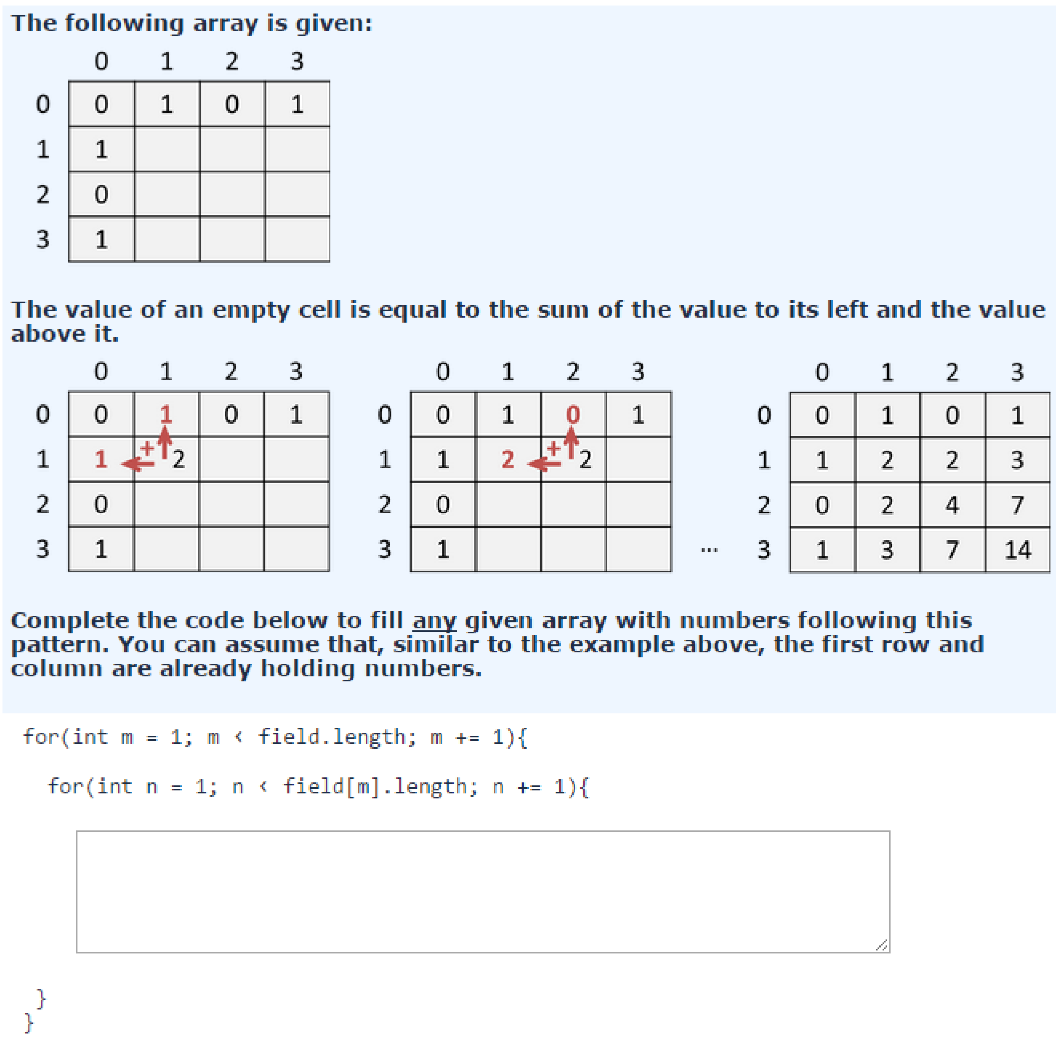


Figure 6.8: Knowledge question example for a more advanced question (Q3.3).

The questions will be referred to by Q<s>.<n> where s is the set number and n the question number in the set, e.g. Q2.3 is the third question in the second question set. Using this notation, the three question sets are Q1.1-4, Q2.1-4, and Q3.1-3.

Evaluation methods for knowledge questions Answers were analysed following two different scoring and one rating methods:

- *Method one - correct/incorrect:* The first evaluation method was simply to check for correctness of the answer. For statistical evaluation, correct answers were assigned a score of one and incorrect answers a score of zero.
- *Method two - marking, number of correct cells/values:* The second evaluation method assigns one point to every correctly picked cell for Q1.1-4 and Q3.1-2 or value for Q2.1-4. As Q3.3 has an open answer field, a high variation in answers is possible, which is why it was again only marked for being correct or incorrect. According to the correct number of cells to be picked or the number of values to be chosen, a maximal score can be achieved for each question, as listed in Table 6.5.

  For Q2.1-4, a fixed number of variables has to be selected, which each can be either right or wrong. For Q1.1-4 and Q3.1-2 however, too many cells can be selected, which needs to be considered in the marking. Thus points were given for every correctly selected cell and subtracted for every additional cell above the correct number (with a minimal score of zero). E.g. the code in Figure 6.5 accesses two cells. If three cells were selected of which two were the correct answer, the score is 2 (correct) - 1 (incorrect) = 1.

  This scoring method enables a more fine-grained analysis through incremental grading than the first one. A drawback is that it does give points for answers, which are mostly incorrect, e.g. if from the four cells, which are accessed by the code in Figure 6.7, three were selected wrongly and only one correctly, there is a high chance that the single correct cell was just a lucky guess, but the answer would still receive a score of 1. However, as the thought process of the participant is unknown, in some cases the score of 1 can be deserved, e.g.

	Q1.1	Q1.2	Q1.3	Q1.4	Q2.1	Q2.2	Q2.3	Q2.4	Q3.1	Q3.2	Q3.3
scores	1	2	2	4	2	4	4	6	4	9	1

Table 6.5: Maximal score for each knowledge question, when evaluated with the marking method.

if the cell with the start index was chosen correctly, but mistakes were made with the step size.

- *Method three - analysis of mistakes:* The third evaluation method is aiming to identify what kind of mistakes were made, e.g. index for row and column are switched, or an index is constantly one too big. While such an analysis provides a deeper understanding of the results, it was difficult to determine the kind of problems students had. Especially if multiple mistakes were made it became hard or even impossible to clearly separate them from random mistakes. E.g. if [1][2] is the correct result, it may be switched [2][1] or both index may be one too big [2][3]. If both mistakes were made, it would result in a solution [3][2] for which it is difficult to tell if it is a combination of mistakes, or a random guess. For this reason the method turned out to be not feasible, as most students made more than one mistake and thus for most answers the type of mistakes were not detectable with enough certainty.

Comparing the three evaluation methods, the first one is the most reliable one, as the risk of misinterpreting an answer is the lowest. An answer is either correct or incorrect and distortion of the data can only be caused by lucky guesses, which due to the high number of different answer options are rather unlikely. The second method is less reliable, but instead offers more gradation. Answers which are partially wrong are still awarded marks, which leads to a more detailed evaluation. But since each value is evaluated individually, there is a higher chance for lucky guesses. The last method is the most informative but turned out to be not feasible. If most of the answers cannot be assigned to an error type with certainty, there is too much missing data for a meaningful statistical evaluation.

Experience questionnaire (ExpQ): After interacting with the application and answering KnowQ, participants were asked about their experience and evaluation of the applications with some general questions, designed by us, and the game experience questionnaire (GEQ), aiming to evaluate their game play experience (IJsselsteijn et al., 2008). As the e-learning tool is not a game, it could not be evaluated with the GEQ. To evaluate the quality of the tool in comparison to the DEGs, the AttrakDiff2 questionnaire was additionally used (Hassenzahl and Monk, 2010). Via open text fields, qualitative data was gathered on what participants did or did not like about the applications. More detailed information on the general questions, the GEQ, and the AttrakDiff2 used in ExpQ is provided subsequently.

General questions Five general statements were presented for the basic evaluation of the applications (DEGs and e-learning tool). They were rated on a five-point scale from 1: "strongly disagree" to 5: "strongly agree".

1. I would describe the application as a game.
2. I would describe the application as a learning tool.
3. I would like to use this or similar applications in addition to the lecture.
4. I would recommend the application to my peers.
5. Overall I like the application.

The first two statements were to evaluate the participants' perception of the DEGs regarding their two components: learning and playing. Of particular interest is whether the DEGs are perceived as games or mostly as learning applications and whether the e-learning tool significantly differs from the DEGs. The neutral term "application" was used throughout the whole experiment when referring to DEGs as well as to the e-learning tool in order not to bias the responses, i.e. by suggesting that it is or is not meant to be a game. The third statement was to find out the perceived usefulness and willingness to use the applications for learning and the last two statements were about an overall positive or negative perception.

To gather further information about positive and negative aspects, participants were asked to elaborate in two comment fields what they did and did not like about the application. Using content analysis (Krippendorff, 2013) the students' answers were coded by the author and a colleague who is experienced in qualitative HCI evaluation methods. Comments were segmented and analysed for issues which were mentioned. About 20 issues were categorised independently by both evaluators for one of the applications. Results were then compared and discussed, to find an agreement on the definition of categories and the assignment of issues to categories. The same process was then repeated in steps of about 20 issues, while switching between applications and positive/negative comments. After a few iterations, an appropriate set of categories was found and evaluators had high consistency across identified issues and selected categories.

Game experience questionnaire (GEQ) The game experience questionnaire (GEQ) is a self-reported multidimensional measure to evaluate player's experience for the seven factors: Immersion (sensory and imaginative), Flow, Competence, Tension, Challenge, Positive Affect, and Negative Affect. It was developed by IJsselsteijn et al. (2008) and is one of the most frequently used standardized questionnaires for investigating game enjoyment (Mekler et al., 2014). The original version consists of 42 items of which six are optional, which are measured on a five-point scale. The In-Game Experience Questionnaire (iGEQ) is an abbreviated version with two items per factor, used as a faster measuring tool during play (van den Hoogen et al., 2008). Another version, focusing on children's game play experience is called KidsGEQ, with three items per factor (Poels et al., 2008). Aiming for a shorter but still informative evaluation of game experience, we selected three or four items (out of the original six) for each factor, mostly following the iGEQ, but also taking into account that the evaluated games were rather small. The list of items and reasoning for the exclusion of certain items can be found in Table A.8 and Table A.9. To evaluate each factor, the mean value for the associated items was calculated.

For Study 1, some of the items were slightly reworded. The control group was not testing a game but an e-learning tool, so DEGs and tool were addressed as "application" throughout the whole experiment. Therefore the term "game" was changed to "application" for its instances in GEQ items. In addition, since none of the games had a real story, but just a brief introduction of the setting to give reasoning for why to pursue the goals of the game, the term "story" was changed to "setting or theme" for GEQ item #3. GEQ item #11 was also slightly adapted, by extending "tiresome" with "and exhausting", a term which we felt was more commonly used.

Since the games which we evaluated were educational games, we expected the experienced challenge to be influenced by the cognitive load of learning about the access of 2D arrays. We therefore added the following item to the Challenge factor: "It was cognitive demanding". To make this apparent, we refer to the modified factor as ChallengeDEG.

AttrakDiff2 In our study the three DEGs are compared with an e-learning tool, which cannot be evaluated with the GEQ as it is not a game. We therefore chose an additional instrument for this study, the AttrakDiff2, a standardised questionnaire for assessing users experience with products (Hassenzahl and Monk, 2010)[2]. It is thus applicable for measuring the perceived quality of any kind of application and for comparing the three games and the tool. The short version of the questionnaire was chosen, as there was only limited time available in the experiment. The simplified AttrakDiff2 was developed by Hassenzahl and Monk (2010), consisting of ten

[2]web appearance: http://attrakdiff.de/

items, four to measure the pragmatic quality (confusing - structured, impractical - practical, unpredictable - predictable, complicated - simple), four to measure the hedonic quality (dull - captivating, tacky - stylish, cheap - premium, unimaginative - creative), and two additional ones to evaluate the overall quality and aesthetic quality (good - bad, beautiful - ugly). Each item is rated on a seven-point scale between opposing terms, expressing tendency to one or the other.

Pre-study programming questionnaires (ProgQ 1-3): A detailed description of the ProgQ 1-3 questionnaires can be found in Section 5.2.3, since these questionnaires were used in a pre-study to identify requirements for the DEGs. Comparing results from the pre-study with results from the current study allowed us to verify if the current participants had the same requirements for the DEGs as the cohort in the pre-study, based on whose responses the applications were developed.

In addition, the ProgQ 1-3 were used for measuring the covariate learning preconditions. Students' attitude to programming and perceived ease of learning programming were detected in every ProgQ questionnaire based on a five point scale. Values for the study were taken from ProgQ 2, as this questionnaire was conducted closest before the time of the experiment. In ProgQ 1, information on participants' math skills were gathered by a self-reported measure, and it was recorded whether or not a participant had prior programming knowledge.

Learner characteristics questionnaire (CharQ): To collect more information on the learners' characteristics, a post-gameplay survey was distributed, asking participants about their usage of games in general as well as assessing their learning style and personality traits. Together with demographic and personal skills information gathered with the ProgQ, this information allowed us to further analyse the results of the study regarding potential factors of learner individual characteristics influencing the success of a DEG.

The questionnaire was divided in three sections:

- *Play preconditions section:* To take players preferences for certain game genre into account, we asked them to rate how much they like the five game genre, which we identified in Chapter 4, providing a short description of each. For each genre they were given a five point scale reaching from "don't like it at all" to "like it a lot". In addition we asked participants to rate their prior game experience on a five point scale from "no experience" up to "very experienced".

- *Big Five personality trait section:* As there are slight variations between the results from different authors describing the Big Five personality traits, one source had to be chosen for the study. For the data collection we decided to use John's BFI questionnaire (John and Srivastava, 1999), for the following reasons: (1) reported validity and reliability (coefficient alpha mean .83) are similar to other established measures (NEO, TDA), while the BFI is fast and easy to complete (John et al., 2008, p. 131) (2) a German version is available with "psychometric properties similar to the original" (Rammstedt and John, 2007, p.205), which is important as Study 2 was conducted in Germany (3) the BFI was also used by related literature (Johnson et al., 2012).

- *Learning style section:* Felder and Soloman (2015) developed a 44 item inventory, called the index of learning style (ILS), to measure the four dimensions and assess an individual's learning style. Out of the two possible styles per dimension (Section 6.1.2), a preference for one is identified by answers to eleven items, depending on which one is chosen more frequently, which also shows how strong this preference is.

Interviews (Interv): For collecting qualitative data, participants were invited to a semi-structured interview subsequent to the experiment. Guided by a prepared series of questions we sought to gain deeper insight into problems encountered while playing the DEGs, suggestions for improvement, participants' general impression of the usage of such games, and more details on the perceived differences between the games. Particularly for Study 1 we asked the interviewee to play all three DEGs and compare them. Participant's interaction with the applications were recorded during the experiment and the reconstructed course of the game (or use of tool) was viewed in preparation for each interview. Although multiple participants agreed to be interviewed, only a very small number of interviews were conducted, since the selected learning topic for the DEGs was taught close to the end of the semester and most students were busy preparing for the exams. As the interview questions depended on a good memory of the experiment, it had to be conducted shortly afterwards to be accurate and we thus relied on the qualitative data gathered with the open text field questions in the ExpQ.

Design and procedure

The comparison experiment was run in a lab session, which took place after a lecture on the topic 2D arrays. The experiment followed the procedure displayed in Figure 6.9.

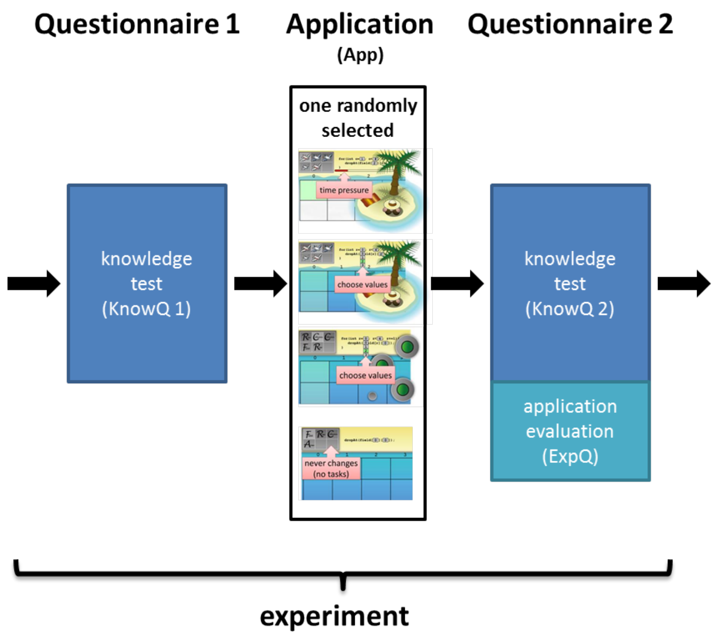


Figure 6.9: Procedure for the DEG comparison experiment in Study 1.

Since there were three DEGs to be compared as well as the e-learning application for the control group in a limited period of time, a between-subject experimental design was chosen. This is suitable due to a sufficiently large number of participants, attending the same course at the same university. Bias was reduced by randomly assigning the applications to the participants. Out of the 185 (83+102) students testing the applications (DEGs and TOOL) 51 got assigned the PUZZLE, 53 the ACTION, 42 the ABSTRACT, and 39 the TOOL. The slightly higher numbers for PUZZLE and ACTION are caused by the fact that in the UK2013 study, these two games were favoured with a ratio of 2:1 compared to the other applications, due to an uncertain number of participants. This was to ensure that at least the pair of GEAM attributes of the element Challenge could be compared, which being part of the core gameplay was assumed to have potentially more impact than the element Setting.

6.2.3 Results

Following our conceptual model as described in Section 6.1.3, we analysed the data gathered in our study regarding differences between DEGs in (1) learning outcome, (2) player experience, as well as (3) impact of participant's individual differences on both learning outcome and player experience. To validate whether there were any major issues with the games, which might have had impact on the results, we also analysed (4) the qualitative feedback participants gave on the DEGs. All statistical analyses were conducted using SPSS 22.0.

To verify if a similar amount of time was spent on each game, we measured the duration of play during the experiment. Time spent on the DEGs was on average very similar for the three DEGs (PUZZLE: mean=28:42 min, SD=7:41 min; ACTION: mean=27:31 min, SD=12:35 min; ABSTRACT: mean=30:16 min, SD=10:21 min), but significantly higher than for the TOOL (mean=6:49 min, SD=5:05 min) as indicated by ANOVA ($F(3,181)=54.371$, $p<0.01$) and confirmed by a Scheffe test ($p<0.01$ for the comparison with all three DEGs).

Learning outcome

For statistical analysis of the learning outcome, the difference between post- and pre-test scores was calculated. Based on the resulting difference scores, we further calculated the mean over all knowledge questions Q1.1-3.3 for each participant. This was done for both scoring methods - correct/incorrect and marking - but difference scores from the marking method were standardized first. To evaluate the question sets Q1-3 individually, separate mean scores were calculated for each set.

Preliminary analysis: We first conducted preliminary analysis to make sure that the data was not compromised by outliers, and to determine whether to use parametric or non-parametric statistical methods, by testing the variables distribution for normality.

Outliers The calculated mean scores were screened for outliers by inspecting the boxplots. Three cases were excluded from analysis of learning outcome as well as player experience, reducing the total number of cases to 173 (PUZZLE: 48, ACTION: 53, ABSTRACT: 34, TOOL: 38).

Normality According to the Shapiro-Wilk test, none of the calculated score mean values showed a normal distribution ($p<.05$) for all four different applications. The skewness factor of the post-test mean scores is always lower than that of the pre-test, due to the participants' overall improvement. The differences of mean values (post- minus pre-test) are all positively skewed (except for PUZZLE on the Q3.1-3 mean mark). In addition, due to the ceiling effect, the maximum score has a particularly high frequency, especially for the post-test scores. Transformation of data did not result in a normal distribution. We therefore decided to use the non-parametric Wilcoxon signed ranks and Kruskal-Wallis test, with Mann-Whitney U for follow-up analysis, to evaluate the learning outcome.

Main analysis: To identify possible differences in knowledge gain between the four different applications, a Kruskal Wallis test was used. First we compared the mean score, calculated over all answers from the pre-test, to verify that the prior knowledge level was comparable across the groups. The assumption was met, since no significant differences were found for mean pre-test scores, using both scoring methods.

The knowledge gain was calculated by subtracting pre-test from post-test scores. For each case we then built the mean score over all questions (Q1.1-3.3) as well as over each question set Q1-3. Results of the comparison with Kruskal Wallis are presented in Table 6.6.

Throughout all comparisons, the ACTION game always achieved the highest mean rank score. For the overall mean difference score, as well as for question set Q2, the results were significant, but only for the marking scoring method. Follow up analysis was conducted with the Mann-Whitney U test, which revealed that for the mean over all questions (Q1.1-3.3) the participants playing the ACTION game improved significantly more from pre- to post-test than the ones playing the PUZZLE ($U=903.50$, $p<.05$, $r=.25$), the ABSTRACT ($U=624.00$, $p<.05$, $r=.26$) game, or using the TOOL ($U=721.00$, $p<.05$, $r=.24$). For the question set Q2.1-4, participants playing the ACTION game also improved more significantly than the ones playing the PUZZLE ($U=913.00$, $p<.05$, $r=.24$), the ABSTRACT ($U=627.00$, $p<.05$, $r=.26$) game, or using the TOOL ($U=747.50$, $p<.05$, $r=.22$). As no significant results were found for question set Q1.1-4, we rejected the Hypothesis 1a, which states that the ACTION game leads to better learning outcome for Q1.1-4 - the questions most similar to the tasks in this game.

		Comparison pre-/post-test difference between apps			
		overall	for each question set		
scoring	value	Q1.1-3.3 mean	Q1.1-4 mean	Q2.1-4 mean	Q3.1-3 mean
correct	χ^2	6.505	3.779	6.395	1.767
	mr PUZZLE	79.95	82.58	80.72	83.19
	mr ACTION	100.93	97.72	99.30	91.42
	mr ABSTRACT	77.40	81.00	75.63	82.97
	mr TOOL	85.07	83.00	87.95	89.26
marking	χ^2	9.437*	4.749	8.896*	4.131
	mr PUZZLE	79.17	82.88	79.59	81.80
	mr ACTION	104.58	99.23	103.84	98.00
	mr ABSTRACT	78.32	81.90	78.03	78.62
	mr TOOL	80.14	79.72	80.89	85.72

Table 6.6: Comparison of knowledge gain (post-test minus pre-test score) between the four applications (df=3), based on the mean score for all questions, as well as each question set, using the Kruskal-Wallis test; N=173, mr=mean rank, *=p<.05.

To further investigate the reason why the ACTION game players received better results, we analysed the post- minus pre-test score difference for each knowledge question individually, but only considered the marking scoring method. Except for the most advanced question Q3.3 ACTION had the highest mean rank for every question. Significant differences between the applications were found for Q2.2 ($\chi^2(3)$=9.483, p<.05), Q2.4 ($\chi^2(3)$=8.646, p<.05), Q3.1 ($\chi^2(3)$=9.580, p<.05), and close to significant results for Q1.4 ($\chi^2(3)$=7.159, p=.067). The question Q1.4 and Q2.4 are both about accessing an area of cells in a 2D array. Question Q3.1 is very similar to question Q1.4 (Figure 6.7), by only adding a print statement to the task. From the basic ways of accessing a 2D array, which were taught by the DEGs, the access of an area is the most complex one, using a double for-loop and six variables. Since the ACTION game had the highest scores for all other basic ways of accessing cells in 2D arrays (Q1.1-3, Q2.1-3), it is likely that this higher improvement paid off when learning the most difficult area access, which combines the knowledge on accessing a cell, row, and column. The behaviourist approach on learning would thus prove to be more effective for learning such small entities of knowledge in limited time. For the most advanced question Q3.3, the participants using the ACTION game even had the lowest mean rank. This would be in line with our assumption that the constructivist learning approach is more effective for a deeper understanding, which can then be applied to solve more advanced tasks. However, the results were not significant and thus we rejected the Hypothesis 1b.

Interestingly, no significant differences were found between the TOOL and the ABSTRACT game. They both have the same graphics, interface, and interaction functionality, but the ABSTRACT DEG has additionally a level structure, goals, and gives feedback on the correctness of player's solutions. This implies that an e-learning tool can be as effective as a DEG. However it could be that the TOOL was used somewhat similar to a DEG, as the players may have aimed for self-defined goals and self-checked their solution by the animation option, which shows the access of cells according to the code. This assumption is supported by the findings on player experience as reported below, which show that some participants had perceived the TOOL as a game.

Player experience

This section is on the evaluation of the player experience, which was measured by five general questions about the applications, as well as the GEQ and the AttrakDiff2 (Section 6.2.2). In preparation for the analysis, the mean values for the seven factors of the GEQ and the pragmatic and hedonic quality factor of the AttrakDiff2 were calculated.

Preliminary analysis: Preliminary analysis was conducted to detect outliers, and to test the variables for normal distribution to decide on the type of statistical methods to be used.

Outliers Inspecting the boxplots for each variable and screening the data for irregular answer patterns, we found five cases which were excluded.

After excluding the identified outliers and conspicuous answers, 171 cases (PUZZLE: 48, ACTION: 52, ABSTRACT: 32, TOOL: 39) remained for analysing the AttrakDiff2 results, and without the TOOL 132 cases for analysing the GEQ results for the DEGs.

Normality All of the general questions had answer distributions which were negatively skewed. The Shapiro-Wilk test confirmed, that none of them were normally distributed.

The Immersion, Competence, Positive Affect and ChallengeDEG factor of the GEQ and the hedonic quality factor of the AttrakDiff2 showed a normal distribution amongst the groups using the different applications, according to the Shapiro-Wilk test (p>.05). A positive skewness and thus predominantly lower scores were found for the remaining GEQ factors Flow, Tension, and Negative Affect. From the AttrakDiff2 answers, only beauty was positively skewed, while the distributions of the pragmatic quality factor, as well as goodness were negatively skewed.

Since many of the GEQ and AttrakDiff2 factors were not normally distributed, we decided to use non-parametric tests for the statistical analysis.

Main analysis: Statistical tests were chosen in accordance with the preliminary analysis and results on differences in player experience between the four applications are reported below.

Results for general statements The only difference in median and mode value between the applications was found for the first general statement on how much participants agreed that the application was a game. Results confirm that the TOOL was with a median value of 3.0 and mode of 3 indeed perceived less as a game than the DEGs, which all had a median value of 4.0 and mode of 4. However, with a median of 3.0 and max score of 4, one may infer that some participants found that the TOOL was, at least to some degree, a game.

No differences in median were found for the statements on the application being a learning tool, the interest to use it, recommendation to peers, and overall likeability, which all had a median value of 4.0 and mode of 4 for the three DEGs as well as the TOOL. So overall participants agreed to these statements.

Results for GEQ Comparing the three DEGs on their GEQ factor scores, the Kruskal-Wallis test revealed no significant differences.

Although the PUZZLE game did have a higher mean rank (mr=71.65, n=48) for the Immersion factor, than the ABSTRACT game (mr=61.44, n=32), the results were not significant ($\chi^2(2)$=1.537, p>.05, n=132) and thus we rejected Hypothesis 2a, which was based on the assumption that the realistic graphics of the PUZZLE game would lead to a higher degree of immersion.

Results for AttrakDiff2 The Kruskal-Wallis test showed no significant differences (p>.05) for the pragmatic ($\chi^2(3)$=5.245) and hedonic ($\chi^2(3)$=1.054) quality, or the goodness ($\chi^2(3)$=1.441) and beauty ($\chi^2(3)$=3.135), among the four applications (n=171).

Individual differences

As described in Section 6.1.2, the dependent variables (DVs) learning outcome and player experience may not only be influenced by the independent variable game type, but also by the covariates learning precondition, learning style, play precondition and personality traits, each consisting of several variables. The literature suggests that the first two covariates influence the learning outcome, while the latter two influence the player experience. We decided to split the model and analyse both dependent variables separately, due to the following reasons:

1. The dependent variables learning outcome and player experience (GEQ and AttrakDiff2) show only low correlations, which detected by the Pearson product-moment correlation are all below 0.3. Results for the non-parametric Spearman's rho are very similar, with a maximum of 0.345 for the pragmatic quality of the AttrakDiff2. The low correlation indicates, that the dependent variables are mostly independent from each other.

2. If both dependent variables were analysed together, the covariates for both DVs had to be considered, although they only impact one DV. However, if covariates were included in the analysis, which are not expected to have an impact, this would reduce the power of the analysis, due to an increased number of degrees of freedom.

As we predicted an impact of the player experience on the learning outcome (Hypothesis 3a, Hypothesis 3b), we conducted a separate analysis to evaluate this prediction.

In addition to splitting the model (Figure 6.1) in the top half (learning outcome with covariates: learning precondition and learning style) and bottom half (player experience with covariates: play precondition and personality trait), we also decided to analyse the two covariates for learning outcome separately. The reason for this decision is the data collection. Since the experiment took part in a limited time frame, the data for the covariates was collected at a different time. The learning precondition was measured with a different questionnaire (ProgQ), than the learning style (CharQ), which is why we only have 34 cases with full datasets for evaluating the learning outcome. Since this number is too low for a meaningful analysis, we decided to evaluate the two covariates independently and thus have 67 case for learning preconditions, and 94 cases for learning styles. The covariates for player experience were both measured in the same questionnaire (ExpQ) and, after subtracting the number of missing data, 86 cases were further analysed.

To measure the learning outcome in our experiment, we used a pre- and post-test design (Section 6.2.2). There are multiple options on how to analyse pre-/post-test data (Dimitrov and Rumrill Jr., 2003). While for the previous analysis we used the difference between pre- and post-test scores, in the current analysis with covariates we decided to use the post-test results as dependent variable and include the pre-test results as additional covariate to reduce error variance (Dimitrov and Rumrill Jr., 2003; Rausch et al., 2003).

For learning outcome we report only the analysis of the marking scoring, as the correct scoring had shown no significant differences.

Preliminary analysis: To decide which statistical method to use for analysing the conceptual model, we conducted preliminary analysis.

Outliers We detected and removed outliers as described in the preceding analysis for learning outcome and player experience.

Normality Both DVs were tested for normality. The mean score for the post-test marks showed, according to the Shapiro-Wilk test, only a normal distribution for the PUZZLE game group (p=.138), but not for the groups who used the other applications (p<.05). As reported above, many of the GEQ and AttrakDiff2 factors were also not normally distributed. There are however some controversies about how stringent the non-violation of the normality assumption is, when using a (M)ANCOVA (Williams et al. (2013), Tabachnick and Fidell (2007), p.202).

For comparison between non-parametric and parametric test results, we conducted an ANCOVA for learning outcome, with the mean post-test mark as DV, game type as IV and mean pre-test mark as covariate (CV). Levene's test was not violated (p=.904) and a significant difference (F(3,168)=3.136, p<.05, η_p^2=.053) was detected between the applications. In accordance with the Kruskal-Wallis test, the ANCOVA also revealed significantly higher improvement (p<.05) for the ACTION game as compared to the PUZZLE game, the ABSTRACT game, and the TOOL. For the learning experience we also received the same results for the non-parametric Kruskal-Wallis test, as when comparing the applications regarding the GEQ factors with the parametric MANOVA. Although a significant difference between groups was detected, the follow up showed no significant differences for all of the GEQ factors. For AttrakDiff2, no significant differences were found with MANOVA.

Due to the robustness of (M)ANCOVA and the detected similar results for our data when evaluated with non-parametric and parametric tests, we decided to report the results of the parametric tests for evaluating our conceptual model (Section 6.1.3).

Results on learning outcome: Subsequently we present the results for impact of the IV game type on the DV learning outcome taking into account the covariates (CVs). Due to missing data, we conducted the analysis for the covariates separately, first for the CV learning precondition and thereafter for the CV learning style, using ANCOVA. Finally we analysed the correlation between player experience and learning outcome, using Spearman's rho.

Learning preconditions The confounding variables - ease of learning, attitude to programming, math skills, and prior programming knowledge - were considered to influence the learning outcome. The assumptions for the use of an ANCOVA were met: (1) no high correlation amongst the covariate variables, with a maximum Pearson product-moment correlation of .501 ($p<.01$), between 'ease of learning' and 'attitude to programming'; (2) linearity between CVs and DV; (3) homogeneity of regression slopes, with none of the interaction terms between CVs and IV being significant; (4) Levene's test of equality of error variance was not significant ($p=.659$).

There was no significant effect of game type on learning outcome after controlling for the learning preconditions ($F(3,58)=0.652$, $p=.585$, $\eta_p^2=.033$). Only 'attitude to programming' ($F(1,58)=5.076$, $p<.05$) and the pre-test mean mark score ($F(1,58)=75.437$, $p<.01$) had a significant impact on learning outcome, with a high partial effect size of the pre-test mean mark ($\eta_p^2=.565$), accounting for more than half of the variance in the post-test mean mark, and low partial effect size of 'attitude to programming' ($\eta_p^2=.080$). Since the earlier reported partial effect size of the game type difference without considering covariates had been low ($\eta_p^2=.053$), it is unclear, if controlling for the learning preconditions has indeed removed the effect, or if the increased degree of freedom has made the weak effect disappear.

Based on the results we rejected Hypothesis 4a, with the exception of the variable attitude to programming, for which we found an impact on learning outcome.

Learning styles Measured with Felder and Soloman's (2015) ILS (Section 6.2.2), the CV learning style consists of four factors - active/reflective, sensing/intuitive, visual/verbal, sequential/global. Except for the normality, assumptions for conducting an ANCOVA were mostly met: (1) only low correlation (<0.3) amongst the learning style factors; (2) linearity between CVs and DV only for visual/verbal; (3) homogeneity of regression slopes, with none of the interactions terms between CV and IV being significant; (4) Levene's test of equality of error variance was not significant (p=.986).

No significant effect was found for game type on learning outcome after controlling for learning style (F(3,85)=2.114, p=.104, η_p^2=.069). Also, none of the learning style factors showed a significant impact on the learning outcome. Only the pre-test mean mark effect was significant (F(1,85)=164.258, p<.01) with a partial effect size of η_p^2=.659.

Since no significant influence was found for the covariate learning style, we decline Hypothesis 5a, Hypothesis 5b, and Hypothesis 5c.

Player perception Hypothesis 3a predicts a positive correlation between player experience (GEQ factors) and learning outcome (mean difference mark between pre- and post-test), except for the GEQ factor Negative Affect, for which the correlation is expected to be negative (Hypothesis 3b). To validate the hypotheses, the one-tailed non-parametric Spearman's rho correlations were observed (Table 6.7).

	Immersion	Flow	Competence	Tension	Challenge DEG	Positive	Negative
post-pre test mean mark	.086	.182*	.079	.123	.222**	.171*	-.101

Table 6.7: One-tailed Spearman's rho correlation between learning outcome (post - pre test mean mark) and player experience (GEQ factors); N=132, *=p<.05, **=p<.01.

All correlations are very low (<0.3), with the highest and only highly significant between ChallengeDEG and learning outcome. The directions of correlations are as predicted (positive for all except for Negative Affect), but due to the weak and mostly insignificant correlations, we rejected the Hypothesis 3a and Hypothesis 3b.

Results player experience: After controlling for play preconditions and personality traits, the effect of game type on the measured GEQ factors and AttrakDiff2 factors was analysed respectively, using MANCOVA. The results are reported below.

GEQ factors To determine the necessity of step-down analysis, in case of high correlations between the DVs, the two-tailed non-parametric Spearman's rho correlation was inspected accordingly (Table 6.8). Correlations above 0.5 were found between Immersion, Flow, ChallengeDEG, and Positive Affect. This is in line with the literature presented in Section 5.3.1, where similarities between immersion and flow have been discussed, which both are believed to depend on a balanced challenge and to lead to enjoyment. As only one correlation is above 0.7, we decided to use MANCOVA without step-down analysis.

Besides the assumption of normality and absence of outliers which was discussed earlier, the following assumptions were met, as required for conducting a MANCOVA: (1) multivariate normality was tested by calculating the Mahalanobis distance, which with a maximum of 15.98 is below the critical value of 24.32 for seven DVs (Tabachnick and Fidell, 2007, p.949); (2) homogeneity of variance-covariance matrices was confirmed by the Box's test (p=.488) (3) linearity was assessed by inspecting scatterplots between each variable pair and overall confirmed; (4) absence of multicollinearity and singularity, with moderate correlations between DVs, as shown in Table 6.8.

	Flow	Competence	Tension	ChallengeDEG	Positive	Negative
Immersion	**.610****	.262**	.148	**.658****	**.748****	-.285**
Flow		.276**	.304**	.496**	**.629****	-.231**
Competence			-.123	.069	.487**	-.042
Tension				.321**	.028	.284**
ChallengeDEG					**.583****	-.205*
Positive						-.421**

Table 6.8: Spearman's rho correlation between GEQ factors; N=132, *=p<.05, **=p<.01.

According to Wilk's Lambda, differences between the three DEGs (p=.066) were only close to significant, with the GEQ factor Competence being closest to a significant result (F(2,56)=3.681, p=.031, η_p^2=.116), when considering the Bonferroni adjustment (p=0.05/7=0.007). None of the covariate variables considered in the model showed a significant impact, with p values above 0.1 except for prior game experience (p=.086) and BFI agreeableness (p=.055).

Since no significant results were found, Hypothesis 6a, Hypothesis 6b, and Hypothesis 7a were rejected.

AttrakDiff2 factors Correlations between the four AttrakDiff2 DVs were calculated, using Spearman's rho. Due to the overall moderate correlation values (Table 6.9), and the decision to use a parametric test as discussed earlier, we conducted a MANCOVA without step-down analysis. Prior to the analysis the following assumptions were assessed: (1) multivariate normality was confirmed with the Mahalanobis distance, which with a maximum value of 18.15 was just below the critical value of 18.47 for four DVs (Tabachnick and Fidell, 2007, p.949); (2) homogeneity of variance-covariance matrices was close to being violated, with the Box's test (p=.005) being just above the limit of p=.001 suggested by Tabachnick and Fidell (2007, p.252) (3) linearity was confirmed based on the scatterplots between each variable pair; (4) absence of multicollinearity and singularity was met with moderate correlations between DVs (Table 6.9).

	hedonic	goodness	beauty
pragmatic	.036	.236**	.033
hedonic		.478**	.423**
goodness			.407**

Table 6.9: Spearman's rho correlation between AttrakDiff2 factors; N=171, **=p<.01.

Wilk's Lambda showed no significant differences between the four applications as well as no impact of any of the covariates, with p values being all above 0.1.

Qualitative Feedback on DEGs

As part of the ExpQ we asked the participants to comment on what they liked or did not like about the application they had used in the experiment and analysed their answers as described in Section 6.2.2. Our interest in this qualitative feedback was mainly for the following reasons: (1) To verify that there were no major design flaws in the games, which would have distorted the results; (2) to gain more in depth understanding on the player experience; (3) to find out how the applications were perceived in terms of learning.

Categorisation of positive/negative comments: In this section we present the results of the analysis of the positive and negative comments for each application.

The comments have been categorised and an overview on number of comments per category can be found in Table 6.10. Findings for each category are presented subsequently.

	Number of comments for each application							
	Positive				Negative			
Type of comment	PUZ	ACT	ABST	TOOL	PUZ	ACT	ABST	TOOL
Invalid or generic								
Unclear meaning	1	5	4	9	2	6	1	3
Not applicable	1	0	1	1	4	0	4	7
Nothing	0	0	1	0	7	6	5	7
Everything	0	1	1	0	0	0	1	0
Missing	2	1	9	0	4	2	9	7
Usability	8	5	5	7	2	2	3	2
Graphics	2	3	0	0	3	2	1	3
Experience	7	8	1	0	17	11	7	6
Application	10	12	3	3	8	9	4	3
Instructions	5	2	0	1	2	1	2	1
Visualisation	10	4	9	14	0	0	1	0
Content in steps	1	2	4	1	0	2	0	0
Content								
In general	2	1	3	1	3	1	1	0
About arrays	11	11	10	7	2	0	2	2
About coding	0	1	0	1	0	3	1	4
Through examples	0	2	0	2	1	0	1	0
Game design decisions								
Hints	1	-	1	-	4	-	13	-
Level restart	-	8	-	-	-	4	-	-
Time pressure	-	4	-	-	-	11	-	-
Learning application	4	16	6	4	3	0	2	0

Table 6.10: Number of different types of positive and negative comments for the applications: PUZ=PUZZLE, ACT=ACTION, ABST=ABSTRACT, TOOL. - = not applicable, as it is not a functionality of the game.

Comments about the application: One main category of comments were addressing the usability, graphics, experience, or other aspects of the application, including some invalid or generic answers.

Invalid or generic answers: Some answers were either invalid, very broad, or unclear.

- *Unclear meaning:* Either there were mistakes in the comment, or formulations were used which were incomprehensible. Furthermore, sometimes it was not clear which aspect was addressed, e.g. a statement about "it" was made, without specifying what "it" was.
- *Not applicable:* Some comments were not applicable, as they did not answer the question, what the participant did or did not like about the application. The negative statements, which were found to be not applicable, were mostly related to the pre-/post-test questionnaires instead of the application.
- *Nothing/Everything:* Some participants gave the generic answer to either like or dislike nothing or everything.
- *Missing:* A few comment fields were left blank.

Usability: Positive comments were made about the easy or intuitive use of the applications interface. For the TOOL, the modification of variables and the instant representation of results were positively mentioned in two comments.

Usability issues were identified in not being able to skip the animations in the PUZZLE game, which showed the order in which the cells of the array were accessed. As the order of access is an important aspect which needs to be learned, it would be problematic to skip them. Similarly for the ABSTRACT game it was criticised that the animations are too slow. A fast forward functionality could fix the issue. There were also single negative comments on the way of how to change the variable values in the code fragment, however the small number of negative, and higher number of positive comments indicate an overall good usability.

Graphics: Positive comments about the graphics or the animations, visual appeal and colours were made.

Negative comments about the graphics were besides a general dislike due to individual's preferences, e.g. one player of the PUZZLE and one of the ACTION game found that there was too much water; a user of the TOOL found it looked unprofessional, and another player of the PUZZLE game thought the graphics were aimed at a younger audience.

Experience: Positive experiences were mentioned, like fun, amusement, finding it interesting. For the PUZZLE game, two comments stated, that the learning was more engaging and someone said about the ACTION game that it is repetitive without feeling boring.

A negative experiences which was frequently mentioned was that the game was too long and became boring, or caused a loss of concentration. The ACTION game was also referred to as being tedious and repetitive. For each of the three DEGs there was one comment which stated that the games were slightly confusing. Some found that the PUZZLE and ABSTRACT game felt like a children's application. For the TOOL, one person thought that it was a game, which he/she did not understand. In addition there were comments stating that it was not too engaging and not entertaining or captivating.

Application: Some stated that in general they liked the game or the application. More specifically the interactivity and responsiveness was mentioned as positive aspects. Further positive comments for the PUZZLE and ACTION game were their challenging nature and that it was possible to work on their own pace in the PUZZLE game.

The most frequently mentioned negative aspects for all three DEGs was, that they are too repetitive. The PUZZLE game was criticised for having too many levels. Two suggestions for improvements were made for the PUZZLE game: not to highlight the accessed cells in higher level to increase difficulty and to include an option for adjusting the difficulty. A comment for the ABSTRACT game stated that the puzzles got too hard at the end and a suggestion was made to give freedom to play around with different settings at the end, which means to provide the TOOL functionality. Main criticism for the TOOL was the missing purpose.

Comments about the learning: Another main category of comments were addressing the learning aspect of the applications, like the instructions explaining the topic 2D arrays, structure and presentation of the learning content, and how learning content and game were linked.

Instructions: Positive comments were made about the easy understandable and clear explanations before each level, with example solutions. On the other hand some students thought that the instructions could have been clearer.

Visualisation: The good visual representation of the array and how the code accesses different cells was commented on by multiple students. For the ACTION game, the graphical representation of mistakes was also positively mentioned. One of the comments for the ABSTRACT games was that it was good for visual learners.

There was only one negative comment about the visualisation of the array. An ABSTRACT game player stated that the highlights, indicating which cells were accessed by the current code, made it too easy.

Introducing new content in small steps: Students commented positively on the fact that complex ideas were broken down in simple steps, which progressed in logical order, with increasing difficulty for each level. For the TOOL the steps in which the content was presented at the beginning was also mentioned as positive aspect.

Two players of the ACTION game had similar concerns, as they did not like the strict sequential presentation of the learning content. One would have liked to be able to skip steps and the other found that the difficulty did not increase fast and high enough.

Content: Information was given about what and how student's felt they had learned.

- *In general:* Several students stated that they were learning. Others criticised that they had not learned much, were just guessing, or forgot the newly acquired knowledge right after the game.
- *About arrays:* Besides the very general comments on learning, some stated explicitly that they were learning about the foundational concepts of arrays. It was found to be particularly good for newbies, or to notice misconceptions about arrays.

  Two players stated that the PUZZLE game had not helped to learn the topic arrays. For the ABSTRACT game a comment criticised the focus on the variable values instead of the syntax and another indicated that due to prior knowledge nothing new was learned.
- *About coding:* Specific comments about what was learned were also made about the code, used to access arrays. However, a player of the ACTION game criticised to only have learned how to interpret, but not how to write code, or that the code explanation was not thorough enough (line by line). Others would have liked to learn more commands, like the usage of the arrays

".length" attribute, or how to write more complex programs. Particularly for the TOOL there were multiple comments that asked for more options to manipulate the code, like using a "larger equals" comparison, changing variable names, or assigning the value of a variable to another.

- *Through given examples:* Students also liked the examples given in the applications. A player of the PUZZLE game asked for more examples and a player of the ABSTRACT game for different examples of the same code.

Game design decisions: Basic design decisions made according to the selected genres were the *time pressure* in the ACTION game and the penalty of restarting the level for each failure. However some decision were made to ensure the comparability of the DEGs and to meet the time restrictions of the experiment. The tasks in the ACTION game would have been randomised when restarting a level, but were kept the same for comparability and to ensure that a level was passed after several tries. Similarly the hints given in the PUZZLE and ABSTRACT game would have been voluntary but the players were not given this choice in the experiment to make sure that the game would be finished in time. Particularly the compromises which we had to make for the experiment were expected to receive negative comments.

- *Hints in PUZZLE and ABSTRACT game:* One comment for the PUZZLE and the ABSTRACT game did say that it was good to receive help when stuck on an answer for too long, but several students did not like that the hints were given too quickly and automatically. They would have preferred to have a choice on when and how much hints to receive.

- *Level restart after error in ACTION game:* Again one student stated that it was good to have lots of chances to restart and think about the answer. However mostly the restart in general, but also the fact that there was no randomisation of the tasks was criticised.

- *Time pressure in ACTION game:* The time pressure seemed to encourage the learning for some students, as they stated that they liked how it forced them to think fast. For others the time pressure did hinder the learning, as they found that it broke their concentration and did not give them enough time to think. Some did not like the pressure in general, or would have liked to be able to change the speed, or to have more time.

Learning application: Positive comments regarding the use of a game for learning were that it was a good and fun way to learn, which some found better compared to the lecture. In particular the practical aspect was mentioned positively. Specific functionalities that were perceived to be helpful for learning were the manipulation of values in the PUZZLE game, the repetition which ensures that the concept is understood before moving on, and the immediate feedback given on answers in the ACTION game. For the TOOL, the practise and simplicity were mentioned positively as well as the option to select different values for the code.

When learners play a DEG, one concern is that they purely focus on achieving the goals in the game, but get around the learning, or that what they learn only helps them to be successful in the game, without being able to use this knowledge outside of the game. The concern was confirmed in two comments for the PUZZLE game, which stated that the player had only focused on solving the puzzle, but not on learning, or that they had stopped thinking about the code and directly mapped the numbers to the fields of the grid. Similarly for the ABSTRACT game someone stated that it only taught how to highlight boxes. Furthermore, one comment said that the ABSTRACT game was not as rewarding as regular programming.

6.2.4 Discussion

Regarding the learning outcome the ACTION game proved to be more successful than the other applications in our study. The ACTION game follows a behaviourist learning approach, as it repetitively trains how to read Java code for accessing cells in a 2D array, until a set number of tasks are solved correctly. Behaviourist learning is expected to train the exact tasks which are repetitively practised, which is why we assumed that the question set Q1.1-4 in the knowledge test, which were about reading code, would benefit most from the ACTION game. However, our results show that the ACTION game players performed particularly well throughout the whole knowledge tests and significantly better on the writing code part than participants training with the other applications. Consequently we rejected our initial hypothesis (Table 6.11).

Participants overall agreed to the statement that the three DEGs were games, while the TOOL was not, although some participants did seem to perceive the TOOL as being a game as well. No significant differences were found between the three DEGs for the GEQ factors, or between all four applications for the AttrakDiff2 qualities. Based on these results we assume that the *time pressure* in the ACTION game and the *puzzle* solving in the PUZZLE game were similarly challenging. Against our expectations the *abstract graphics* in the ABSTRACT game did not result in a

nr	hypothesis	result
Learning outcome		
1a	The ACTION game achieves a better learning outcome for knowledge questions which are very similar to the ones in this game than the PUZZLE and ABSTRACT game.	rejected
1b	The PUZZLE and ABSTRACT game achieve a better learning outcome than the ACTION game for knowledge questions which are beyond the scope of questions trained in the games.	rejected
Player experience		
2a	The PUZZLE game is more immersive than the ABSTRACT game.	rejected
Relation player experience - learning outcome		
3a	The GEQ factors Immersion, Flow, Competence, Tension, Challenge, and Positive Affect are positively related to the learning outcome (Tension only up to a medium level).	rejected
3b	The GEQ factor Negative Affect is negatively related to the learning outcome.	rejected
Individual differences		
4a	The learning preconditions ease of learning programming, attitude to programming, math skills, and prior programming knowledge are positively correlated with the learning outcome, when learning a programming topic with a DEG.	rejected
5a	Active learners have a more positive learning outcome than reflective learners for DEGs from the Action genre (in our case Mini-games with a close tendency to the Action genre).	rejected
5b	Visual learners have a more positive learning outcome than auditive learners for DEGs which make use of visualisation to explain the learning content.	rejected
5c	Sequential learners have a more positive learning outcome than global learners for DEGs which structure the learning content by levels.	rejected
6a	The player's game preferences for a certain type of game are positively related with their player experience when playing this type of game.	rejected
6b	Player's prior game experience is positively related to player experience.	rejected
7a	The BFI personality traits have an impact on player experience.	rejected

Table 6.11: Results for hypotheses in Study 1.

significant difference in player experience compared to the *realistic graphics* in the PUZZLE game, which we thought would be more immersive (Table 6.11).

We only found weak indications for an impact of player experience on the learning outcome, so we rejected the respective hypotheses (Table 6.11). A low impact was found for participant's attitude to programming on the learning outcome. Otherwise there was also no evidence for an impact of individual differences of the participants in learning precondition, learning style, play precondition, and personality traits on learning outcome or player experience (Table 6.11).

The qualitative analysis of the positive and negative comments which were made about the applications, revealed that there were no major issues. One interesting problem, amongst several categories, was that the difficulty level seemed not suitable for every player. The games may need to be adjusted to the skills of individual players, however it is aimed at students with low prior knowledge and participants who were too experienced were expected to more likely be bored.

6.3 Study 2: Comparison of Genre

6.3.1 Introduction and hypotheses

Study 2 compared two DEGs, which differ in game genre and thus by multiple features. Game genres are game types which are defined as elaborated in Chapter 4.

The following two DEGs were compared:

- *ACTION 2:* A Mini-game with slight orientation towards the Action genre, using *time pressure*, *static* goals, *scores*, but missing attributes like *fight*, *opponents*, *character* as would be required for an Action game. The ACTION 2 game is a modified version of the ACTION game used in the first study (Section 5.4.1), which had some design limitations due the requirement of being comparable with the other DEGs in Study 1.
- *ADVENTURE:* A DEG which matches the Adventure genre definition, by including the attributes *character*, *communicate*, *puzzle solving*, *various goals*, *no penalties*, *explore*, *story*. Being a very small version of an Adventure (short story line, not many choices, not very much to explore), the game is also close to the Mini-game genre.

Both games are described in detail in the previous chapter, Section 5.4.2. The assigned names are capitalized for visibility and will be used to address the two different cases.

Consequently the compared genres were Mini-game, with orientation towards Action games, and an Adventure game, which from its characteristics was still close to Mini-games.

Learning outcome and player experience

Effect of game type on learning outcome: Like the ACTION game in Study 1, the ACTION 2 game follows a behaviourist learning approach, while the ADVENTURE game, similar to the PUZZLE game in Study 1, is more based on a constructivist learning approach. According to our argumentation in Study 1 (Section 6.2.1), the behaviourist approach mainly trains behaviour, while the constructivist approach causes deeper understanding. Accordingly we propose the following hypotheses:

Hypothesis 8a *The ACTION 2 game achieves a better learning outcome for knowledge questions which are very similar to the ones in this game than the ADVENTURE game.*

Hypothesis 8b *The ADVENTURE game achieves a better learning outcome than the ACTION 2 game for knowledge questions which are beyond the scope of questions trained in the games.*

Effect of game type on player experience: There has been little research on the effect of game type on player experience. As described in Section 6.2.1, Johnson et al. (2012) presented a study, but their results were difficult to transfer to our study, since they chose a different set of game genres, without providing clear definitions for them.

Both games have potential to immerse the player. The *story* in the ADVENTURE game is supporting imaginative immersion (Ermi and Mäyrä, 2005), while the Challenge *time pressure* in the ACTION 2 game can lead to increased immersion as well (Cox et al., 2012). The two causes of immersion may thus lead to a similarly strong experience.

In terms of the experience of challenge, Cox et al. (2012) reported that the *time pressure* in a game led to a higher perception of challenge. While *puzzle* solving can also be challenging, the ADVENTURE game in our study has less difficult *puzzle*, as it focuses more on the *story*. The player has to solve rather simple tasks, while constantly receiving feedback, which tells her if she is on the right track. This should lead to a higher experience of challenge for the ACTION 2 game in comparison with the ADVENTURE.

Hypothesis 9a *Playing the ACTION 2 game results in a higher experience of challenge than playing the ADVENTURE game.*

Effect of player experience on learning outcome: A possible effect of player experience on learning outcome has been discussed for Study 1 and hypotheses have been proposed, which we also evaluated for Study 2.

Hypothesis 3a *The GEQ factors Immersion, Flow, Competence, Tension, Challenge, and Positive Affect are positively related to the learning outcome (Tension only up to a medium level).*

Hypothesis 3b *The GEQ factor Negative Affect is negatively related to the learning outcome.*

Individual differences

Following our argumentation in Study 1, we propose the same hypotheses for Study 2 regarding the impact of participant's individual differences on the learning outcome and player experience.

Learning preconditions: We assume learning preconditions to have an impact on the learning outcome:

Hypothesis 4a *The learning preconditions ease of learning programming, attitude to programming, math skills, and prior programming knowledge are positively correlated with the learning outcome, when learning a programming topic with a DEG.*

Learning style: We expect to find an impact of Felder and Spurlin's (2005) learning style dimension on learning outcome:

Hypothesis 5a *Active learners have a more positive learning outcome than reflective learners for DEGs from the* Action *genre (in our case* Mini-games *with a close tendency to the* Action *genre).*

Hypothesis 5b *Visual learners have a more positive learning outcome than auditive learners for DEGs which make use of visualisation to explain the learning content.*

Hypothesis 5c *Sequential learners have a more positive learning outcome than global learners for DEGs which structure the learning content by levels.*

Both DEGs in this study use visualisation and structure the content, in the ACTION 2 game by levels and similarly in the ADVENTURE game by story parts. They should thus be particularly suited for visual and for sequential learners. However, differences between the two games should be found for active versus reflective learners. The fast paced ACTION 2 game should be more suitable for active learners while the slow paced ADVENTURE game should be better for sequential learners.

Play preconditions: We assume play preconditions to have an impact on the player experience:

Hypothesis 6a *The player's game preferences for a certain type of game are positively related with their player experience when playing this type of game.*

Hypothesis 6b *Player's prior game experience is positively related to player experience.*

Personality: We expect to find an impact of the BFI personality traits (John and Srivastava, 1999) on player experience:

Hypothesis 7a *The BFI personality traits have an impact on player experience.*

6.3.2 Method

Participants

The study was conducted with Computer Science students from the Furtwangen University (Germany) during their first semester, where they were introduced to Java programming. Data was collected using the instruments described in the subsequent "Material" section. Out of the 87 students participating in the experiment, 82 full data sets were retrieved (41 for both groups, starting with different DEGs), but less for the additional information acquired in further questionnaires during the semester. Table 6.12 shows how many participants answered which questionnaire.

Experiment							
KnowQ 1	App 1	ExpQ 1	KnowQ 2	App 2	ExpQ 2	KnowQ 3	CompQ
87	87	87	87	86	86	83	85
Learner Characteristics							
ProgQ 1	ProgQ 2	ProgQ 3	CharQ	Interv			
53	54	30	11	3			

Table 6.12: Number of students participating in the experiment, starting with the first knowledge test (KnowQ 1), followed by testing two applications (App 1+2), each evaluated by another knowledge test plus experience questionnaire (ExpQ 1+2, KnowQ 2+3) and finishing with a comparison questionnaire (CompQ); as well as numbers of participants who filled in the three questionnaires about learning programming (ProgQ1-3), one about game experience, learning style and personality traits (CharQ) or took part in an interview (Interv).

Comparison to results from Pre-survey For choosing the learning topic and to identify requirements for the DEGs, a Pre-survey had been run beforehand (Section 5.2). To differentiate between students from the Pre-survey and from Study 2, we assign the following code, consisting of country and year of study:

- academic year 2012/13: GER2012 (Pre-survey)
- academic year 2013/14: GER2014 (Study 2)

To verify that the GER2014 cohort were similar in their needs for support in learning programming to the GER2012 students, based on whose response the DEGs were developed, the questionnaire series from the Pre-survey (ProgQ 1-3) was also distributed to the GER2014 students. Numbers of participation for each questionnaire in GER2014 are shown in Table 6.13 and are higher than in the previous Table 6.12, where only students which also participated in the experiment were considered. Due to an issue with the distribution of the third questionnaire , numbers here are exceptionally low.

	Qn#1			Qn#2	Qn#3	Qn#1-3
university	N	female	preKnow	N	N	N
Furtwangen	73	30	31	68	36	17

Table 6.13: Participants from the Furtwangen University (GER2014) for each ProgQ questionnaire, as well as number of students who had filled in the full set of all three questionnaires (Qn#1-3).

Due to the data being not normally distributed, we used the non-parametric Mann-Whitney U test to compare students responses from GER2012 and GER2014. Overall results were very similar, including the reported ease of learning for the topic "arrays".

Material

Applications: The ACTION 2 and the ADVENTURE game are described in detail in Section 5.4.2. The DEGs were developed as HTML5/JavaScript applications to be played in any browser. The computers provided in the laboratory (CPU Intel Core I7 4770K, 32 GB RAM, Nvidia GTX 770 4GB) ran Windows 7 and the participants could choose between three commonly used browsers (Google Chrome, Internet Explorer, Firefox). Some students did use their personal laptops instead of the provided PCs, but since the technical requirements for running the DEGs were not demanding and the games' dimensions were limited (800 x 600 pixel), play conditions were very similar.

Overview of Questionnaires: For collecting the data, the same questionnaires as in Study 1 were used, which are summarized in Table 6.4 and described in detail in the corresponding section. In the experiment a knowledge questionnaire (KnowQ) was used to measure the learning outcome and an experience questionnaire (ExpQ) to measure the player experience. Outside of the experiment we used the Pre-survey programming questionnaires (ProgQ 1-3) and a learner characteristics questionnaire (CharQ) to capture individual differences between participants.

Questionnaire adjustments compared to Study 1: The ExpQ and the KnowQ differed slightly from the ones used in Study 1.

ExpQ: For Study 2, the German translation of GEQ was used, which can be found in Nacke (2009, p.269-270). In Study 1 we had to make small adjustments to a few questions, since the experiment comprised not only DEGs as in Study 2, but also a learning tool, and since unlike the ADVENTURE game, the games in Study 1 did not have an actual story. Hence, for Study 2 the original version of the German GEQ was used. Due to the learning tool being not included in Study 2, we also removed the AttrakDiff2 questions from ExpQ, which had been used in addition to the GEQ to compare not only games, but also the games with the tool.

KnowQ: As students had to answer the same knowledge questions three times (KnowQ 1-3), the procedure was slightly changed to avoid frustration and to save time. After answering the questions the first time, participants were for the second time presented with their previous answers and could choose to keep or revise them. The same was applied for the third time, where they got presented their most recent answer. Furthermore, due to time constraints, the first and easiest of the advanced knowledge questions (Q3.1) was removed from KnowQ for this study.

Order of KnowQ and ExpQ: Another slight change compared to Study 1 was that after playing the games, the ExpQ was filled in first, before answering the KnowQ. The reason for this was to ensure that participants would remember the applications as well as possible, when evaluating their experiences, since they were testing two different DEGs this time, as described in the subsequent "Design and procedure" section.

Comparison questionnaire (CompQ): Two DEGs were compared and unlike the procedure in Study 1, each participant tested both of them. To gather data on participants' preferences for one DEG over another and their underlying reasoning, we added the Comparison questionnaire (CompQ) in Study 2. The CompQ was designed by us.

For the comparison of the two DEGs, the following three questions were asked:

- Which one of the two applications did you like better?
- In your opinion, which one of the two applications is better suited to teach access to 2D arrays?
- In your opinion, how much do the two applications differ?

For the first two questions, the answer options were: the first, second, both, or none of the applications. The third question had three answer options to indicate whether the two DEGs differed: strongly, somewhat, or barely. A comment field was provided to elaborate the answers on why players liked one game more than the other (or both similarly). The comments were evaluated by assigning them to categories, following the same procedure as for the general comments on positive and negative game aspects in the ExpQ (Section 6.2.2).

Since Study 2 is a comparison of two different game genres as defined in Chapter 4, we wanted to know if participants would be able to identify the DEGs' genre correctly. We therefore presented the definitions of the five identified genres and asked to rate how well the definitions match each DEG on a scale from 1: "does not match at all" to 5: "matches very well". An ordinal measure was chosen as a game may only belong partially to a genre (i.e. mixed genres). Based on the responses we can compare the players' perception of the genres with the designers' intended implementations of the genres, which is of particular interest, if we want to consider the target groups' genre preferences when developing DEGs later on.

Design and procedure

The comparison experiment was run in two different lab sessions, as the three courses who participated were run by two different lecturer. The procedure which was followed is displayed in Figure 6.10.

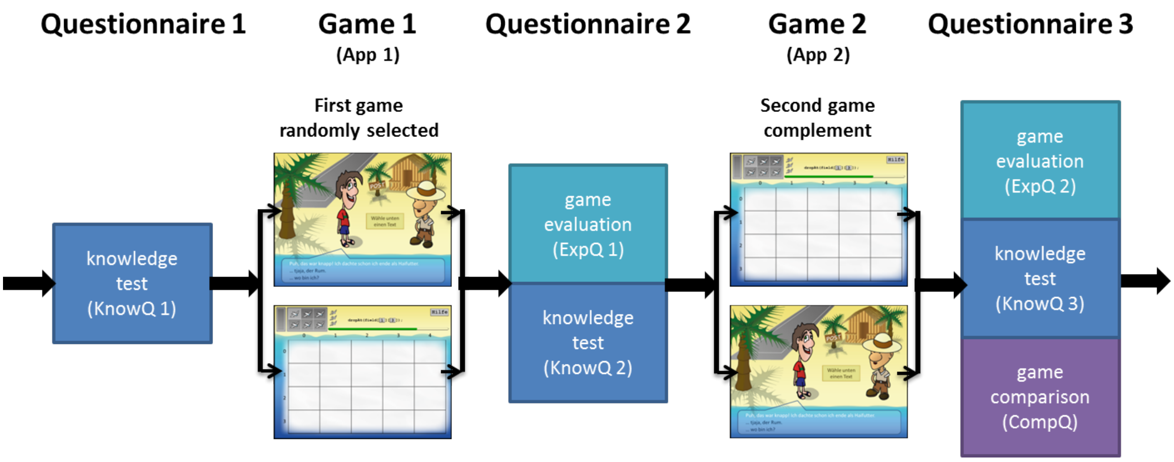


Figure 6.10: Procedure for the DEG comparison experiment in Study 2.

A within-subject design was chosen to compare the two DEGs. Because of their stronger variation in contrast to the applications in Study 1, our aim was to avoid the further influence of differences between participants. Counterbalancing was used to control the impact of order effects. The first game was randomly assigned, by alternative linking to one of the DEGs at the end of the first knowledge questionnaire KnowQ 1; the second game was then assigned accordingly. Number of participants were allocated evenly to the two cases with 44 starting with the ACTION 2 and 43 with the ADVENTURE game.

Three students did not manage to finish the experiment during the lab session, but later during the day. Knowledge answers given after leaving the lab were not taken into account, but their experience and comparison questionnaire responses were.

As described in Section 6.2.2, data on learner characteristics were collected with the additional questionnaires ProgQ 1-3 and CharQ. The latter was set up and distributed online, as it would have required too much of the lecture time for students to fill them in on paper. Unfortunately, the response rates were very low, so the game preference, learning style, and personality trait data gathered with CharQ could not be analysed in this study. This limits the number of covariates to be evaluated to the CV learning preconditions, which is assumed to have an impact on the learning outcome. Nonetheless, we used a within-subject design, therefore group differences are not an issue when analysing the effect of the game type on game experience.

The within-subject design entails the issue that knowledge will be gained when testing the first DEG, which influences the knowledge gain for the second DEG. So while the game experience can be compared within-subject, the learning outcome was also analysed between-subjects, e.g. to compare the improvement of both groups after playing the first DEG.

6.3.3 Results

For analysing the gathered date, we followed our conceptual model described in Section 6.1.3, to identify differences between the two DEGs in learning outcome and player experience. We also considered the impact of participant's individual differences, as well as the qualitative feedback based on the comments for both DEGs.

We measured how much time participants spent on each game, to see if it was similar for the two different DGEs. Time spent playing could not be measured with full accuracy, due to server caused latencies in loading the game graphics, which resulted in some waiting time for the players. Assuming that it overall affected both games similarly, statistical analysis was done with the data as recorded. Students who started with the ACTION 2 game spent on average 21:10 min (SD=6:02 min) on this DEG, and only 15:28 min (SD=3:08 min) on the ADVENTURE, which is significantly less time (related t-test: $t[41]=7.189$, $p<0.01$). Participants starting with the ADVENTURE game spent on average 16:41 min (SD=4:37 min) on this game and 16:47 min (SD=3:21 min) on the ACTION 2 DEG, which is significantly less time than the cohort starting with this game ($t[83]=4.096$, $p<0.01$).

Learning outcome

Like in Study 1, mean values were calculated for the correct/incorrect and the marking scores of each participant over all questions Q1.1-3.3, as well as for each question set Q1-3.

Preliminary analysis: To identify outliers and to test the distribution of the mean scores for normality, the data was analysed accordingly.

Outliers For detecting outliers, the boxplots for the calculated mean values of scores were observed and all conspicuous cases inspected. As no invalid answers were found, all cases were considered in the statistical analysis of the learning outcome.

Normality Except for the mean value of the score difference between KnowQ1 and KnowQ3, none of the mean scores were normally distributed, according to the Shapiro-Wilk test. Therefore non-parametric tests were used for analysing the learning outcome.

Main analysis: Participants were randomly assigned to either start with the ACTION 2 or the ADVENTURE game. Both groups played both games, but the knowledge gain caused by the second DEG was affected by what was learned in the first DEG. Therefore the learning outcome needs also to be tested between subjects, i.e. comparing the two groups regarding the first DEG they played.

To identify if both groups had equal prior knowledge, we compared the results from the first knowledge questionnaire using a Mann-Whitney U test. The groups did not differ significantly on any of the question sets (Q1-3) mean scores, for both the correct/incorrect and the marking scoring method. For the overall mean score (Q1.1-3.3), a significant difference was found with a smaller mean rank of 38.77 for ACTION 2 than for the ADVENTURE with 49.35, but only for the less fine grained correct/incorrect scoring method ($U=716.00$, $p<.05$, $r=0.22$), where differences were magnified due to dividing answers only into the two categories “correct” and “incorrect”.

Overall knowledge gain The knowledge gain was significantly higher for the first DEG that participants played than for the second DEG, regardless of whether they first played the ACTION 2, or the ADVENTURE game. This was tested for the overall score and scores for each question set with both scoring methods, using the Wilcoxon signed rank test. The only exception was the scores for the first question set Q1, which increased significantly less, when the ACTION 2 game was played first. This is due to the ACTION 2 being more effective than the ADVENTURE game in teaching how to answer Q1, as reported in the subsequent paragraph.

Comparison between DEGs, between subject For the between subject comparison, the learning outcome was calculated by subtracting the KnowQ 1 scores from KnowQ 2 scores and the KnowQ 2 from KnowQ 3 scores to detect the effect of the first and the second DEG respectively.

Using a Mann-Whitney U test, significant differences were found for the knowledge gain after playing the first game. Comparing the mean scores over all questions (Q1.1-3.3) with both scoring method, higher improvements were found for participants who played the ACTION 2 as opposed to the ADVENTURE game (Table 6.14). Analysing the question sets Q1-3 individually, we found that only the mean scores for Q1 differed significantly for both scoring methods. This confirms Hypothesis 8a, which predicted that the behaviourist approach of the ACTION 2 game would best train to answer the type of questions used in the game, which were most similar to the questions in question set Q1.

Comparison first DEG		KnowQ 2-1 score differences between DEGs			
		overall	for each question set		
scoring	value	Q1.1-3.3 mean	Q1.1-4 mean	Q2.1-4 mean	Q3.1-3 mean
correct	U	683.00*	661.50*	765.00	812.50
	mr ACTION 2	49.98	50.47	48.11	40.97
	mr ADVENTURE	37.88	37.38	39.79	47.10
marking	U	675.50*	633.00**	813.50	873.50
	mr ACTION 2	50.15	51.11	47.01	45.65
	mr ADVENTURE	37.71	36.72	40.92	42.31

Table 6.14: Comparison of the knowledge gain from the first DEG (KnowQ 2 - KnowQ 1 scores) between participants playing the ACTION 2 or the ADVENTURE game (df=1), based on the mean score for all questions, as well as each question set, using the Mann-Whitney U test; N=87, mr=mean rank, *=p<.05, **=p<.01.

Very similar results were found when comparing the knowledge gain for the second DEG played in the experiment. Again, higher improvements were found for the ACTION 2 game as opposed to the ADVENTURE game (Table 6.15), with significant results for the overall score and highly significant results for the first question set Q1, providing further evidence for accepting Hypothesis 8a.

Comparison second DEG		KnowQ 3-2 score differences between DEGs			
		overall	for each question set		
scoring	value	Q1.1-3.3 mean	Q1.1-4 mean	Q2.1-4 mean	Q3.1-3 mean
correct	U	634.00*	604.00**	731.00*	820.00
	mr ACTION 2	47.54	48.27	45.17	41.00
	mr ADVENTURE	36.60	35.88	38.90	42.98
marking	U	661.00*	610.00**	755.00	828.50
	mr ACTION 2	46.88	48.12	44.59	41.21
	mr ADVENTURE	37.24	36.02	39.48	42.77

Table 6.15: Comparison of the knowledge gain from the second DEG (KnowQ 3 - KnowQ 2 scores) between participants playing the ACTION 2 or the ADVENTURE game (df=1), based on the mean score for all questions, as well as each question set, using the Mann-Whitney U test; N=83, mr=mean rank, *=p<.05, **=p<.01.

Although the mean ranks are less different for the third question set Q3 and even slightly in favour of the ADVENTURE game, we have to reject Hypothesis 8b, which predicted that the constructivist approach of the ADVENTURE game would lead to a higher improvement in Q3.

Order effect When comparing the knowledge gain over the whole experiment using the Mann-Whitney U test, no significant differences were found between the two groups who each started with a different DEG. The order in which the games are played thus did not have an impact on how well the topic was learned.

Player experience

Preliminary analysis: To determine the validity of all cases and the statistical methods to be used for the analysis of the player experience, we screened the data for outliers and tested its distribution for normality.

Outliers Inspecting the boxplots for the seven GEQ factors for both applications, up to three potential outliers were found for each factor. Closely examining each case, no indication was found to assume that they were outliers due to corrupted data. Similarly for the general questions and the comparison questions no outliers were identified.

Normality All of the answers for the general questions were negatively skewed due to rather high ratings, and were thus not normally distributed. The comparison question, about how different the two applications were perceived to be, did also not follow a normal distribution.

Testing the distribution of the GEQ factors for normality with the Shapiro-Wilk test, about four of the seven factors were normally distributed for the two different groups and two games played by each group. To be consistent with the Study 1 analysis, and due to the reported violations of normality, we decided to use non-parametric tests to evaluate the player experience.

Main analysis: Comparing the two DEGs regarding differences in player experience, we selected statistical tests in accordance with the preliminary analysis and report results for the general statements on each DEG, the GEQ, and the comparison questions below.

Results for general statements Participant's answers to the general statements were compared by inspecting the median and mode values of the responses (Table 6.16). The data file was split by the two groups, starting with a different game, as order effects may occur.

		general statements (GSs)									
		GS1 game		GS2 learning		GS3 use		GS4 recommend		GS5 like	
first DEG	value	md	mo	md	mo	md	mo	md	mo	md	mo
ACTION 2	ACTION 2	4.0	4	5.0	5	4.0	4	4.0	4	4.0	4
	ADVENTURE	4.0	4	4.0	4	3.0	3	3.0	3	3.0	3
ADVENTURE	ACTION 2	4.0	4	4.0	4	4.0	4	4.0	5	4.0	4
	ADVENTURE	4.0	4	4.0	4	4.0	4	4.0	3	4.0	4

Table 6.16: Comparison of the five general statements (GSs) between the two DEGs ACTION 2 and ADVENTURE; md=median, mo=mode.

Results were highly similar for both games for the group of students which started with the ADVENTURE game. They only differed in the mode value on how much they agreed to recommend each game with a higher value for the ACTION 2 game as compared to the ADVENTURE game. When starting with the ACTION 2 game, results differed on four of the five statements, with the ADVENTURE game being rated lower not only on the recommendation, but also on how much students thought it was a learning tool, would like to use it, and liked it overall.

In general students agreed with the five general statements, as the median values were all above neutral. The only exception was the less well perceived ADVENTURE game when played as second DEG, which had neutral median ratings for the GS3-5.

Results for GEQ The ACTION 2 game had significantly higher ratings for five of the seven GEQ factors (Section 6.2.2): Immersion, Flow, Tension, ChallengeDEG, and Positive Affect (Table 6.17). Results for Immersion and Positive Affect were only significant for participants who had played the ACTION 2 game first. The higher experience of challenge in the ACTION 2 game confirms Hypothesis 9a, according to which *time pressure* increases this experience. The increased experience of Tension may be related to the higher level of Challenge, which would have to be confirmed by further research.

		GEQ factors						
first DEG	value	Immersion	Flow	Competence	Tension	Challenge DEG	Positive Affect	Negative Affect
AC (n=44)	Z	-3.606**	-3.471**	-1.524	-3.695**	-5.295**	-2.704**	-5.100**
	mr AC	19.07	21.12	22.82	21.53	21.49	19.40	8.00
	mr AD	13.71	20.50	17.38	12.94	11.50	16.15	21.89
AD (n=42)	Z	-0.670	-4.043**	-0.665	-5.053**	-4.706**	-1.340	-2.741**
	mr AC	17.63	21.19	14.38	20.85	21.02	15.43	14.39
	mr AD	13.37	14.57	15.50	9.83	9.50	20.60	17.98

Table 6.17: Within subject comparison of the seven GEQ factors between the two DEGs ACTION 2 (AC) and ADVENTURE (AD) (df=1), using the Wilcoxon signed rank test; mr=mean rank, **=p<.01.

The differences in Immersion and Flow are of particular interest. As discussed in Section 6.3.1 immersion is facilitated by the *story* of the ADVENTURE game and by the *time pressure* in the ACTION 2 game. According to the results, the experienced immersion was indeed similar, but only if the ADVENTURE was played first. Independent of the order in which the DEGs were player, the flow experience was reported to be significantly higher in the ACTION 2 game. This aspect needs to be further researched, as the causes for the difference in flow experience are unclear. Possible assumptions are that the ADVENTURE game was too short for the player to get deeply immersed in the *story* and experience flow, students may not have been able to relate to the *story*, or *time pressure* is a stronger trigger for flow.

The significantly lower Positive Affect rating for the ADVENTURE game when playing it as the second DEG is in accordance with the aforementioned findings for the general statements. Why the ADVENTURE game also scores significantly higher for the Negative Affect needs to be clarified, by evaluating the comments about what participants did not like about the game.

Results for the comparison When asked at the end of the experiment for their preference of game, the ACTION 2 game was clearly favoured by most students (Table 6.18). In total, 71% chose ACTION 2, 15% ADVENTURE and 14% liked them both equally, while no one voted to not like both. For learning the topic 2D arrays, more participants (27%) thought that both were suited well, while again most voted for the ACTION 2 game (67%), but only a few for the ADVENTURE (6%).

question	first app	ACTION 2	ADVENTURE	both	N
liked more	ACTION 2	32	6	5	43
	ADVENTURE	28	7	7	42
	total	60	13	12	85
better for learning 2D arrays	ACTION 2	32	2	9	43
	ADVENTURE	25	3	14	42
	total	57	5	23	85

Table 6.18: Frequencies for DEG preferences.

In average the difference between the two DEGs was rated 1.55 (SD=0.52; median=2.0), which means that they were perceived to differ between somewhat and strongly.

		Assignment to genre		
first DEG	value	MINI-GAME	ACTION	ADVENTURE
ACTION 2	Z	-4.092**	-3.679**	-3.767**
(n=44)	mr ACTION 2	12.32	9.50	17.63
	mr ADVENTURE	5.00	0.00	16.34
ADVENTURE	Z	-.456	-4.134**	-3.187**
(n=42)	mr ACTION 2	14.83	9.88	6.50
	mr ADVENTURE	14.12	6.50	14.94

Table 6.19: Within subject comparison of the assignment of the two DEGs ACTION 2 and ADVENTURE (df=1) to the different game genre, using the Wilcoxon signed rank test; mr=mean rank, **=p<.01.

The genres were identified correctly by the participants. Significant differences were found for the ADVENTURE and ACTION genre (Table 6.19) with higher ratings for the ADVENTURE and ACTION 2 game accordingly. ACTION 2 was also rated higher on being a MINI-GAME, but only when being played first in the experiment. Perceptions of the players did thus match the game genre intentions of the designer for both DEGs.

Individual differences

As mentioned before, the evaluation of the conceptual model described in Section 6.1.3 was only possible to a limited extent. Due to low response rates for the CharQ questionnaire, only the covariate learning preconditions measured by the ProgQ 1+2 could be considered. Accordingly the hypotheses concerning the covariates learning style (Hypothesis 5a, 5b, 5c), play preconditions (Hypothesis 6a, 6b), and personality traits (Hypothesis 7a) were not evaluated.

Preliminary analysis: To verify the use of ANCOVA and Pearson correlation as statistical methods for analysing the conceptual model, preliminary analysis were conducted.

Outliers As reported earlier, no outliers were found for the learning outcome.

Normality The mean scores for the learning outcome were not normally distributed. However, following our reasoning for Study 1 (Section 6.2.3), we decided to conduct a parametric ANCOVA to control for the pre-test results and the learning preconditions when analysing the effect of the game type.

As a considerably higher knowledge gain was found for the first as compared to the second DEG played by the participants, we decided to compare the two groups on the randomly assigned first DEG. Accordingly we used the mean score from KnowQ 1 (pre-test) as CV and from KnowQ 2 (post-test) as DV. With game type as IV we conducted an ANCOVA to compare the parametric with the previously described non-parametric test results. Levene's test for the ANCOVA was not violated. In contrast to the Mann-Whitney U test, no significant results were found between the two game types for the correct/incorrect (F(1,84)=2.235, p=.139, η_p^2=.026) and the marking scoring method (F(1,84)=3.505, p=.065, η_p^2=.040). As a consequence we also tested the scores for the first question set Q1, which according to the non-parametric test had differed between the game types with high significance. Again the Levene's test assumption was met. For the Q1 scores the ANCOVA results did indeed confirm this difference, as significant differences were found between the game types for both scoring methods: correct/incorrect ($F(1,84)=4.261$, $p<.05$, $\eta_p^2=.048$) and marking ($F(1,84)=5.586$, $p<.05$, $\eta_p^2=.062$). We therefore considered the overall score as well as the Q1 score for the analysis of the conceptual model.

Results on learning outcome: To evaluate the impact of learning preconditions (CV) on the game type's (IV) effect on learning outcome (DV), we conducted an ANCOVA as reported below. Furthermore, to research the relation between player experience and learning outcome, we analysed the Spearman's rho correlations between the two DVs.

Learning preconditions Considering the confounding variables - ease of learning, attitude to programming, math skills, and prior knowledge - as well as the pre-test scores from KnowQ 1, an ANCOVA was conducted, analysing the effect of game type on learning outcome (post-test scores from KnowQ 2). Number of cases with complete data for the analysis were 18 starting with the ACTION 2 and 19 starting with the ADVENTURE game. For a more concise presentation of the results, we only report the evaluation of the marking scoring method for the overall and Q1 mean scores. Assumptions for conducting an ANCOVA were met for both scores: (1) there was no high correlation amongst the covariates, with the highest Pearson correction of .577 ($p<.01$) being identified between ease of learning and attitude to programming; (2) linearity between CVs and DV; (3) homogeneity of regression slopes were confirmed by finding none of the interaction terms between CVs and IV to be significant; (4) Levene's test of equality of error variance was not significant for the overall score ($p=.596$) and the Q1 score ($p=.166$).

No significant effect was found for game type on the overall score as measure for the learning outcome, after controlling for the learning preconditions ($F(1,30)=1.716$, $p=.200$, $\eta_p^2=.054$). Also, none of the CVs had a significant impact. Considering the first question set Q1 as measure for the learning outcome does also not result in a significant effect of the game type ($F(1,30)=2.830$, $p=.103$, $\eta_p^2=.086$, CI=.042 to .434), again with no significant impact of any of the CVs. Results were similar for the correct/incorrect scoring method. As a consequence we rejected Hypothesis 4a, which predicted an impact of the learning preconditions on the learning outcome.

Player perception To evaluate Hypothesis 3a and Hypothesis 3b, which predict an impact of the player experience on the learning outcome, Spearman's rho correlation was inspected for the first and the second DEG played by the participants (Table 6.20). For the learning outcome, the overall score (Q1.1-3.3) from the marking scoring method was used.

mean mark	Immersion	Flow	Competence	Tension	Challenge DEG	Positive	Negative
KnowQ 2 - 1	.128	.095	.093	.093	.122	.077	-.078
KnowQ 3 - 2	.087	.199*	-.141	.170	.255**	.042	-.121

Table 6.20: One-tailed Spearman's rho correlation between learning outcome (post - pre test mean mark) and player experience (GEQ factors); N=87, *=p<.05, **=p<.01.

Correlations were all very low, with the highest being found between ChallengeDEG and learning outcome for the second application. The direction of the correlation was as predicted negative for the Negative Affect factor of the GEQ and positive for the other GEQ factors. One exception was that Competence was negatively correlated for the second application. This may be explained by the ceiling effect of not being able to gather further knowledge in the second application while having increased Competence from the knowledge gained in the first application. Since the correlations were all very weak, we have to reject Hypothesis 3a and Hypothesis 3b.

Qualitative Feedback on DEGs

Similar to Study 1 we analysed participants' comments to get further insight into their play and learning experiences and any issues they may have had with the games. In addition we analysed participants' reasoning regarding their preferences for one of the DEGs, as due to the within-subject design in Study 2, they were able to compare both games.

Categorisation of positive/negative comments: Comments on what participants did or did not like about the DEGs were categorised as described in Section 6.2.2. For each category the number of statements are listed in Table 6.21 and a summary of the results is given below.

	Number of comments for each application							
	Positive				Negative			
	First DEG		Second DEG		First DEG		Second DEG	
Type of comment	ACT	ADV	ACT	ADV	ACT	ADV	ACT	ADV
Invalid or generic								
Unclear meaning	4	6	8	6	1	4	3	8
Not applicable	1	2	0	3	1	0	1	4
Nothing	1	0	0	1	4	6	5	4
Everything	2	0	0	0	0	0	0	0
Missing	0	0	1	2	4	2	3	1
Usability	1	3	0	1	5	2	4	3
Graphics	3	0	0	9	1	2	1	1
Experience	3	2	7	3	4	8	5	13
Application	7	2	5	3	2	7	3	4
Instructions	5	2	8	2	2	2	1	3
Visualisation	4	5	4	2	0	0	0	0
Content in steps	8	3	1	0	3	3	1	0
Content								
In general	3	9	3	4	0	0	1	0
About arrays	7	9	2	6	0	1	0	0
About coding	2	1	1	2	1	0	0	0
Game design decisions								
Time pressure	0	0	7	2	10	-	14	-
Story	-	1	-	7	-	2	-	7
Further issues and requests	-	-	-	-	3	1	0	1
Learning application	16	17	10	7	4	11	3	3

Table 6.21: Number of different types of positive and negative comments for the applications.

Comments about the games: One main part of comments were addressing positive or negative aspects of the games in terms of usability, graphics, experience, or other concerns.

Invalid or generic answers: The categories for the different types of invalid or generic comments were the same as in Study 1 and their description can be found in Section 6.2.3.

Usability: The ADVENTURE game was described as easy to use. One player found particularly useful that in the final task, the position in the code which was currently processed was pointed out.

Possible improvements mentioned for the ACTION 2 game were to indicate more clearly which variables belong to the row and which to the column in the array and to make the planes more easily distinguishable. The design of the planes was already adapted based on a similar request in a pre-test of the game (planes differ in colour and size, as shown in Section 5.4.1), but may still not be obvious enough. Furthermore the help button seemed to have been not obvious enough, as two comments asked to be able to revisit the help pages, which were also automatically shown at the beginning of a level.

For both games there were requests to add a keyboard input. According to the comments this was to some extent caused by the use of laptops. The game was not designed to be played with a touchpad, but with a mouse, which had been offered to the small group of participants using their own laptop.

Graphics: Graphics, animations and design were positively mentioned for both games. However, two comments regarding the ACTION 2 game asked for a more modern style and two regarding the ADVENTURE game criticised that there were too many animations.

Experience: Several positive experiences were reported for both DEGs, e.g. for the ACTION 2 game someone found it good to have had “a sense of achievement” but also to have been “annoyed if something was wrong”. The story aspect of the ADVENTURE was also source for positive experiences, as comments stated that it was funny and someone even liked the South Sea feeling, which shows that the choice of theme can have a noticeable impact.

Negative experiences for both games were that they were too long, or boring. One person stated for the ACTION 2 game that it was not possible to concentrate for so long, others found it exhausting or too stressful. Only one participant thought that the ADVENTURE was too challenging. In four cases however, it was described as childish.

Application: Positive game aspects mentioned by the participants were the theme with planes and islands for the ACTION 2 game, as well as the bonus level (for practising previously learned knowledge, as described in Section 5.4.2), and the score system, which someone suggested would even work better with a leaderboard. For the ADVENTURE the idea with the islands and the postman did also receive positive comments. In addition the different answer options in the dialogue, which is a typical feature of Adventure games, were mentioned positively.

The island theme was on the other hand also criticised, which shows that it is difficult to meet all tastes. Similarly for the ADVENTURE game someone disliked the dialogues, which had been positively mentioned by others. Of particular interest was that multiple aspect which are typical aspects for the Adventure genre were criticised, like not to have a score, no penalties, and not much increase in difficulty.

Comments about the learning: The other main part of comments were addressing the learning in both DEGs, which was supported through instructions, visualisation, but also game specific aspects depending on how the learning content was integrated in each game.

Instructions: Several students liked the clear and easily understandable instructions at the beginning of each level in the ACTION 2 game, as well as in the ADVENTURE game, where different players stated that it was well explained what they were expected to do.

Not many issues were reported with the instructions in the ACTION 2 game. Someone suggested to test players understanding with one example task prior to the actual level, but overall the understanding seemed to have been fine. For the ADVENTURE game, four students had issues understanding the tasks at the beginning, while one student stated that there were too many hints.

Visualisation: The visualisation of how an array is accessed and how the code is processed was positively mentioned several times for both games. No negative comments were made about this aspect.

Introducing new content in small steps: Comments on the stepwise introduction of new content were mainly positive. Players of the ACTION 2 game liked the revising bonus levels and the slow increase in difficulty. Less, but similar comments were found for the ADVENTURE game.

While the difficulty seemed to have suited most players, single comments stated that the bonus level was too difficult, or that there should have been more, or less tasks. For the ADVENTURE game the criticism was that there was too much at once, or that it was too often the same task.

Content: Comments about the learning aspect of the game were categorised depending on what participants stated to have learned, i.e. about arrays, or coding.

- *In general:* Several comments mentioned that it was easy to understand or well explained, but did not further specify, what it was exactly that they understood. This information is important, as a concern with DEGs is that players may learn how to play the game without learning the learning content included in the game. E.g. someone stated to have been “trained to detect the fields more quickly”, which is not necessarily helpful in writing programs which use arrays. However most comments seemed to have addressed the learning of arrays, including the only negative comment which stated that playing the ACTION 2 game was not helpful.
- *About arrays:* The way in which arrays were taught in both DEGs received multiple positive comments, like “now I think I really understand arrays” or “a good/short overview on the topic”. With only one negative comment, the students overall liked how the topic was presented.
- *About coding:* Positive comments referring to the coding were mainly about learning how to read a for-loop. The only negative comment was that the order in which row and column were accessed was fixed.

Game design decisions: The DEGs were designed in accordance with the characteristics of the selected genres MINI-GAME and ADVENTURE. Comments on their individual features are discussed below.

- *Time pressure in ACTION 2 game:* Positive comments on the *time pressure* were only made by students who had played the ACTION 2 game as second application, or who commented on the ADVENTURE based on prior experience with the ACTION 2 game. Multiple stated to have liked the *time pressure* in the ACTION 2 game, but on the contrary two comments for the ADVENTURE game stated that they liked that there was no *time pressure*. This controversy was also visible in several negative comments about the *time pressure* in the ACTION 2 game, who mainly criticised the stress that it caused, and not to be able to perform well under pressure. A few players only complained that it was too fast, but not that they disliked the *time pressure* in general.

- *Story in ADVENTURE game:* Several students stated that they liked the *story* and the dialogues in the ADVENTURE game, with one even saying that it was motivating. However, there were also several negative comments from players who did not like the *story*, since they thought that it was unnecessary or too long.

- *Further issues and requests:* Two participants requested an option to repeat a level in the ACTION 2 game, e.g. to try to correct a wrong answer. This functionality was implemented, but was removed for the experiment to meet the time restrictions and to ensure comparability between participants.

  As mentioned before there were unexpectedly long loading times for the games, which a few participants complained about in their negative comments.

Learning application: Several students stated that they liked the playful way of learning in both DEGs. More specifically some found the learning by doing and the immediate feedback on their answers positive. There even was a comment which stated that it would be good to have a DEG for every programming topic. Functionalities which were mentioned positively in supporting learning were the fact that wrong answers were not accepted in the ADVENTURE game, but had to be corrected, and the mail system metaphor which was chosen to teach arrays. For the ACTION 2 game it was the repetition which was seen as beneficial for learning, in combination with the variation of the tasks.

A suggestion for improvement of the ACTION 2 game was to show the instructions again, if multiple mistakes were made, although there was a help button with this functionality. Particularly important is one comment where a student described that after a while he/she did no longer look at the variables but only focused on the numbers, indicating that mechanisms were developed to more effectively solve the task in the game, but possibly learn less about arrays. For the ADVENTURE game some students criticised that the dialogues did not contribute to the learning. One main concern raised for both DEGs was to not be able to apply the knowledge gained in the game for solving the tasks in the worksheet later on.

Categorisation of comparison comments: Comments on student's preferences for one or possibly also both DEGs were also categorised (Table 6.22). Each category is further explained below, except for categories with a minimal number of comments.

	DEG preferences: number of comments		
Type of comment	ACT	ADV	Both
Invalid or generic			
Unclear meaning	9	1	0
Not applicable	0	0	1
Missing	3	1	3
Graphics	0	0	1
Experience	16	3	0
Application	21	7	4
Instructions	3	0	1
Visualisation	1	0	0
Content in steps	2	0	0
Content			
In general	10	0	0
About arrays	3	0	2
About coding	1	0	0
Game design decisions			
Time pressure	8	2	0
Story	11	0	0
Learning application	4	2	2

Table 6.22: Number of different types of comments which reason for the preference for one DEG over the other (ACT=ACTION 2, ADV=ADVENTURE), or for no preference (column: Both).

Comments about the games: Some comparison comments were concerned with the game aspects of the DEGs.

Invalid or generic answers: Descriptions for the invalid or generic comments categories are given in Section 6.2.3.

Experience: An often stated reasons for the preference of ACTION 2 over the ADVENTURE game were the player's experiences. Several students stated that they preferred ACTION 2 because it was exciting, fun, interesting, and could also cause anger. The ADVENTURE was favoured for being fun and less stressful.

Application: Difference in game features were the most mentioned argument for preferring one DEG over the other, or for liking both. Supporters of the ACTION 2 game liked above all that it was more demanding and also its point system. The ADVENTURE was preferred for being easier. Single comments for both DEGs argued that one was more a game than the other. Students who liked both games thought that the ADVENTURE was better to start with and the ACTION 2 to train the knowledge afterwards. This is in accordance with previous results which indicated that the games were perceived more positively in this order.

Comments about the learning: Several comments addressed differences in how the participants were learning when playing each DEG.

Instructions: A few students thought the instructions were clearer in the ACTION 2 game, however one comment mentioned that both DEGs were explained well.

Introducing new content in small steps: The ACTION 2 game was found to be structured more clearly.

Content: Several students thought that they had learned better with the ACTION 2 game, either without further explanation about what they felt they had learned, or particularly mentioning the topic array, or a better understanding of the code. There were only two comments which found that both games taught the topic well, but none who explicitly preferred the learning in the ADVENTURE game.

Game design decisions: The *time pressure* in the ACTION 2 DEG was for many players a reason to like the game, but few also disliked it for this reason. Several comments criticised the *story* in the ADVENTURE DEG, mainly because they thought it was time consuming and did not help them with the learning.

Learning application: Two students found that the increase in difficulty in the ACTION 2 game worked better for their learning process than getting too much help in the ADVENTURE game. An additional two comments preferred how the programming topic was linked with the ACTION 2 as opposed to the ADVENTURE game. Supporters of the ADVENTURE game liked the immediate correction of mistakes and thought that it was about learning, while the ACTION 2 was only about practising. Some did like both games due to the playful way of learning.

6.3.4 Discussion

The learning outcome for the ACTION 2 game was significantly better than for the ADVENTURE game, specifically for the type of knowledge questions which were trained in the ACTION 2 game. This result is in accordance with our assumption that the repetitive nature of the ACTION 2 game, following a behaviourist learning approach, is particularly suited for learning how to answer questions which are highly similar to the ones presented in the game. While we found our hypothesis to be confirmed (Table 6.23), we noticed that this learning success was not limited to highly similar tasks, but that players of the ACTION 2 game performed well throughout the whole knowledge questionnaire.

In terms of player experience, the ACTION 2 game exceeded the ADVENTURE game by having significantly higher ratings on three of the seven GEQ factors: Flow, Tension, and ChallengeDEG, and significantly lower ratings on the GEQ factor Negative Affect. While we did not predict the extend of the difference in player experience, it confirmed our hypothesis of a higher perception of challenge in the ACTION 2 game due to the *time pressure* (Table 6.23).

We also identified differences in player experience depending on the order in which games were played. When playing the ACTION 2 game first, the GEQ factors Immersion and Positive Affect were found to be significantly higher too, for the ACTION 2 game, as well as ratings on how much participants overall liked the game, would want to use/recommend it, and thought it was a learning tool. Our explanation for this order effect is that the ADVENTURE game is more thoroughly introducing the topic and thus better suited to be played first as opposed to the ACTION 2 game which is rather training the knowledge, once understood.

Although the order did play a role, the ACTION 2 game was clearly favoured by the participants when asked to compare both games, even for the group which started with the ADVENTURE and rated both DEGs more similarly.

Analysis of the impact of participants individual differences on the learning outcome and player experience was very restricted due to low response rate on the respective questionnaire. We were only able to evaluate the impact of learning preconditions on the learning outcome, which were not found to have an influence (Table 6.23). According to our results, the player experience did also show no impact on the learning outcome, which is why we rejected the corresponding hypotheses (Table 6.23).

nr	hypothesis	result
Learning outcome		
8a	The ACTION 2 game achieves a better learning outcome for knowledge questions which are very similar to the ones in this game than the ADVENTURE game.	confirmed
8b	The ADVENTURE game achieves a better learning outcome than the ACTION 2 game for knowledge questions which are beyond the scope of questions trained in the games.	rejected
Player experience		
9a	Playing the ACTION 2 game results in a higher experience of challenge than playing the ADVENTURE game.	confirmed
Relation player experience - learning outcome		
3a	The GEQ factors Immersion, Flow, Competence, Tension, Challenge, and Positive Affect are positively related to the learning outcome (Tension only up to a medium level).	rejected
3b	The GEQ factor Negative Affect is negatively related to the learning outcome.	rejected
Individual differences		
4a	The learning preconditions ease of learning programming, attitude to programming, math skills, and prior programming knowledge are positively correlated with the learning outcome, when learning a programming topic with a DEG.	rejected

Table 6.23: Results for hypotheses in Study 2.

Both DEGs received positive feedback, especially on the presentation of the learning content. Looking at the negative comments, no major issues were found which would indicate design flaws. The difference in player experience which we found in the quantitative analysis, was also found in the comments, where the ACTION 2 game was described as being more challenging. For the ADVENTURE game the story may be a weak point, as it leads through the game, but needs to catch the interest of the player. Somewhat similarly the time pressure in the ACTION 2 game, which overall led to a more intense player experience, was also criticised by some players for being too stressful.

6.4 Discussion of Both Studies

The learning outcome in both studies was better for the game from the ACTION genre (or more precisely MINI-GAME with tendency to the ACTION genre). These games did follow a behaviourist learning approach in training the player with simple, repetitive tasks. The remaining games in the comparison rather followed a constructivist approach. It is possible that the behaviourist approach was particularly successful for the narrow learning goal which we selected, of learning how to access cells in a 2D array.

The comparison of DEGs which are of the same game genre and differ only by a single game attributes showed an impact on the learning outcome, but only when modifying the Challenge attribute, not the graphic attribute which further specifies the Game World/Setting. We assume that since Challenge is a central element of the game play, it has a higher impact than attributes of other game elements.

When comparing DEGs from different genres, which differ in multiple attributes, we found again an impact on the learning outcome, but also on multiple dimensions of the player experience. The higher difference between the games seems to result in more potential for a difference in player experience.

6.4.1 Limitations

To measure the impact of game type, target group and topic had to be fixed in our studies. The results are therefore not generalizable for other target groups and topics. As mentioned, it is a starting point to identify how large the impact of game type can be. The relation between topic, target group, and game type needs then further investigation (Section 7.3).

As our results are limited to the selected learning topic and goal, it is possible that the behaviourist learning approach is only suited better for this specific learning topic. When researching the impact of game type for other learning topics, it will be of particular interest to validate the success of the behaviourist as compared to the constructivist learning approach.

A drawback of Study 2 are confounding interactive effects. Due to the multiple variations in features between the two games, it is difficult or impossible to determine the exact cause of observed differences.

Both games may be further modified and possibly improved, but there were no major complaints, i.e. multiple negative comments on the same aspect of the game. It would still be interesting to see, how much impact further changes of the games might have on the results.

Chapter 7

Conclusion

Summarizing the results from this thesis, we discuss the findings for our main research question along with the more detailed sub-questions. Finally we point out limitations and directions for future work.

7.1 Discussion

The main research question underlying this thesis is the impact of game type on the success of a DEG in terms of learning outcome and player experience. To approach this question we have answered several sub-questions, which were listed in Introduction (Section 1.2). The results for each sub-question are presented below.

7.1.1 Research sub-questions

Each of the chapters answered a set of research questions, which were identified as essential sub-steps to approach the main question.

1. Game features - Chapter 3
    - *RQ 1a: What are fundamental game features, representing similarities and differences between games?*
    We have identified two types of game features: game elements (i.e. the main building blocks of every game) and game attributes (i.e. the various options on how each game element can be implemented in different games). The game elements derived from the definition of games are:

Player, Input/Output, Actions, Challenges, Goals, Rewards/Penalties, Setting/World, Perspective, and Structure. When analysing how each of these elements is realised in different games, similarities and differences between games can be identified. We compiled a list of attributes for each element, aiming to cover the most important ones without adding too much detail, which would result in an impractical collection of possibly overlapping features. The relation between game elements, and their associated game attributes has been summarized and visualised in the Game Elements-Attributes Model (GEAM: Figure 3.1).

- *RQ 1b: How can we ensure to have found a comprehensive set of features?*
  There are several reasons why we can argue that the game elements and game attributes included in the GEAM represent a comprehensive model for games. (1) The game elements have been extracted from the definitions of games by different authors (Juul, 2011; Salen and Zimmerman, 2004) who have well justified their definitions based on their own research as well as existing literature. Therefore we can assume that the game elements are indeed the main components of every game. (2) The game attributes have been selected from the results of an extensive literature search. Reasons have been given for the features that were not included in our model. Additionally we conducted interviews with gamers, using the repertory grid technique, to further consolidate our findings. (3) Structuring the game attributes by the elements provides a much clearer picture of their purpose in the game (e.g. being a Challenge, or a Reward). Since every element was assigned several attributes it also indicates that the GEAM does indeed provide a comprehensive selection of game features.
- *RQ 1c: How can we structure the identified features to facilitate the comparison of games?*
  As mentioned above, game attributes were assigned to the game elements. Since game elements are what defines games, they are included in every game. This allows us to compare games regarding all their basic components. For each element it can be analysed which attributes are implemented in the game and if they are the same or different. Our GEAM therefore allows a structured comparison based on a selected set of attributes.

2. Game types - Chapter 4

- *RQ 2a: What relations exist between the game features?*
  We proposed the assumption that there are common sets of attributes

which appear in different types of games. An indication for this assumption is the existing classification systems which group similar games. Since these classification systems were found to have several flaws, we conducted a survey to study the relation between attributes. The relation between game elements had already been extracted from the game definitions and visualised in our game model (Figure 3.1). Participants analysed a variety of games for the game attributes implemented in the games. Using a cluster analysis we were then able to search for similarities and differences between games, and identify the sets of related attributes for each cluster.

- *RQ 2b: How can we make use of these relations to identify significantly different types of games?*
  To analyse how much the identified clusters differ, we used multidimensional scaling to visualise their difference in distances on a 2D plane. The four corners of the plane and the central area are furthest apart from each other. They were therefore used as a foundation for a classification with five game types, which differ most significantly. While the areas are still connected, the overlap between the types is reduced to a minimum. Each area was then analysed to find attributes, which characterise the games located in the clusters of this area. As a result we determined the defining set of attributes for the five identified game types.
- *RQ 2c: Can we rely on game genre as a commonly used classification system?*
  Game genres have been identified to have several flaws, like inconsistent definitions, unclear relations between genres and different sets of genres. However, existing game genres are commonly used. Hence, it is desirable to substantiate them with more precise definitions instead of discarding them. To be able to tell if and which genre names fit the five identified game types, we asked participants of the survey to assign the analysed games to one or multiple commonly known genres which we had extracted from our comparison between several different descriptions of game genres. The genre names turned out to be assigned to neighbouring clusters, proving that their usage was still relevant to our results. Consequently, we decided to reuse the names ACTION, ADVENTURE, and ROLE-PLAY genre. The Puzzle and Strategy genre names were found to be misleading, because the term *puzzle* was already used for a GEAM attribute, which is not even most characteristic for the genre and strategy is a concept relevant to many different games. They were replaced by the terms MINI-GAME and RESOURCE genre, respectively.

3. Educational game design - Chapter 5

   - *RQ 3a: How can we approach the design of digital educational games?*
     We have proposed the Tripartite Educational Game Model (Figure 5.2), according to which the design of a DEG needs the identification of a target group, a learning topic, a suitable game to teach the learning topic, and a way to embed the learning content into the game. Following these basic steps for the design of DEGs for our study, we first decided on the target group (computer science students). We then surveyed the target group to find a learning topic for which a DEG as learning tool would be desirable (access of 2D arrays in Java). Finally we needed to decide on a game genre, along with a selection of game attributes, and a concept on how to integrate the content. This step is further described in the subsequent answers.

   - *RQ 3b: How can learning content be incorporated into a game?*
     Based on our GEAM we conducted a thorough research in analysing each attribute individually on how it can contribute to the learning process, either by supporting the approach of one of the learning theories, or by enhancing the player's motivation and engagement. To further understand how the selected topic can be included in a game, we studied several DEGs for learning programming and found four main concepts: the content was either (1) mediated through a quiz, or (2) through visualisation of code in combination with puzzles, but could also be (3) multiple concept 1+2 games, linked by a story, or (4) a program coded by the player, which is executed to solve a task. We applied the findings by giving design guidelines for our five game genres, according to their defining attributes. MINI-GAMES and ACTION games can follow the behaviourist learning approach, by using *time pressure* as main Challenge. ADVENTURE games are suitable for the constructivist approach with the Challenge *puzzle*, and by *exploring* the game world, guided by a *story*. RESOURCE games work well with constructivism/constructionism, by using *limited resources* to *produce* something. ROLE-PLAY games combine many of the attributes from the other genres and provide a particularly strong relation between the player and the *character* in the game, which supports social learning.

   - *RQ 3c: What are possible design solutions for DEGs teaching the access of fields, rows, columns, and areas in 2D arrays with Java?*
     As the topic to be taught is rather small, MINI-GAMES or the genres which are still closely related to MINI-GAMES are suitable: ADVENTURE,

Action and Resource. Since Mini-games have a small set of defining attributes and are thus limited in possible variations, we reckoned that the Challenge can be varied to be either *time pressure* or *puzzle* and that the type of graphics, which is usually *abstract*, can also be *realistic* or *fantasy*. Accordingly we designed two pairs of games, with one DEG being part of both pairs. The Adventure genre offers an alternative solution, by also using *puzzles* as a Challenge, but additionally explaining the topic via a *story*. Games of the Resource genre can teach the topic by making the different possible values of the variables a *limited resource*.

4. Comparison of DEGs with different type - Chapter 6

- *RQ 4a: How can the success of a DEG in terms of learning outcome and player experience be measured?*
To determine the learning outcome, clear learning goals need to be set. We used a pre-post test design with questions that evaluated the knowledge taught by the game, as well as some more advanced questions to examine a possible transfer of knowledge. The learning experience was measured with the standardized game experience questionnaire (GEQ), developed by IJsselsteijn et al. (2008). In addition, qualitative data was gathered via free text comment fields and analysed by categorising the individual statements. The most relevant categories were concerned with the application's usability, graphics, and experiences, and its aim to teach, using instructions, visualisations, and the attributes of each genre (e.g. *time pressure*, *story*, and progression throughout multiple *levels*).

- *RQ 4b: What additional characteristics of the learner may have an impact on the success of a DEG?*
Besides the type of game, we considered other factors to have an impact on the success of a DEG. Based on a literature search we developed a conceptual model (Figure 6.1), which suggests that the learning outcome of a DEG can be influenced by learning preconditions (e.g. the attitude towards the topic and prior knowledge) and learning style, and the player experience can be influenced by play preconditions (e.g. game preferences) and personality traits. To validate the model we measured the factors in the context of our studies, using Felder's scale (Felder and Spurlin, 2005) to determine the learning style and the BIG 5 measures by John and Srivastava (1999) to determine the personality traits. Based on the conceptual model we proposed several hypotheses, most of which were rejected. Only for Study 2 the Hypothesis 8a and Hypothesis 9a were confirmed, which state that the ACTION 2 game, which uses *time pres-*

sure and the behaviourist learning approach, achieves a better learning outcome for the set of knowledge questions particularly trained by this game and is experienced as more challenging than the ADVENTURE game.

- *RQ 4c: What is the impact of game type for DEGs which only differ by one game feature (more precisely GEAM attribute)?*
  To approach this question we conducted a study comparing three DEGs, which in pairs of two differed by single attributes. Significant differences in learning outcome were found, showing that even a single game attribute can have an impact. The ACTION game, which uses *time pressure* as Challenge was more effective for learning to answer some of the knowledge questions than the PUZZLE game, which used *puzzle* as Challenge. A difference in graphics however (*abstract* versus *realistic*) had no noticeable impact on the learning outcome. Interestingly no significant differences in learning outcome were found between the e-learning TOOL and two of the DEGs, showing that DEGs are not necessarily more effective than non-playful learning applications, however there also seems to be potential that they are, as the ACTION game was more successful. For the player experience, no significant differences were found between the three DEGs, indicating that a DEG can be more successful in teaching a topic, without differing in the player's experience from less successful games.
- *RQ 4d: What is the impact of game type for DEGs which differ by more than one game feature?*
  To answer this question we compared DEGs of different game genres. Again playing the ACTION 2 game, which used *time pressure* and the behaviourist learning approach, resulted in a better learning outcome, compared to the ADVENTURE game, which used the constructivist learning approach with *puzzles* and a *story*. In addition the two DEGs differed in multiple dimensions of player experience, namely Flow, Tension, Challenge, and Negative Affect. The ACTION 2 game had higher scores for the first three dimensions, and lower scores on the last, which is consistent with the observation, that the majority of players preferred the ACTION 2 game to the ADVENTURE game. However, the positive comments for the ADVENTURE game and the small group of participants who prefer this game, indicate that the results are not simply caused by bad game design. This leads us to the assumption that game genres do indeed have an impact on the learning outcome and player experience, and that some genres may be preferable to others for teaching a certain

topic. The results also show that individual differences are an important factor which means that not all learners may be satisfied by one DEG.

7.1.2 Research main question

Overall our aim is to gain an understanding of the impact of game type on learning outcome and player experience. The presented research contributes to the attainment of this aim by laying some groundwork to address the issue in a structured way and by describing the methodological framework developed and empirical results obtained. The GEAM and the Game Genre Map are the viable tools developed for identifying which game attributes may be of particular interest for comparison. The redefined game genres narrow down the number of different games to be researched, by highlighting characteristic attributes shared by multiple games, allowing us to design representative games for each genre and to identify possible variation within a genre.

Our results showed more significant differences in learning outcome and player experience for games from different genres, which differed in multiple attributes, but even single attributes showed an impact on the learning outcome. This indicates the importance of further researching the impact of game type, to support DEG designers in building games which are effective and enjoyable tools for learning, while also being cost-effective.

To further understand the factors which may influence the impact of the game type, we proposed a conceptual model which considered learning and play preconditions, as well as the player's learning style and personality traits. For Study 1, only the covariate 'attitude to programming' showed a significant impact on the learning outcome, but with a low effect size. This impact was not found for Study 2. If learning precondition had none, or only a weak impact on the learning outcome for DEGs as opposed to more traditional teaching methods (for which several studies found such an impact) it would lend support to the use of DEGs for students with poorer preconditions. For the remaining covariates we could not confirm any influences as suggested by related research and it needs further investigation to find out how and when they impact learning outcome and player experience and if there are other important factors to be considered, which have not been identified yet.

In conclusion, while we have only just started to understand how the game type impacts learning outcome and player experience for a DEG, we have shown that it is a relevant question to ask, revealing a whole field of related research questions, and how it can be approached systematically.

7.2 Limitations

While we aimed to consolidate and justify our approach to research the impact of game types by considering the related literature and thoroughly designing our studies, each step in developing our methodological framework as well as the empirical studies has been restricted by certain limitations.

- *The GEAM is bound to limitations in its selection of game attributes.*
  The GEAM attributes capture the different game features only up to a certain degree of detailedness. As the GEAM is not exhaustive and cannot be so, it basically cannot describe individual games with full details. Attributes which are too specific or too unusual to be included in GEAM may still be of interest for educational purposes (e.g. GEAM includes the generic attribute "special controller", but not "electric guitar", which is for example used in the DEG Rocksmith[1], where the player learns how to play songs). Our comparison studies were limited to the attributes considered in GEAM and are bound to the identification of possible implementations. Nonetheless, the GEAM was designed to provide a basic structure for the comparison of game types and is not supposed to be prescriptive, leaving room for game designers to exercise their creativity to implement individual game features.

- *Due to the methodology used to develop the Game Genre Map it can only be an approximation.*
  One of the main limitations of the Game Genre Map is its two-dimensionality, which compromises the spatial arrangements of the identified clusters. Furthermore, due to the diversity of games and the creativity involved in the design of a game, not every game may be classified distinctively and therefore the positioning of individual games on the map cannot be exact.

- *Only a particular target group and learning topic were considered in our comparison studies.*
  For a feasible comparison of DEGs of different game types, we had to optimise the number of experimental variables and their respective levels, by deciding on a target group and a learning topic. The generalisability of our findings is

[1] http://rocksmith.ubi.com/rocksmith/en-GB/home/index.aspx

therefore limited to computer science students learning programming and in particular the topic 2D arrays in Java.

- *The interaction effect of attributes for game genres comparison in Study 2 is not known.*
  As reported above, more differences between DEGs were found in Study 2, which compared games of different genres, differing by multiple GEAM attributes. Yet, the individual effect of each attribute on the DVs cannot be isolated.

- *The measurement instruments for gathering data may have limitations.*
  For our empirical evaluation we decided to use certain instruments for measuring the variables of interest: the GEQ for player experience, the AttrakDiff2 for perceived quality of the applications, Felder and Soloman's (2015) LSI for learning style, and John et al.'s (2008) BFI for personality traits. These instruments could all have impact on the results and need to be questioned critically. Several of the GEQ factors were found to have a significant medium correlation (Section 6.2.3, Section 6.3.3), which raises the question, if more independent dimensions need to be found. In addition, Cox et al. (2012) argued that Flow as opposed to Immersion is not a graded experience, as someone either is or is not in flow. However GEQ measures both with a five-point scale. Criticism on learning styles has already been presented in Section 6.1.2 and personality traits were chosen due to a lack of a robust player model as indicated by Bateman et al. (2011).

- *Results for the conceptual model to investigate causes of game type differences were mainly not significant.*
  We need to explain the detected differences between the DEGs compared in Study 1+2. The conceptual model, which we proposed in Section 6.1.3, is not yet powerful enough, as our results were mainly not significant. From the qualitative analysis of the participants' comments on the different games, we derived that individual preferences play an important role in how players perceive and perform in a DEG. It will be challenging to identify variables which capture a confined set of player characteristics that allow us to explain the detected differences. On the other hand, it is not practical to include too many potential variables in one model.

7.3 Future Work

Based on the issues discussed in the previous section we propose possible solutions as future work.

- *The GEAM is bound to limitations in its selection of game attributes.*
  The GEAM was developed to identify key game features which are considered to capture the main characteristics of and differences between game genres. The selection of GEAM attributes is prone to debate, as some researchers may argue that certain attributes are missing, or should be renamed. We invite researchers to discuss possible modifications to further improve the model's applicability.

  In general there are two approaches to further investigate the impact of game type, as shown in this thesis: By studying single attributes or the more strongly different game genres.

- *Due to the methodology used to develop the Game Genre Map it can only be an approximation.*
  The Game Genre Map is a fundamental tool for identifying promising game genres, or mixtures of genres, when designing a game, or in particular a DEG. We have already pointed out how a DEG of the Resource genre could be designed for the 2D array topic (Section 5.4.3) and compared against the existing games. Following our approach we propose that further research investigates the impact of genres differences for DEGs. In particular, we have not considered the Role-play genre, as it requires more effort to build such a rather complex game and as it may be better suited for a more sophisticated topic (e.g. for programming possibly "object orientation" or a combination of introductory concepts), to make use of the broad range of GEAM attributes it facilitates.

- *Only a particular target group and learning topic were considered in our comparison studies.*
  While our results cannot be generalised to other topics and target groups, they do show an impact of game type on learning outcome and player experience. Thus we would expect to find differences when repeating the study with games on a different topic and for a different group of learners. However, we also expect that both, topic and learners, have an impact on how effective a certain type of game is, as different learning approaches may be more or less applicable and player preferences may vary. One of the main steps of future work

will be to enhance the methodological framework to also consider the different values of the two variables - learner and content - besides game type. It will be important to find out, which type of game suits which type of content and target audience.

- *The interaction effect of attributes for game genres comparison in Study 2 is not known.*
  While the comparison between genres appears to be more promising for DEGs in terms of impact on learning outcome and player experience, further research on the modification of single attributes is also needed, if we want to gain further understanding on their individual impact. Depending on the genre it may not always be sensible to modify each attribute, e.g. an ADVENTURE game is typically a single-player and not a multi-player game. Game genres are therefore to be carefully selected when studying particular attributes.

- *The applied instruments for gathering data may have limitations.*
  For player experience, there are several alternative questionnaires (Nordin et al., 2014), which should be examined and compared to the GEQ (Section 6.2.2). There are also approaches to use eye-trackers or other objective measures for player experience, as opposed to self-reported questionnaires, which can also be a valuable attempt for future research. The comparison between e-learning tool and DEGs with the AttrakDiff2 was much less specific than the GEQ. Since we found that the tool was possibly gamified by the participants in setting self-defined goals, we need a measure which provides more in depth understanding on differences in experience when using a tool as compared to a game. Valuable input may be found in Frazer's (2010) work, who understands DEGs as another form of e-learning tools. As pointed out, learning styles are somewhat controversial. A qualitative approach may be more promising in observing how students learn during play. Instead of personality traits the research on player types should be considered, even if it still has limitations.

- *Results for the conceptual model to investigate causes of game type differences were mainly not significant.*
  Besides revising the instruments, we also need to consider further variables which may explain individual preferences for DEGs and find instruments to measure them. We have recorded the players' moves in the games, but did not yet manage to replay and screen all of them for an in depth analysis of players' behaviour. In combination with interviews, this could be a way to identify why individuals prefer one game over another and why a certain DEG may work particularly well for learning a topic.

Due to the increasing interest in DEGs we think that our research question is highly relevant and have laid some groundwork for future investigations. With the large diversity of games, more collaborative effort of the research community is needed to find more conclusive answers and deeper insight into the issue of how game type is related to the success of a DEG.

Appendix A

A.1 Results from Game Model and Attributes Search

Database	Total	3D	GT	CG	S	OM	GM	UM	NF	NA
ACM DL	73	5	9	13	6	14	12	5	9	0
Sage	22	0	0	0	3	2	15	0	2	0
ScienceDirect	116	12	19	6	5	6	14	4	41	9

Table A.1: Categorized search results for "game model"; 3D = 3D model, GT = game theory, CG = concept single game (type), S = simulation/simulated game, OM = other kind of model, GM = game model, UM = unclear meaning, NF = term not found, NA = no access

game element	representative term	terms found during search
players	competition cooperation communication sociability	competitive aspect, one player versus another diplomacy, collaboration interaction between players (speech, gesture, mimic), communication with other players number of players, multiplayer gaming
input/output	input devices	controls, tangible interfaces
actions (interactions)	communicating character	communicating with non-player characters, language: verbal and text improving avatar, role playing piloting, choosing objects
challenge (conflict, problem)	*time pressure* *opponent*	racing against the clock, allotted time playing against the computer, moving/aggressive targets
goals (objectives)		quests, specific goals quantifiable outcome, win-lose victory conditions
rewards (achievements)	feedback - *score* - *praise* *impact on gameplay* - new level - *resources* - power-up	 points, point system, achievement points, assessment leaderboards, medal system, badges, change of the players virtual status, title, reputation gain levels, leveling up, access to new game space/ people/levels, promotions, greater responsibility experience points, virtual currency, gaining new tools/clothes/currency, money, various items power-ups, improving avatar/city/civilization /business to highest level or their own goal
game world (environment, space)	 explore sound	visual style, graphics, fantasy elements, fantasy, realism, mystery, lightning effect, interface design, game fiction space that the player can perform in and explore sound effects, music, sounds
structure	*continuity* story time	levels story, storyline, narrative, fantasy, game fiction, dramatic arc game time

Table A.2: Game attributes identified in the broad literature search, sorted by element and grouped by similarity with choice of representative term for each group. First column: elements with synonyms found during search in brackets; second column: representative terms, italic if not found during the search, but chosen apposite to the remaining results; third column: search results, excluding terms which do not describe games as systems, but from another perspective.

game element	representative term	terms found during interview
players	competition cooperation communication sociability	competitive teams, cooperative party game single, play alone, multi, play in company
input/output	input devices	keyboard, mouse, gamepad, special input device
actions	move fight/destroy collect produce/build character	no/a lot of walking, free movement, restricted/ predefined path destroy (everything), gun, shoot collect (resources) build something, build up avatar, main/virtual/no character, no/multiple heroes, one character, masses, army, many characters, sophisticated/personalised character, with/without personality
challenge	time pressure limited resources opponents puzzle	slow/fast pace, time limit inventory, resource management combat, eliminate enemies thinking, logic
goals	static variant	always same target, determined mission, free choice of tasks / strict goals
rewards	feedback - score *impact on gameplay* - new level - resources - power-up	no improvements collect points improvements, mistakes minor or big impact level up collect money character improves, gain features
game world	 graphical detail dimension music	realistic, fantasy, comic, science fiction simple/elaborate graphics, functional, graphically nifty 2D, 3D music important/unimportant
perspective	view flexibility	direct/indirect navigation (not through/through a character), camera fixed/vagrant, camera bound/ free, ego/god view, view whole game, view limited
structure	continuity end of game duration story	level, no clear level, different level, continuous, open world, level areas (fixed levels) game has/has no end long/short game, a lot/little game time to finish game story important/unimportant for course of game, no story

Table A.3: Game attributes identified in the repertory grid interview, arranged as in Table A.2.

A.2 Sets of Game Genre

genre	Rollings and Adams (2003)	Fullerton (2008)	Wolf (2001)	McCann (2009)	en.wikipedia.org	amazon .co.uk	amazon .com
action	x	x	-	x	x	& shooter	x
strategy	x	x	x	x	x	x	x
role-playing	x	x	x	x	x	x	x
adventure	x	x	x	x	x	x	x
puzzle	x	-	x	-	x	x	x
sports	x	x	x	x	x	x	x
simulation	vehicle simulations	flight & other simulations	x	x	x	x	x
racing/driving	-	x	(separate)	-	(part of vehicle sim.)	racing	racing
fighting	-	-	x	-	x	x	-
flying	-	-	x	-	-	-	x
fitness	-	-	-	-	-	x	-
edutainment	-	x	-	-	-	-	-
educational	-	-	x	-	x	& reference	-
children's	-	x	-	-	-	x	-
adult video	-	-	-	-	x	-	-
games for girls	x	& women	-	-	-	-	-
casual	-	x	-	-	x	-	-
music	-	-	rhythm & dance	x	x	& dancing	rhythm
artificial life	x	-	x	life simulation	-	-	-
online	x	-	-	mmorpg	mmorpg, strategy	-	-
board, card & casino	-	-	(separate) 3.gambling	-	board/card	x	(separate)
arcade & platform	-	-	platform	-	-	x	arcade
quiz & trivia	-	-	quiz	-	trivia	x	trivia
action-adventure	-	-	-	x	x	-	-
construction & management simulations	x	simulation/ building	management simulation	-	x	-	-

Table A.4: Sets of game genres from different authors/sources

A.3 Game Classification Survey

Player	
1	When you are playing the game, are there other players in the game with you? (players = humans or computer players replacing humans)
a	There are no players other than me.
b	There can be other players.
c	There are always other players.
1a -> 2	While other players play the game on their own as well and they are not present simultaneously with you in the game, are their game performance/behaviour relevant to you?
a	Yes, I compete against others, e.g. through a highscore.
b	Yes, I cooperate with others, e.g. by playing the game in turns, exchange items.
c	Yes, I communicate with others in the scope of the game (e.g. through a forum or when you watch each other play).
d	No, other players are not relevant to me.
1b/c -> 3	In what way(s) do you interact with the other players?
a	I compete against them (who is faster, better, stronger).
b	I cooperate with them (e.g. help them, play together or exchange items).
c	I communicate with them.
d	I play on my own, even if there are other players around.
4	How would you rate the importance of the following aspects? (individually out of 100)
2/3a -> a	Competition [0:not important at all -> 100:very important]
2/3b -> b	Cooperation [0:not important at all -> 100:very important]
2/3c -> c	Communication with other players [0:not important at all -> 100:very important]
3d -> d	Be able to play alone [0:not important at all -> 100:very important]
Interaction	
1	On which device do you most often or prefer to play the game?
a	Computer / Laptop
b	Console (e.g. Playstation, Xbox, Wii)
c	Handheld Console (e.g. Gameboy, portable Playstation)
d	Mobile Phone, Smartphone, Tablet, MP3 Player
2	What input device(s) do you usually use to play the game on the [chosen device]?
1a -> a	Mouse

1a -> b	Keyboard
1a -> c	Touchpad
1a/b -> d	Gamepad
1a/b -> e	Joystick
1a/c/d -> f	Touchscreen
1a/b -> g	Special Controller (wheel, guitar, etc.)
1a/b -> h	Tracking / no Controller (e.g. Kinect)
1b/c/d -> i	Motion Sensor (e.g. Wii-Controller, PS3-Move / Gameboy with a game featuring the motion sensor)
1c/d -> j	Keys / Buttons (not on Touchscreen)
Actions	
1	Which of the following actions are available in the game?
a	Move: you can move objects or characters in the game
b	Place or Position: you can place objects or characters in the game at a position
c	Fight or Destroy: you can shoot or attack/destroy objects or opponents
d	Collect: you can collect items or resources
e	Produce or Build: you can combine or use resources or items to build new ones
f	Communicate: there is a textual or verbal communication
2	How would you rate the importance of the following action(s) in the game? (individually out of 100)
1a -> a	Move [0:not important at all -> 100:very important]
1b -> b	Place or Position [0:not important at all -> 100:very important]
1c -> c	Fight or Destroy [0:not important at all -> 100:very important]
1d -> d	Collect [0:not important at all -> 100:very important]
1e -> e	Produce or Build [0:not important at all -> 100:very important]
1f -> f	Communicate [0:not important at all -> 100:very important]
1a -> 3	How do you move objects or characters through the game?
a	I can only move them to fields / restricted positions.
b	I move them mainly in one direction / one way (along a path or route).
c	I can move them rather freely in the game environment.
1b -> 4	How do you place or position objects or characters in the game?
a	I can place or position them only on fields / restricted positions.
b	I can place or position them rather freely in the game environment.
5	Do you mainly control one character in the game, which performs the actions you chose above? This can be different characters over time. (character = person or car, or a similar token assigned to you)
a	Yes, I control mainly one character and perform the action with it.

b	No, but there is (at least) one character, that I have a sepcial connection or emotional attachment to.
c	No, I don't control such a character.
5a/b -> 6	Please rate the following aspects of this character. (if multiple characters please choose the one you are most attached to)
a	Emotional connection to the character
b	How much can you personalise the character (appearance & abilities)
Challenges	
1	Which of the following challenges do you face in the game?
a	Time Pressure or Reaction: fast reaction is needed or something has to be achieved in limited time
b	Limited Resources: limited life/lives, energy or resources that are needed to perform actions or produce something
c	Opponents: objects or characters with (limited) intelligence that attack the player or play against him/her
d	Obstacles or Forces: something that makes it harder to achieve the goal or to reach it in a straightforward way (e.g. walls or gravity)
e	Puzzle: you need to use skills (e.g. logical thinking, pattern recognition) to be successful or move on in the game
f	Quiz: you are required to answer questions, which are usually short
g	Search and Find: you need to search for something
h	Limited or No Savability: you can't save the game status at any time or at all
2	How would you rate the importance of the following challenge(s) in the game? (individually out of 100)
1a -> a	Time Pressure or Reaction [0:not important at all -> 100:very important]
1b -> b	Limited Resources [0:not important at all -> 100:very important]
1c -> c	Opponents [0:not important at all -> 100:very importan])
1d -> d	Obstacles or Forces [0:not important at all -> 100:very important]
1e -> e	Puzzle [0:not important at all -> 100:very important]
1f -> f	Quiz [0:not important at all -> 100:very important]
1g -> g	Search and Find [0:not important at all -> 100:very important]
1h -> h	Savability [0:not important at all -> 100:very important]
Goals	
1	How would you describe the goal(s) that you try to reach in the game?
a	Static (always the same goal, e.g. reach the end of the level)
b	Various (different goals, like missions, tasks or quests)
c	Self-defined (you set your own goals)

1b -> 2	How free are you to choose which goals (missions, tasks or quests) you want to complete?
a	(0:not at all (I have to achieve all goals) -> 100: very much (I can choose freely which goals I want to aim for))
Rewards / Penalties	
1	What rewards do you get for reaching a goal in the game?
a	Access to new levels / regions of the map or game world
b	Power-ups (e.g. buffs), improvements, access to new items
c	Score / highscore, points
d	Resources, money, lives (something you need/use in the game)
e	Praise (through sound or text)
2	How are mistakes made during the gameplay penalised?
a	Losing time (go back to last saving point, game over - play again)
b	Losing power-ups (e.g. buffs), improvements, losing items
c	Lower score / less points
d	Losing resources, money, lives (something you need/use in the game)
e	No penalties
World / Setting	
1	How would you describe the visual appearance of the objects and the environment in the game? (along the three dimensions given below)
a	Abstract (not a world, but a grid or board) [0: not at all -> 100: very much]
b	Realistic (similar to real world) [0: not at all -> 100: very much]
c	Fantasy/Fictional (not real) [0: not at all -> 100: very much]
2	How would you rate the level of detail of the graphics and animations of the game?
a	Level of graphical detail [0: very low -> 100: very high]
3	What is the dimension of the game environment.
a	2D (you can only interact inside the screen plane)
b	3D (interaction towards or away from you are also possible)
4	How much can you explore in the game (how rich and large is the game environment)?
a	How much to explore [0: nothing -> 100: very much]
5	Are you free to choose the environment you want to play in (independent from your progress in the game)?
a	Yes, out of a selection of maps / levels / environments.
b	No, it is predefined.
6	What view do you have in the game?

a	My view in the game is bound to a character (from the character's perspective or following it).
b	My view in the game is freely movable through the whole game world/environment (which is accessible at the time).
c	My view in the game is fixed (does not move, but might change between different positions).
Structure	
1	Does the game have a story?
a	Yes
b	No
c	Don't know
1a -> 2	How important is the story for the game?
a	The story is [0: not important at all -> 100: very important]
3	How is the game structured?
a	The game is broken down into separate parts - you finish/get through a part (level or map) and move on to the next one.
b	The game does not have separate parts, but is continuously progressing.
4	Did you ever finish the game?
a	Yes, I finished it at least once.
b	No, because I am still playing it.
c	No, because I didn't have the time to finish it.
d	No, because it is too long.
e	No, because the game doesn't really have an end (you can play on forever).
f	No, because it became too difficult.
g	No, because I lost the interest.
h	No, because I never played the game, I only know about it from other sources (e.g. heard about it, watched someone playing it),
5	Did you play the game multiple times (starting from the beginning)?
a	Yes, I played the game multiple times.
b	No, I played the game only once.
5a -> 6	Why did you play the game again?
a	I played it again to improve my skills.
b	I played it again to make new experiences (i.e. explore different options in the game).
c	I played it again on a higher difficulty level.
d	I played it again to reach better results.
e	I played it again because I wanted to try again (after a game over).
f	I played it again after a long time to re-experience it.

g	I played it again because I like the game so much.
5a -> 7	How many times did you replay the game / restart it (approximately)?
a	[Numerical input]
5b+4a -> 8	Why did you play the game only once?
a	I played it only once because it is too long.
b	I played it only once because I don't expect (many) new experiences, if I would play it again.
c	I played it only once because I don't like the game so much.
5a+4a -> 9	How long did it take you to play the game once (pure playtime)? Please state the duration as a range between an approximate minimum and maximum in hours. (minutes can be expressed through a decimal, e.g. 30 min = 0.5 h)
a	[Numerical input: range]
5a-4a -> 10	How long did you play the game before you started again from the beginning (pure playtime)? Please state the duration as a range between an approximate minimum and maximum in hours. (minutes can be expressed through a decimal, e.g. 30 min = 0.5 h)
a	[Numerical input: range]
5b -> 11	How many hours (approximately) did you spend playing the game in total? (minutes can be expressed through a decimal, e.g. 0.5 h = 30 min)
a	[Numerical input]
Classification	
1	What category/categories would you assign the game to? Please make use of the comment fields to state why you chose a category or if you have any comments on your choice.
a	Action
b	Adventure
c	Puzzle
d	Role-playing
e	Simulation
f	Sports
g	Strategy
2	Would you use a category name that is different from the ones provided above?
a	Yes, I would use the following category name(s): (please state them in the comment field and separate them through a semicolon [;]) [text field]
b	No, I think the category name(s) chosen above is/are appropriate.

3	Do you have any final comments about the questionnaire or the game? Please state them here.
a	[text field]

Table A.5: Questions from the Game Classification Survey. Arrows (->) indicate dependencies between questions.

A.3.1 List of non-predefined games selected in the Game Classification Survey

4 story
Age of Empires
Battlefield, Battlefield 3 (2x)
Borderlands
Company of Heroes
Counter Strike (3x)
Cube 2
DDO
Defense Grid Awakening
Devil May Cry
Doom
Dragon Age
Eufloria
EVE Online
Fable
Faster than light
Geneforge
Guild Wars 2
Heavy rain
League of Legends (4x)
Left 4 Dead 2
Lord of the Rings Online
Mass Effect (3x), Mass Effect 3
Metal Gear Solid
Rennsimulationen/iRacing u.a.
rFactor 2
RIFT
Romance of the Three Kingdoms
Saints Row The Third
Shadow of the Colossus
Skyrim
Spellforce
Starcraft, Stracraft 2
Team Fortress 2

A.4 Hierarchical Clustering of Games from Survey

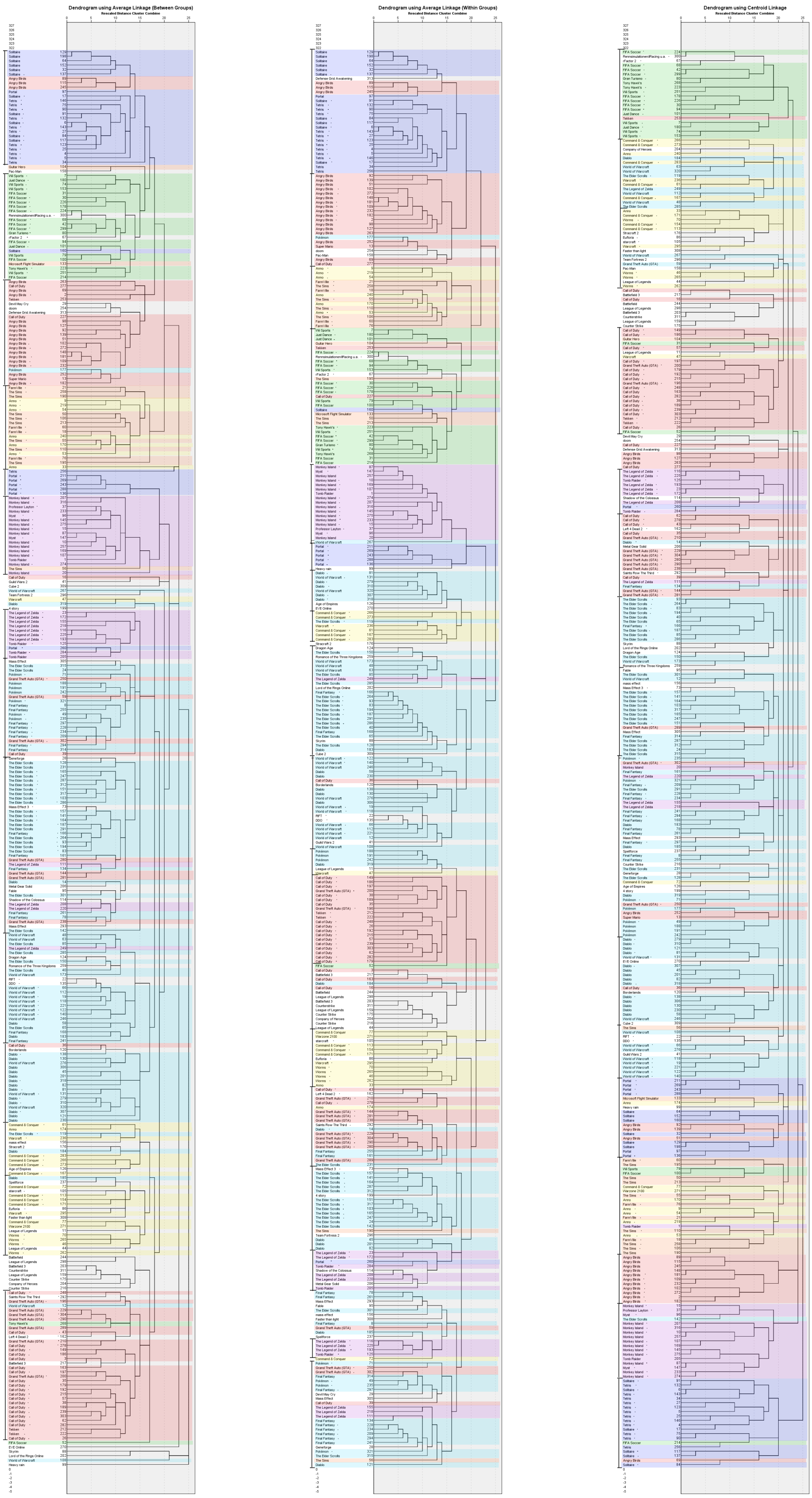


Figure A.1: Dendrograms of different hierarchical clustering methods (method 1 to 3 from left to right).

A.5 Analysis of DEGs for Learning Programming

Game	Learning Topic	Game Genre	Player	Inter-action	Action	(main) Challenge
1. group: Content mediated through challenges: Question/Answer, Multiple Choice, Sorting						
Duck-shooting	remember & identify programming terms	Mini-game (mini online game)	single	mouse?	destroy (shoot ducks)	quiz (pick duck with correct answer), time pressure?
Crossword	remember & identify programming terms	Mini-game (mini online game)	single	mouse & key-board?	place (letters in boxes)	quiz (answer questions), time pressure?
SpaceOut	remember & identify programming terms	Mini-game (mobile arcade game)	single	keys of mobile phone?	move, destroy (shoot answer), character (car)	quiz, obstacles (wrong answers turn into bombs)
Snail, Doggy	remember & identify programming terms	Mini-game (mobile arcade game)	single	keys of mobile phone?	move, character (snail/dog), destroy (eat code lines)	quiz (code lines in correct order), time pressure
2. group: Content mediated through Challenges, Actions & Visualisation						
Wu's Castle	loops and arrays	Mini-game/ Adventure (2d role playing game)	single	mouse & key-board?	move, character, place (loop parameters)	quiz (questios), puzzle (setting for machina with for-loops)
Robozzle (roboz-zle.com)	loops, conditions, functions, recursion	Mini-game (puzzle game)	single (hints from others)	mouse	move (token), place (assemble commands), collect (stars)	puzzle (use commands to navigate token)
List pro-cessing	process list, loop, condition, counter, flag	Mini-game	single	mouse	place (assemble commands)	puzzle (use commands to solve task)
3. group: Like first group, but variant tasks/quests linked by a story						
Saving Sera	loops, print statements, quicksort	Mini-game/ Adventure (exploratory game)	single	mouse & key-board	move, character, place (answers)	quiz (questions)
The Cat-acombs	conditions and loops	Adventure/ Roleplay(3d fantasy role playing game)	single	mouse & key-board?	collect, fight, communicate, character	quiz (questions), search (correct spell hidden amongst wrong ones), opponents (monster)
4. group: Visualisation of executed programs, written by player						
Kernel Panic	fundamental programming skills	Mini-game/ Resource (real-time strategy game)	single/ multi	mouse & key-board?	move (via code), fight (command army), collect (units)	opponents (battle against others), limited resources (units)
Colobot, Ceebot	fundamental programming skills	Action/ Adventure (real time game of strategy)	single	mouse & key-board	move (command robots & machines), character (robot), destroy?	puzzle, opponents?

Table A.6: Analysis of DEGs for learning programming, based on the GEAM (? = unclear from description)

Game	Goal	Rewards	Setting/ World	Perspective	Structure	Reference
1. group: Content mediated through challenges: Question/Answer, Multiple Choice, Sorting						
Duck-shooting	static (finish level)	new level, score	realistic, 2D, not much to explore/detail (sky with ducks, question)	fixed	parts (3 levels)	Ibrahim et al. (2011)
Crossword	static (solve cross-word)	new level, score	abstract, 2D, not much to explore/detail (crossword frame, questions)	fixed	parts (3 levels)	Ibrahim et al. (2011)
SpaceOut	static (finish level)	score, lose resources (3 lives)	realistic, 2D, not much to explore/detail (car on ground, words fall from sky, questions)	fixed	parts? (multiple level?)	Hamid and Fung (2007)
Snail, Doggy	static (finish level)	?	realistic, 2D, not much to explore/detail (answers as food, character, questions)	fixed	parts? (multiple level?)	Hamid and Fung (2007)
2. group: Content mediated through Challenges, Actions & Visualisation						
Wu's Castle	various	next level	fantasy, 2D/3D?, a little to explore/detail (visualisation of loop, code execution)	bound to character/fixed?	continuous + parts (4 levels), story	Eagle and Barnes (2008)
Robozzle (robozzle.com)	static (finish level)	new level	abstract, 2D, not much to explore/detail (board with coloured tiles & token)	fixed	parts (multiple levels)	`www.robozzle.com`
List processing	static (finish level)	new level, praise	abstract, 2D, not much to explore/detail (board with tiles for letters and commands)	fixed	parts (multiple levels)	Heintz and Law (2012)
3. group: Like first group, but variant tasks/quests linked by a story						
Saving Sera	various (different tasks)	new level	fantasy, 3D (between quests)/2D (quests), a little to explore/detail (castle to collect quests, visualisation for each quest)	bound to character (between quests), fixed (quests)	coniuous + parts (quests), story	Barnes et al. (2007)
The Catacombs	various	gain resources	fantasy, 3D, some to explore/detail (catacombs)	bound to character	continuous, story	Barnes et al. (2007)
4. group: Visualisation of executed programs, written by player						
Kernel Panic	various	?	abstract, 3D, some to explore/detail (grid forming a landscape)	?	parts (missions/ battle), story (campaign)	Muratet et al. (2009)
Colobot, Ceebot	various (missions)	new level	realistic/fantasy, 3D, some to explore/detail (space with prepared mission + code window)	bound to character	parts (missions)	`ceebot.com`

Table A.7: Analysis of DEGs for learning programming based on the GEAM (? = unclear from description)

A.6 Questionnaires for DEG Comparison Studies

nr	GEQ item	iGEQ	kidsGEQ	selected	reasoning for exclusion
GEQ-Immersion					
3	I was interested in the game's story	Yes	No	Yes	
14	It was aesthetically pleasing	No	Yes	Yes	
20	I felt imaginative	No	(Yes)	No	small games, limited room to imagine
21	I felt that I could explore things	No	No	No	small games, limited room to explore
30	I found it impressive	Yes	Yes	Yes	
33	It felt like a rich experience	No	No	Yes	
GEQ-Flow					
5	I felt completely absorbed	Yes	No	Yes	
15	I forgot everything around me	Yes	Yes	Yes	
28	I lost track of time	No	No	Yes	
31	I was deeply concentrated in the game	No	(Yes)	No	similar to #5
34	I lost connection with the outside world	No	(Yes)	No	similar to #15
40	I was fully occupied with the game	No	No	No	optional GEQ item
GEQ-Competence					
2	I felt skilful	Yes	(Yes)	No	vague, can be covered by #17
12	I felt strong	No	No	No	vague, can be covered by #17 and #19
17	I was good at it	No	Yes	Yes	
19	I felt successful	Yes	No	Yes	
23	I was fast at reaching the game's targets	No	No	Yes	
42	I felt competent	No	Yes	No	optional GEQ item

Table A.8: Reasoning for items selected for an abbreviate version of the game experience questionnaire (GEQ) used in the two comparison studies. Some of the kidsGEQ items are in brackets, as they were rephrased and did not fully match the original GEQ item.

nr	GEQ item	iGEQ	kidsGEQ	selected	reasoning for exclusion
GEQ-Tension					
7	I felt tense	No	(Yes)	Yes	
9	I felt restless	No	No	Yes	
24	I felt annoyed	No	(Yes)	No	covered by #32
27	I felt irritable	Yes	(No)	No	similar to #7 and #9
32	I felt frustrated	Yes	(Yes)	Yes	
39	I felt pressured	No	No	No	optional GEQ item
GEQ-Challenge(DEG)*					
8	I felt that I was learning	No	No	Yes	
13	I thought it was hard	No	No	No	covered by #29
26	I felt stimulated	Yes	(Yes)	No	vague, somewhat covered by #29
29	I felt challenged	Yes	Yes	Yes	
36	I had to put a lot of effort into it	No	Yes	Yes	
37	I felt time pressure	No	No	No	optional GEQ item
*	It was cognitive demanding			Yes	
GEQ-PositiveAffect					
1	I felt content	Yes	No	No	covered by #6
4	I could laugh about it	No	(Yes)	No	unexpected that students would laugh
6	I felt happy	No	No	No	similar to #41
16	I felt good	Yes	Yes	Yes	
22	I enjoyed it	No	No	Yes	
41	I thought it was fun	No	Yes	Yes	
GEQ-NegativeAffect					
10	I thought about other things	No	No	Yes	
11	I found it tiresome	Yes	Yes	Yes	
18	I felt bored	Yes	Yes	Yes	
25	I was distracted	No	No	No	more behavioural than experiential
35	I was bored by the story	No	No	No	opposite of #3, somewhat covered by #18
38	It gave me a bad mood	No	(Yes)	No	optional GEQ item

Table A.9: Reasoning for items selected for an abbreviate version of the game experience questionnaire (GEQ) used in the two comparison studies. Some of the kidsGEQ items are in brackets, as they were rephrased and did not fully match the original GEQ item. * Item was added specifically for the evaluation of educational games.

Bibliography

Aarseth, E., Smedstad, S. M., and Sunnanå, L. (2003). A multidimensional typology of games. In *DiGRA '03 - Proceedings of the 2003 DiGRA International Conference: Level Up*, volume 2.

Abt, C. C. (1970). *Serious games*. Viking Press.

Adams, E. (2010). *Fundamentals of game design*. New Riders, 2nd edition.

Adams, E. (2014). *Fundamentals of game design*. New Riders, 3rd edition.

Allport, G. W. and Odbert, H. S. (1936). Trait-names: A psycho-lexical study. *Psychological monographs*, 47(1):i–171.

Amory, A. (2007). Game object model version II: a theoretical framework for educational game development. *Educational Technology Research and Development*, 55(1):51–77.

Amory, A., Naicker, K., Vincent, J., and Adams, C. (1999). The use of computer games as an educational tool: identification of appropriate game types and game elements. *British Journal of Educational Technology*, 30(4):311–321.

Anagnostou, K. and Pappa, A. (2011). Developing videogames for physics education. In Felicia, P., editor, *Handbook of Research on Improving Learning and Motivation through Educational Games: Multidisciplinary Approaches*, chapter 43, pages 939–960. IGI Global.

Apperley, T. H. (2006). Genre and game studies: toward a critical approach to video game genres. *Simulation & Gaming*, 37(1):6–23.

Astleitner, H. (2000). Designing emotionally sound instruction: The FEASP-approach. *Instructional Science*, 28(3):169–198.

Avedon, E. M. (1971). The structural elements of games. *The study of games*, pages 419–426.

Barbe, W. B. and Milone Jr, M. N. (1981). What we know about modality strengths. *Educational Leadership*, 38(5):378–380.

Barbosa, A. F. S. and Silva, F. G. M. (2011). Serious games - design and development of OxyBlood. In *Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology*, ACE '11, pages 15:1–15:8. ACM.

Barnes, T., Powell, E., Chaffin, A., and Lipford, H. (2008). Game2Learn: improving the motivation of CS1 students. In *Proceedings of the 3rd international conference on Game development in computer science education*, GDCSE '08, pages 1–5. ACM.

Barnes, T., Richter, H., Powell, E., Chaffin, A., and Godwin, A. (2007). Game2Learn: building CS1 learning games for retention. *SIGCSE Bull.*, 39(3):121–125.

Bateman, C., Lowenhaupt, R., and Nacke, L. E. (2011). Player typology in theory and practice. In *Proceedings of DiGRA 2011 Conference: Think Design Play.*

Bates, B. (2004). *Game design.* Premier Press, 2nd edition.

Baumgarten, F. (1933). *Die Charaktereigenschaften.* Francke.

Bedwell, W. L., Pavlas, D., Heyne, K., Lazzara, E. H., and Salas, E. (2012). Toward a taxonomy linking game attributes to learning: an empirical study. *Simulation & Gaming*, 43(6):729–760.

Bennedsen, J. and Caspersen, M. E. (2005). An investigation of potential success factors for an introductory model-driven programming course. In *Proceedings of the First International Workshop on Computing Education Research*, ICER '05, pages 155–163. ACM.

Bergeron, B. P. (2006). *Developing Serious Games.* Charles River Media.

Bergervoet, E., van der Sluis, F., van Dijk, B., and Nijholt, A. (2011). Let the game do the talking: The influence of explicitness and game behavior on comprehension in an educational computer game. In *CW'11, 2011 International Conference on Cyberworlds*, pages 120–127.

Björk, S. (2013). On the foundations of digital games. In *Proceedings of the 8th International Conference on the Foundations of Digital Games (FDG 2013)*, pages 1–8.

Björk, S. and Holopainen, J. (2005). *Patterns in game design.* Charles River Media.

Bopp, M. M. (2008). Storytelling and motivation in serious games. *Part of the Final Consolidated Research Report of the Enhanced Learning Experience and Knowledge Transfer - Project (ELEKTRA).*

Brathwaite, B. and Schreiber, I. (2008). *Challenges for Game Designers.* Charles River Media.

Brown, E. and Cairns, P. (2004). A grounded investigation of game immersion. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '04, pages 1297–1300. ACM.

Bruckman, A. (1999). Can educational be fun? In *Game Developer's Conference, San Jose, California.*

Byrne, P. and Lyons, G. (2001). The effect of student attributes on success in programming. *SIGCSE Bull.*, 33(3):49–52.

Caillois, R. (1961). *Man, Play and Games.* Translated by Meyer Barash. New York, Free Press of Glencoe.

Calvillo-Gámez, E. H., Cairns, P., and Cox, A. L. (2010). Assessing the core elements of the gaming experience. In Bernhaupt, R., editor, *Evaluating User Experience in Games*, pages 47–71. Springer.

Chong, Y., Wong, M., and Thomson Fredrik, E. (2005). The impact of learning styles on the effectiveness of digital games in education. In *Proceedings of the Symposium on Information Technology in Education, KDU College, Patailing Java, Malaysia.*

cmdscale (2015). R library for multidimensional scaling, available at: https://stat.ethz.ch/r-manual/r-devel/library/stats/html/cmdscale.html (accessed: July 2015).

Coffield, F., Moseley, D., Hall, E., Ecclestone, K., et al. (2004a). *Learning styles and pedagogy in post-16 learning: A systematic and critical review.* Learning and Skills Research Centre.

Coffield, F., Moseley, D., Hall, E., Ecclestone, K., et al. (2004b). *Should we be using learning styles? What research has to say to practice.* Learning and Skills Research Centre.

Costa Jr, P. T. and McCrae, R. R. (1995). Domains and facets: Hierarchical personality assessment using the revised NEO personality inventory. *Journal of personality assessment*, 64(1):21–50.

Costikyan, G. (1994). I have no words and I must design. Interactive Fantasy # 2. *British roleplaying journal*.

Cowley, B., Charles, D., Black, M., and Hickey, R. (2008). Toward an understanding of flow in video games. *Comput. Entertain.*, 6(2):20:1–20:27.

Cox, A., Cairns, P., Shah, P., and Carroll, M. (2012). Not doing but thinking: The role of challenge in the gaming experience. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 79–88. ACM.

Crawford, C. (1982). *The Art of Computer Game Design (Electronic version)*. Department of History, Washington State University Vancouver.

Csikszentmihalyi, M. (1975). *Beyond boredom and anxiety*. Jossey-Bass.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. Harper and Row.

Csikszentmihalyi, M. (1997). *Finding flow: The psychology of engagement with everyday life*. Basic Books.

Dahlskog, S., Kamstrup, A., and Aarseth, E. (2009). Mapping the game landscape: Locating genres using functional classification. In *DiGRA '09 - Proceedings of the 2009 DiGRA International Conference: Breaking New Ground: Innovation in Games, Play, Practice and Theory*, volume 5. Brunel University.

de Freitas, S. and Neumann, T. (2009). The use of 'exploratory learning' for supporting immersive learning in virtual environments. *Computers & Education*, 52(2):343–352.

Deterding, S., Dixon, D., Khaled, R., and Nacke, L. (2011). From game design elements to gamefulness: Defining "gamification". In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, MindTrek '11, pages 9–15. ACM.

Dimitrov, D. M. and Rumrill Jr., P. D. (2003). Pretest-posttest designs and measurement of change. *Work IOS Press*, 20(2):159–165.

Djaouti, D., Alvarez, J., and Jessel, J.-P. (2011). Classifying serious games: The G/P/S model. In Felicia, P., editor, *Handbook of Research on Improving Learning and Motivation through Educational Games: Multidisciplinary Approaches*, chapter 6, pages 118–136. IGI Global.

Djaouti, D., Alvarez, J., Jessel, J.-P., Methel, G., and Molinier, P. (2007). The nature of gameplay: a videogame classification. In *Symposium proceedings "Cybergames 2007", Manchester, UK.*

Djaouti, D., Alvarez, J., Jessel, J.-P., Methel, G., and Molinier, P. (2008). A gameplay definition through videogame classification. *International Journal of Computer Games Technology*, 2008:4:1–4:7.

Dominguez, M., Young, R. M., and Roller, S. (2011). Automatic identification and generation of highlight cinematics for 3D games. In *Proceedings of the 6th International Conference on Foundations of Digital Games*, FDG '11, pages 259–261. ACM.

Dunn, R. S. and Dunn, K. J. (1979). Learning styles/teaching styles: Should they... can they... be matched? *Educational leadership*, 36(4):238–244.

Eagle, M. and Barnes, T. (2008). Wu's castle: Teaching arrays and loops in a game. In *Proceedings of the 13th Annual Conference on Innovation and Technology in Computer Science Education*, ITiCSE '08, pages 245–249. ACM.

Eagle, M. and Barnes, T. (2009a). Evaluation of a game-based lab assignment. In *Proceedings of the 4th International Conference on Foundations of Digital Games*, FDG '09, pages 64–70. ACM.

Eagle, M. and Barnes, T. (2009b). Experimental evaluation of an educational game for improved learning in introductory computing. In *Proceedings of the 40th ACM Technical Symposium on Computer Science Education*, SIGCSE '09, pages 321–325. ACM.

Edge, D., Cheng, K.-Y., and Whitney, M. (2013). SpatialEase: Learning language through body motion. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 469–472. ACM.

Egenfeldt-Nielsen, S. (2006). Overview of research on the educational use of video games. *Digital Kompetanse*, 1(3):184–213.

Egenfeldt-Nielsen, S., Smith, J. H., and Tosca, S. P. (2013). *Understanding video games: The essential introduction*. Routledge, 2nd edition.

Elverdam, C. and Aarseth, E. (2007). Game classification and game design: Construction through critical analysis. *Games and Culture*, 2(1):3–22.

Ermi, L. and Mäyrä, F. (2005). Fundamental components of the gameplay experience: Analysing immersion. *DiGRA '05 - Proceedings of the 2005 DiGRA International Conference: Changing Views: Worlds in Play*, 3.

Felder, R. and Soloman, B. (2015). Index of learning styles. Retrieved from http://www.ncsu.edu/felder-public/ILSpage.html (accessed: June 2015).

Felder, R. M. and Silverman, L. K. (1988). Learning and teaching styles in engineering education. *Engineering education*, 78(7):674–681.

Felder, R. M. and Spurlin, J. (2005). Applications, reliability and validity of the index of learning styles. *International Journal of Engineering Education*, 21(1):103–112.

Ferro, L. S., Walz, S. P., and Greuter, S. (2013). Towards personalised, gamified systems: An investigation into game design, personality and player typologies. In *Proceedings of The 9th Australasian Conference on Interactive Entertainment: Matters of Life and Death*, IE '13, pages 7:1–7:6. ACM.

Fleming, N. D. and Mills, C. (1992). Not another inventory, rather a catalyst for reflection. *To Improve the Academy*, 11:137.

Fotouhi-Ghazvini, F., Earnshaw, R. A., Robison, D., and Excell, P. S. (2009). Designing augmented reality games for mobile learning using an instructional-motivational paradigm. In *CW'09, 2009 International Conference on Cyberworlds*, pages 312–319. IEEE.

Fransella, F., Bell, R., and Bannister, D. (2004). *A Manual for Repertory Grid Technique*. John Wiley & Sons, 2nd edition.

Frazer, A. (2010). *Towards better gameplay in educational computer games: a PhD thesis*. PhD thesis, University of Southampton.

Frazer, A., Argles, D., and Wills, G. (2007). Is less actually more? The usefulness of educational mini-games. In *ICALT 2007, Seventh IEEE International Conference on Advanced Learning Technologies 2007*, pages 533–537. IEEE.

Frazer, A., Argles, D., and Wills, G. (2008). The same, but different: The educational affordances of different gaming genres. *ICALT '08, Eighth IEEE International Conference on Advanced Learning Technologies 2008*, pages 891–893.

Fullerton, T. (2008). *Game Design Workshop: A Playcentric Approach to Creating Innovative Games*. Morgan Kaufmann, 2nd edition.

Garner, S., Haden, P., and Robins, A. (2005). My program is correct but it doesn't run: A preliminary investigation of novice programmers' problems. In *Proceedings of the 7th Australasian Conference on Computing Education, ACE 2005*, volume 42, pages 173–180. Australian Computer Society, Inc.

Garris, R., Ahlers, R., and Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. *Simulation & Gaming*, 33(4):441–467.

Giannakos, M. N. (2013). Enjoy and learn with educational games: Examining factors affecting learning performance. *Computers & Education*, 68:429–439.

Goldberg, L. R. (1990). An alternative "description of personality": The Big-Five factor structure. *Journal of personality and social psychology*, 59(6):1216–1229.

Grace, L. (2005). Game type and game genre.

Grace, L. D. (2012). Making and analyzing games: Not as art, but as literature. In *Proceeding of the 16th International Academic MindTrek Conference*, MindTrek '12, pages 237–240. ACM.

Habgood, M. P. J. (2007). *The effective integration of digital games and learning content*. PhD thesis, University of Nottingham.

Halverson, R. (2005). What can K-12 school leaders learn from video games and gaming? *Innovate: Journal of Online Education*, 1(6).

Hamid, S. H. A. and Fung, L. Y. (2007). Learn programming by using mobile edutainment game approach. In *DIGITEL '07, The First IEEE International Workshop on Digital Game and Intelligent Toy Enhanced Learning, 2007*, pages 170–172, Washington, DC, USA. IEEE.

Harteveld, C. (2011). *Triadic game design: Balancing reality, meaning, and play.* Springer.

Harteveld, C. and Bekebrede, G. (2011). Learning in single-versus multiplayer games: The more the merrier? *Simulation & Gaming*, 42(1):43–63.

Hassan, O. A. (2011). Learning theories and assessment methodologies - an engineering educational perspective. *European Journal of Engineering Education*, 36(4):327–339.

Hassenzahl, M. and Monk, A. (2010). The inference of perceived usability from beauty. *Human–Computer Interaction*, 25(3):235–260.

Heintz, S. and Law, E. L.-C. (2012). Evaluating design elements for digital educational games on programming: A pilot study. In *Proceedings of the 26th Annual BCS Interaction Specialist Group Conference on People and Computers*, BCS-HCI '12, pages 245–250, Swinton, UK. British Computer Society.

Heintz, S. and Law, E. L.-C. (2015a). The game genre map: A revised game classification. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play*, CHI PLAY '15, pages 175–184. ACM.

Heintz, S. and Law, E. L.-C. (2015b). Game Elements-Attributes Model: a first step towards a structured comparison of educational games. In *DiGRA'15 - Proceedings of the 2015 DiGRA International Conference*. Digital Games Research Association.

Hense, J. and Mandl, H. (2014). Learning in or with games? In *Digital Systems for Open Access to Formal and Informal Learning*, chapter 12, pages 181–193. Springer.

Herrington, J. and Oliver, R. (1995). Critical characteristics of situated learning: Implications for the instructional design of multimedia. *ASCILITE 1995 Conference*, pages 253–262.

Hohn, R. L. (1995). *Classroom learning & teaching*. Longman.

Huizinga, J. (1949). *Homo ludens: A study of the play-element in culture*. Routledge and Kegan Paul.

Hunicke, R., LeBlanc, M., and Zubek, R. (2004). MDA: A formal approach to game design and game research. In *Proceedings of the AAAI Workshop on Challenges in Game AI*.

Huotari, K. and Hamari, J. (2012). Defining gamification: A service marketing perspective. In *Proceeding of the 16th International Academic MindTrek Conference*, MindTrek '12, pages 17–22. ACM.

Hwang, G.-J., Sung, H.-Y., Hung, C.-M., and Huang, I. (2013). A learning style perspective to investigate the necessity of developing adaptive learning systems. *Educational Technology & Society*, 16(2):188–197.

Ibrahim, R., Yusoff, R. C. M., Mohamed@Omar, H., and Jaafar, A. (2011). Students perceptions of using educational games to learn introductory programming. *Computer and Information Science*, 4(1):205–216.

IJsselsteijn, W., Poels, K., and De Kort, Y. (2008). The game experience questionnaire: Development of a self-report measure to assess player experiences of digital games. *FUGA technical report, deliverable 3.3, TU Eindhoven, Eindhoven, The Netherlands.*

Isbister, K., Flanagan, M., and Hash, C. (2010). Designing games for learning: Insights from conversations with designers. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 2041–2044. ACM.

Jantke, K. P. (2006). Games that do not exist: Communication design beyond the current limits. In *Proceedings of the 24th Annual ACM International Conference on Design of Communication*, SIGDOC '06, pages 35–42. ACM.

Jantke, K. P. (2010). Toward a taxonomy of game based learning. In *2010 IEEE International Conference on Progress in Informatics and Computing (PIC)*, volume 2, pages 858–862. IEEE.

Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., and Walton, A. (2008). Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies*, 66(9):641–661.

John, O. P., Donahue, E. M., and Kentle, R. L. (1991). *The Big Five Inventory – Versions 4a and 54*. Berkeley, CA: University of California, Berkeley, Institute of Personality and Social Research.

John, O. P., Naumann, L. P., and Soto, C. J. (2008). Paradigm shift to the integrative big-five trait taxonomy: History, measurement, and conceptual issues. In John, O. P., Robins, R. W., and Pervin, L. A., editors, *Handbook of personality: Theory and research*, chapter 4, pages 114–158. New York, NY: Guilford Press, 3rd edition.

John, O. P. and Srivastava, S. (1999). The big five trait taxonomy: History, measurement, and theoretical perspectives. In Pervin, L. A. and John, O. P., editors, *Handbook of personality: Theory and research*, pages 102–138. New York, NY: Guilford Press, 2nd edition.

Johnson, D., Wyeth, P., Sweetser, P., and Gardner, J. (2012). Personality, genre and videogame play experience. In *Proceedings of the 4th International Conference on Fun and Games*, FnG '12, pages 117–120. ACM.

Jones, M. G. (1998). Creating electronic learning environments: Games, flow, and the user interface. In *Proceedings of Selected Research and Development Presentations at the National Convention of the Association for Educational Communications and Technology (AECT)*, pages 205–214.

Jung, C. G. (1921). *Psychologische Typen*. Rascher.

Juul, J. (2003). The game, the player, the world: Looking for a heart of gameness. In *Proceedings of the 2003 DiGRA International Conference: Level Up*.

Juul, J. (2005). *Half-real: Video games between real rules and fictional worlds*. MIT Press.

Juul, J. (2011). *Half-real: Video games between real rules and fictional worlds*. MIT Press, paperback edition.

Kelley, D. (1988). *The Art of Reasoning*. Norton.

Kiili, K. (2007). Foundation for problem-based gaming. *British Journal of Educational Technology*, 38(3):394–404.

Kinnunen, P. and Malmi, L. (2008). CS minors in a CS1 course. In *Proceedings of the Fourth International Workshop on Computing Education Research*, ICER '08, pages 79–90. ACM.

Kinnunen, P., McCartney, R., Murphy, L., and Thomas, L. (2007). Through the eyes of instructors: A phenomenographic investigation of student success. In *Proceedings of the Third International Workshop on Computing Education Research*, ICER '07, pages 61–72. ACM.

Klabbers, J. H. G. (2003). The gaming landscape: A taxonomy for classifying games and simulations. In *Proceedings of the 2003 DiGRA International Conference: Level Up*.

Kolb, A. Y. and Kolb, D. A. (2005). *The Kolb Learning Style Inventory, Version 3.1 2005 technical specifications*. Boston, MA: Hay Group.

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Prentice Hall, Englewood Cliffs, NJ.

Kosmadoudi, Z., Lim, T., Ritchie, J., Louchart, S., Liu, Y., and Sung, R. (2013). Engineering design using game-enhanced CAD: The potential to augment the user experience with game elements. *Computer-Aided Design*, 45(3):777–795.

Koster, R. (2005). *A theory of fun for game design*. Paraglyph Press.

Krejtz, K., Biele, C., Chrzastowski, D., Kopacz, A., Niedzielska, A., Toczyski, P., and Duchowski, A. (2014). Gaze-controlled gaming: Immersive and difficult but not cognitively overloading. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*, UbiComp '14 Adjunct, pages 1123–1129. ACM.

Krippendorff, K. (2013). *Content Analysis: An Introduction to Its Methodology*. SAGE Publications, 3rd edition.

Lahtinen, E., Ala-Mutka, K., and Järvinen, H.-M. (2005). A study of the difficulties of novice programmers. In *Proceedings of the 10th Annual SIGCSE Conference on Innovation and Technology in Computer Science Education*, ITiCSE '05, pages 14–18. ACM.

Lainema, T. (2009). Perspective making: Constructivism as a meaning-making structure for simulation gaming. *Simulation & Gaming*, 40(1):48–67.

Lave, J. (1991). Situating learning in communities of practice. In Resnick, L. B., Levine, J. M., and Teasley, S., editors, *Perspectives on Socially Shared Cognition*, chapter 4, pages 63–82. Washington, D. C.: American Psychological Association.

Lecky-Thompson, G. W. (2007). *Video Game Design Revealed*. Charles River Media.

Lepper, M. R., Corpus, J. H., and Iyengar, S. S. (2005). Intrinsic and extrinsic motivational orientations in the classroom: Age differences and academic correlates. *Journal of Educational Psychology*, 97(2):184–196.

Lewis, J. P., McGuire, M., and Fox, P. (2007). Mapping the mental space of game genres. In *Proceedings of the 2007 ACM SIGGRAPH Symposium on Video Games*, Sandbox '07, pages 103–108. ACM.

Linehan, C., Kirman, B., Lawson, S., and Chan, G. (2011). Practical, appropriate, empirically-validated guidelines for designing educational games. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 1979–1988. ACM.

Malone, T. W. (1981). Toward a theory of intrinsically motivating instruction*. *Cognitive Science*, 5(4):333–369.

Malone, T. W. and Lepper, M. R. (1987). Making learning fun: A taxonomy of intrinsic motivations for learning. In Snow, R. E. and Farr, M. J., editors, *Aptitude, Learning, and Instruction*, volume 3, pages 223–253. Lawrence Erlbaum Associates.

McCann, S. (2009). Game genres demystified. *Library Journal*, 134(1):56.

McEwan, G., Gutwin, C., Mandryk, R. L., and Nacke, L. (2012). "I'm just here to play games": Social dynamics and sociality in an online game site. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*, CSCW '12, pages 549–558. ACM.

Mekler, E. D., Bopp, J. A., Tuch, A. N., and Opwis, K. (2014). A systematic review of quantitative studies on the enjoyment of digital entertainment games. In *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems*, CHI '14, pages 927–936. ACM.

Michael, D. and Chen, S. (2005). *Serious games: Games that educate, train, and inform*. Thomason Course Technology PTR.

Milne, I. and Rowe, G. (2002). Difficulties in learning and teaching programming - views of students and tutors. *Education and Information Technologies*, 7(1):55–66.

Mooi, E. and Sarstedt, M. (2011). *A Concise Guide to Market Research: The Process, Data, and Methods Using IBM SPSS Statistics*. Springer.

Moreno-Ger, P., Burgos, D., Martínez-Ortiz, I., Sierra, J. L., and Fernández-Manjón, B. (2008). Educational game design for online education. *Computers in Human Behavior*, 24(6):2530–2540.

Muratet, M., Torguet, P., Jessel, J.-P., and Viallet, F. (2009). Towards a serious game to help students learn computer programming. *International Journal of Computer Games Technology*, 2009:1–12.

Myers, I. B. (1962). *Manual: The Myers-Briggs Type Indicator*. Educational Testing Service.

Na, B. (2006). A proposal for new genre classification of digital games. *Asia Culture Forum 2006*.

Nacke, L. E. (2009). *Affective Ludology: Scientific Measurement of User Experience in Interactive Entertainment*. PhD thesis, Blekinge Institute of Technology.

Nakamura, J. and Csikszentmihalyi, M. (2002). The concept of flow. In *Handbook of positive psychology*, chapter 7, pages 89–105. Oxford University Press.

Nikula, U., Gotel, O., and Kasurinen, J. (2011). A motivation guided holistic rehabilitation of the first programming course. *ACM Transactions on Computing Education*, 11(4):24:1–24:38.

Nordin, A. I., Denisova, A., and Cairns, P. (2014). Too many questionnaires: Measuring player experience whilst playing digital games. In *Seventh York Doctoral Symposium on Computer Science and Electronics*.

Orr, K. and McGuinness, C. (2014). What is the "learning" in games-based learning? In Connolly, T. M., Hainey, T., Boyle, E., Baxter, G., and Moreno-Ger, P., editors, *Psychology, Pedagogy, and Assessment in Serious Games*, chapter 11, pages 221–242. IGI Global.

Owen, M. (2004). An anatomy of games: A discussion paper. *Futurelab*.

Papert, S. (1986). *Constructionism: A new opportunity for elementary science education*. Massachusetts Institute of Technology, Media Laboratory, Epistemology and Learning Group.

Parlett, D. (1999). *The Oxford History of Board Games*. Oxford University Press.

Pashler, H., McDaniel, M., Rohrer, D., and Bjork, R. (2009). Learning styles: Concepts and evidence. *Psychological Science in the Public Interest*, 9(3):105–119.

Paul, F. C., Goh, C., and Yap, K. (2015). Get creative with learning: Word out! a full body interactive game. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA '15, pages 81–84. ACM.

Perry, D. and DeMaria, R. (2009). *David Perry on Game Design: A Brainstorming Toolbox*. Charles River Media.

Pintrich, P. R. and De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82(1):33–40.

Piteira, M. and Costa, C. (2012). Computer programming and novice programmers. In *Proceedings of the Workshop on Information Systems and Design of Communication*, ISDOC '12, pages 51–53. ACM.

Poels, K., Ijsselsteijn, W., and de Kort, Y. (2008). Development of the Kids Game Experience Questionnaire. *Poster presented at the Meaningful Play Conference 2008*.

Prensky, M. (2001). *Digital Game-Based Learning*. McGraw Hill.

Prensky, M. (2007). *Digital Game-Based Learning*. Paragon House.

Pritchard, A. (2014). *Ways of learning: Learning theories and learning styles in the classroom*. Routledge, 3rd edition.

Quenk, N. L. (2009). *Essentials of Myers-Briggs Type Indicator Assessment*. Essentials of Psychological Assessment. Wiley, 2nd edition.

Rammstedt, B. and John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality*, 41(1):203–212.

Rapeepisarn, K., Wong, K. W., Fung, C. C., and Khine, M. S. (2008). The relationship between game genres, learning techniques and learning styles in educational computer games. In Pan, Z., Zhang, X., Rhalibi, A., Woo, W., and Li, Y., editors, *Proceedings of the 3rd International Conference on Technologies for E-Learning and Digital Entertainment*, Edutainment '08, pages 497–508, Berlin, Heidelberg. Springer.

Rausch, J. R., Maxwell, S. E., and Kelley, K. (2003). Analytic methods for questions pertaining to a randomized pretest, posttest, follow-up design. *Journal of Clinical Child and Adolescent Psychology*, 32(3):467–486.

Rawn, R. W. A. and Brodbeck, D. R. (2008). Examining the relationship between game type, player disposition and aggression. In *Proceedings of the 2008 Conference on Future Play: Research, Play, Share*, Future Play '08, pages 208–211. ACM.

Raybourn, E. M. (2007). Applying simulation experience design methods to creating serious game-based adaptive training systems. *Interacting with Computers*, 19(2):206–214.

Reiss, S. (2000). *Who Am I? The 16 basic desires that motivate our actions and define our personalities*. Tarcher/Putnam.

Reiss, S. (2004). Multifaceted nature of intrinsic motivation: The theory of 16 basic desires. *Review of General Psychology*, 8(3):179–193.

Rienties, B. and Rivers, B. A. (2014). Measuring and understanding learner emotions: Evidence and prospects. *Learning Analytics Community Exchange*.

Ritterfeld, U., Cody, M., and Vorderer, P. (2009). *Serious Games: Mechanisms and Effects*. Routledge.

Rogers, S. (2010). *Level Up! The guide to great video game design*. Wiley.

Rollings, A. and Adams, E. (2003). *Andrew Rollings and Ernest Adams on Game Design*. New Riders.

Romero, M. and Barma, S. (2015). Teaching pre-service teachers to integrate Serious Games in the primary education curriculum. *International Journal of Serious Games*, 2(1):45–54.

Rosati, P., Dean, R. K., and Rodman, S. M. (1988). A study of the relationship between students' learning styles and instructors' lecture styles. *IEEE Transactions on Education*, 31(3):208–212.

Rountree, N., Rountree, J., and Robins, A. (2002). Predictors of success and failure in a CS1 course. *SIGCSE Bulletin*, 34(4):121–124.

Ruch, A. W. (2012). Grand Theft Auto IV: Liberty City and modernist literature. *Games and Culture*, 7(5):331–348.

Ryan, R. M. and Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, 25(1):54–67.

Salen, K. and Zimmerman, E. (2004). *Rules of Play: Game Design Fundamentals*. MIT Press.

Sawyer, B. (2002). *Serious Games: Improving Public Policy through Game-based Learning and Simulation*. Woodrow Wilson International Center for Scholars.

Sawyer, B. and Smith, P. (2008). Serious games taxonomy. In *Slides from the Serious Games Summit at the Game Developers Conference*.

Schaubhut, N. A., Herk, N. A., and Thompson, R. C. (2009). *MBTI® Form M Manual Supplement*.

Schell, J. (2008). *The Art of Game Design: A Book of Lenses*. Morgan Kaufmann.

Schmitz, B., Specht, M., and Klemke, R. (2012). An analysis of the educational potential of augmented reality games for learning. In Specht, M., Sharples, M., and Multisilta, J., editors, *Proceedings of the 11th International Conference on Mobile and Contextual Learning 2012*, pages 140–147.

Schneider, E. F., Lang, A., Shin, M., and Bradley, S. D. (2004). Death with a story: How story impacts emotional, motivational, and physiological responses to first-person shooter video games. *Human Communication Research*, 30(3):361–375.

Siang, A. C. and Rao, R. K. (2003). Theories of learning: A computer game perspective. In *Proceedings of the IEEE Fifth International Symposium on Multimedia Software Engineering*, ISMSE'03, pages 239–245. IEEE.

Skalski, P., Tamborini, R., Shelton, A., Buncher, M., and Lindmark, P. (2010). Mapping the road to fun: Natural video game controllers, presence, and game enjoyment. *New Media & Society*, 13(2):224–242.

Sommeregger, P. and Kellner, G. (2012). Brief guidelines for Educational Adventure Games Creation (EAGC). In *IEEE Fourth International Conference On Digital Game And Intelligent Toy Enhanced Learning, 2012*, DIGITEL, pages 120–122. IEEE.

Suits, B. (1990). *The Grasshopper: Games, Life and Utopia*. David R Godine.

Susi, T., Johannesson, M., and Backlund, P. (2007). Serious games - an overview. Technical report, Technical report, HS-IKI-TR-07-001, University of Skövde, Sweden.

Sweetser, P. and Wyeth, P. (2005). GameFlow: A model for evaluating player enjoyment in games. *Computers in Entertainment*, 3(3):1–24.

Szilas, N. and Acosta, M. (2011). A theoretical background for educational video games: Games, signs, knowledge. In Felicia, P., editor, *Handbook of Research on Improving Learning and Motivation through Educational Games: Multidisciplinary Approaches*, chapter 11. IGI Global.

Tabachnick, B. G. and Fidell, L. S. (2007). *Using Multivariate Statistics*. Pearson, 5th edition.

Tan, P.-H., Ting, C.-Y., and Ling, S.-W. (2009). Learning difficulties in programming courses: Undergraduates' perspective and perception. In *2009 International Conference on Computer Technology and Development*, volume 1 of *ICCTD'09*, pages 42–46. IEEE.

van den Hoogen, W., IJsselsteijn, W., and de Kort, Y. (2008). Exploring behavioral expressions of player experience in digital games. In *Proceedings of the Workshop on Facial and Bodily Expression for Control and Adaptation of Games*, ECAG 2008, pages 11–19.

VGChartz (2012). Homepage VGChartz, available at: www.vgchartz.com/gamedb/ (accessed: 20 December 2012).

Villagrasa, S. and Duran, J. (2013). Gamification for learning 3D computer graphics arts. In *Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality*, TEEM '13, pages 429–433. ACM.

Voulgari, I. and Komis, V. (2013). Player characters as devices for supporting learning in massively multiplayer online games. In *Workshop Proceedings of the 8th International Conference on the Foundations of Digital Games*, FDG'03.

Wherry, R. J. and Bartlett, C. J. (1982). The control of bias in ratings: A theory of rating. *Personnel Psychology*, 35(3):521–551.

Williams, M. N., Grajales, C. A. G., and Kurkiewicz, D. (2013). Assumptions of multiple regression: Correcting two misconceptions. *Practical Assessment, Research & Evaluation*, 18(11):1–14.

Wilson, K. A., Bedwell, W. L., Lazzara, E. H., Salas, E., Burke, C. S., Estock, J. L., Orvis, K. L., and Conkey, C. (2009). Relationships between game attributes and learning outcomes: Review and research proposals. *Simulation & Gaming*, 40(2):217–266.

Winn, B. M. (2009). The design, play, and experience framework. In *Handbook of Research on Effective Electronic Gaming in Education*, chapter 58, pages 1010–1024. IGI Global.

Wolf, M. J. P. (2001). Genre and the video game. In *The Medium of the Video Game*, chapter 6, pages 113–134. University of Texas Press.

Woollard, J. (2010). *Psychology for the Classroom: Behaviourism*. Routledge.

Woszczynski, A. B., Haddad, H. M., and Zgambo, A. F. (2005). Towards a model of student success in programming courses. In *Proceedings of the 43rd Annual Southeast Regional Conference*, volume 1 of *ACM-SE 43*, pages 301–302. ACM.

Yap, K., Zheng, C., Tay, A., Yen, C.-C., and Do, E. Y.-L. (2015). Word out!: Learning the alphabet through full body interactions. In *Proceedings of the 6th Augmented Human International Conference*, AH '15, pages 101–108. ACM.

Yau, J. Y.-K. and Joy, M. (2004). Introducing Java: the case for fundamentals-first. In *International Conference on Education and Information Systems, Technologies and Applications*, volume 2 of *EISTA 2004*, pages 229–234.

Yusoff, R. C. M., Zaman, H. B., and Ahmad, A. (2010). Design a situated learning environment using mixed reality technology - a case study. *World Academy of Science, Engineering and Technology*, 47:884–889.

Zammitto, V. L. (2010). Gamers' personality and their gaming preferences. Master's thesis, School of Interactive Arts and Technology, Simon Fraser University.

Zyda, M. (2005). From visual simulation to virtual reality to games. *Computer*, 38(9):25–32.