The Role of Fun in Learning

The Role of Fun in Learning

Doctoral Dissertation

Gabriella Tisza

Department of Industrial Design
Eindhoven University of Technology
2022

ISBN: 978-90-386-5635-9
A catalogue record is available from the Eindhoven University of Technology Library

Cover designed by Gabriella Tisza

The Role of Fun in Learning

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Technische Universiteit Eindhoven, op gezag van de rector magnificus prof.dr.ir. F.P.T. Baaijens, voor een commissie aangewezen door het College voor Promoties, in het openbaar te verdedigen op donderdag 19 januari 2023 om 13:30 uur

door

GABRIELLA TISZA

geboren te Budapest, Hungary

Dit proefschrift is goedgekeurd door de promoteren en de samenstelling van de promotiecommissie is als volgt:

Voorzitter:	prof.dr. L. L. Chen
1e promotor:	prof.dr. P. Markopoulos
2e promotor:	prof.dr.ir. M.M. Bekker
Leden:	prof.dr. P. Markopoulos
	prof.dr.ir. M.M. Bekker
	prof.dr. R.H.M. De Groot (Open Universiteit)
	prof.dr. P.J. den Brok (Wageningen University & Research)
	dr. R. Taconis
	dr. B. Zaman (Katholieke Universiteit Leuven)
	prof. dr. J.C. Read (University of Central Lancashire)

Het onderzoek of ontwerp dat dit in dit proefschrift wordt beschreven is uitgevoerd in overeenstemming met de TU/e Gedragscode Wetenschapsbeoefening.

Summary

Researchers and practitioners in learning sciences, educational technology and child-computer interaction often argue that fun is an essential element of learning. Substantial research efforts are directed towards making learning enjoyable to facilitate engagement in the learning process and to improve the learning outcomes. Despite such wide interest, there has been little systematic effort to provide quantifiable evidence regarding the role of fun in learning not least due to a lack of a common theoretical framework for defining the concept of fun and for supporting its measurement. Aiming to fill this gap we set out first to define fun and to develop a measurement tool for the assessment of fun in learning activities, and particularly for adolescents. We then used this measurement tool to assess fun in various learning activities and aimed to quantify its impact on different learning outcomes and in different contexts.

This thesis addresses four research questions: (1) What is the role of fun in learning? (2) What is fun? (3) How can we measure fun? and (4) Is the relationship between fun and learning affected by personal factors such as self-regulation and environmental factors such as socioeconomic status? The context of the work while investigating these questions was both the formal and non-formal learning setting, with focus on STEM (Science, Technology, Engineering, and Mathematics) education. These questions are answered by evaluating interventions in quasi experiments, using a mix of methods, mostly quantitative.

The work starts by identifying the key aspects of a fun activity, followed by a theoretically grounded and empirically validated definition of fun. The theories that provided ground for this work were gamification and game enjoyment, play, flow, Control-Value Theory, and Self-Determination Theory. We then describe the development of a related measurement tool, FunQ. Using FunQ, we investigate the role of fun in various settings, such as playful coding workshops and digital game-based learning (DGBL) classes, which activities took place in the school environment, however, were not part of the formal curriculum. These case studies provide us with an in-depth understanding on how fun influences learning and lead us to a model that describes the relationship between the experienced fun while learning, children's attitude about the topic, and their learning outcomes. Thereafter, this model is tested in further studies in relation to coding and game-based learning, investigating whether personal (i.e., children's self-regulated learning skills) and environmental (i.e., socioeconomic status) factors influence the effect of fun on learning, and whether fun and learning can be linked to physiological response data. Consequently, the model is developed further, reaching its final state that incorporates all the knowledge and understanding gathered during this work, describing the role of fun in learning, including various factors that nuance the previously discovered relationships.

The work presented in this thesis contributes to the intersection of the child-computer interaction and learning sciences fields, particularly to the subfield focusing on STEM education. The five main contributions of this thesis are described as follows: (1) It provides a theoretically grounded and empirically validated definition for fun, which can serve as a common theoretical framework for researchers in these fields. (2) It provides a much needed, validated measurement tool for the assessment of fun, which fills a gap in the current palette of tools for empirical assessment of learning and other activities. (3) It introduces the fun in Learning (FiL) model that describes the relationship between fun, attitude, and learning, extending our understanding of the role of fun in learning, supported by quantitative research methods. (4) It nuances the effect of fun in learning by considering personal and environmental factors (i.e., self-regulation and socioeconomic status) providing educators cues on tailoring their activities to different audiences. And finally, (5) it links fun to physiological markers derived from physiological response data, allowing for a momentary investigation of the effect of fun in learning. The introduced FiL model can serve as a starting point and an inspiration for researchers and educators for orchestrating meaningful fun learning activities for children that not only serve as entertainment but are also beneficial for achieving the learning goals.

Contents

List of Figures

List of Tables

Part I.

INTRODUCTION

1 Introduction and Motivation

1.1 Introduction

Imagine you are a 12 year old student. Today at school you are going to get a special class. Someone is going to visit and show you how to do programming. You are very enthusiastic. You always wanted to know how to program a robot, so you are looking forward to the class. However, your best friend, Ben, is not that much enthusiastic. He is more into music, drawings and paintings. He is not fond of science and robots. When the time comes, people come into your class to give the programming workshop. You are having the best day of your life. Finally, you learn how to program a robot! You learn how to make it to do what you want, to navigate it where you want. While your friend, Ben, is slightly disappointed. He was not really interested, he finds it too difficult, and he does not find it as much fun as you. Even when you try to help him, he is still a bit unhappy. At the end of the class, the two of you are having different feelings. You are very excited, have a positive attitude about programming and you are really looking forward to doing it again because you thought it was a lots of fun. While your friend, Ben, sees it differently. He thinks it was okay, but it was also difficult, and he does not really want to do it again as he did not find it as much fun as you.

This vignette invites the reader to join a story that illustrates how the same learning activity can be a very different experience for children, how fun can be a natural way to describe whether the experience was positive or not, and how it can also be determinant for some of the outcomes of the learning activity, which differ widely based on individual factors. Nowadays, fun has importance in several fields of life. While it is a defining element in leisure and play, it is increasingly understood how fun is an essential component of learning, interaction design and game play. Interaction design researchers have often considered fun or enjoyment as a key success criterion for their designs, not only while designing games but also for other interactive technologies especially when targeting children users [176]. Since one of the primary aims of educators is to get and keep learners motivated, and given that research has demonstrated that the promise of fun has an inviting effect [169] and that fun increases engagement with learning technologies and with learning activities [127, 169, 232, 317], substantial research effort has been invested within the field of learning sciences in making learning enjoyable. This interest is especially stressed in relation to STEM (Science, Technology, Engineering and Mathematics) learning as there is a worldwide pursuit to increase children's interest in scientific topics, and especially in computer science, as computational thinking and programming are frequently seen as the literacy skills of the 21st century [210].

Despite such wide interest, there has been little systematic effort to provide quantifiable evidence regarding the role of fun in learning, not least due to a lack of a common theoretical framework for defining the concept of fun and for supporting its measurement.

Therefore, answering the simple questions such as *whether you and Ben have learnt evenly from the same playful programming class, and whether the experienced fun has to do anything with it,* is not straightforward. Moreover, we not only have limited scientific evidence for the effect of fun on the learning outcomes, but we know very little about how further personal and environmental factors nuance the aforementioned relationship. Having a thorough understanding on the role of fun in learning, within the specific context of STEM education, is crucial for designing and implementing activities that reach their goal of increasing children's interest and involvement in science related activities on the long-term.

This thesis contributes to this goal by first defining fun and creating a reliable measurement tool for its assessment in the learning context. Then, with the use of this measurement tool we assess fun in various learning activities, to quantify its impact on different learning outcomes and in different contexts, considering personal and environmental factors as well. Therefore, this thesis contributes with a model on the role of fun in learning, which can serve as a starting point and an inspiration for researchers and educators for orchestrating meaningful fun learning activities for children that not only serve as entertainment, but also beneficial for achieving the learning goals.

1.2 Background

In the following sections, we introduce and briefly discuss the related topics and the theories covered throughout the chapters of the thesis to position the scope of this work.

1.2.1 Formal-, Non-formal-, and Informal Learning

While one would typically consider school as a setting where learning takes place, the academic literature generally distinguishes three learning contexts: formal, non-formal, and informal. For our conceptualization, we draw upon the definitions of The Council of Europe[1], and those provided by Eshach [83]. According to them, *formal learning* happens usually at a formal learning space (e.g., school), it is structured and follows a syllabus, and the learning goals are predetermined by the teacher. Accordingly, *formal learning* is led by a teacher or educator, who evaluates the learning outcomes, and students' motivation is typically extrinsic as their participation is compulsory. *Non- formal learning*, on the other hand, happens often outside of the formal learning context, it is structured and might follow a syllabus, but it is typically not part of the school curriculum, and the learning goals usually arise from the learners' conscious decision. The activity is typically guided or teacher-led, the learning outcome is usually not evaluated, and students' motivation is generally more intrinsic as their participation is mostly voluntary. Lastly, *informal learning* can take place everywhere, it is unstructured and does not follow a syllabus, there are no learning goals defined, the learning activity is learner-led, and thus the learning is not

[1] https://www.coe.int/en/web/lang-migrants/formal-non-formal-and-informal-learning

evaluated, and motivation for participation is mainly intrinsic and therefore voluntary. The differences between formal-, non-formal-, and informal learning are summarized in Table 1.1. The non-formal and informal (in other words, out-of-school) settings have special importance in relation to STEM (Science, Technology, Engineering, and Mathematics) education, as they often are aimed at increasing children's interest in scientific topics [303].

Table 1.1 Differences between formal-, non-formal-, and informal learning.

Formal learning	Non-formal learning	Informal learning
At formal learning space	Often outside of formal context	Everywhere
Structured	Structured	Unstructured
Follows a syllabus	Might follow a syllabus	Does not follow a syllabus
Part of the school curriculum	Usually not part of the school curriculum	Not part of the school curriculum
Learning goals are predetermined by the teacher	Learning goals are often determined by the learners	No learning goals are determined
Teacher-led	Guided or teacher-led	Learner-led
Learning is assessed	Learning is usually not assessed	Learning is not assessed
Participation is compulsory	Participation is mainly voluntary	Participation is voluntary
Motivation is extrinsic	Motivation is often intrinsic	Motivation is intrinsic

1.2.2 Learning Approaches

During this dissertation research, three learning approaches had special importance: Constructivist Learning, Lifelong Kindergarten, and Design-Based Learning.

The Constructivist Learning Theory originates from the work of Piaget [219], and is based on the idea that children are active participants in their learning journey, and knowledge is constructed from experiences. Papert [211, 212] developed further this idea, and turned it into a learning theory called Constructivism. The core element of Constructivism is supporting children becoming authors and creators of educational content, rather than passive recipients. This approach is frequently adopted in child-computer interaction research.

A follow-up on Papert's work is the Creative Learning Model of the Lifelong Kindergarten research group at MIT [238]. The Lifelong Kindergarten approach examines how constructionist learning can be implemented into technologies and educational practices. It is described as being "ideally suited to the needs of the 21st century, helping learners to develop the creative thinking skills that are critical to success and satisfaction in today's [digital] society" ([237], p. 1). This approach emulates a traditional kindergarten environment where, during play, children design, create, experiment, and explore continuously. In this approach, learning takes place through a spiraling process that starts with imagining, and followed by creating, playing, sharing, reflecting, before returning to imagining, and so on.

The Design-Based Learning approach is related to the aforementioned in the sense that it requires children to create their own solution for a design challenge or a real-life problem, based on their prior knowledge. In general, Design-Based Learning involves learning from trial and error, open exploration, teamwork, reflection, and supportive tools [343]. All of the aforementioned theories are frequently used in relation to STEM education.

1.2.3 STEM Learning, and Learning to Program or Code

In the recent decades the importance of STEM teaching and learning has become apparent due to the technological advancements of the 21st century. Accordingly, related educational activities are gaining momentum both within formal and informal contexts. Increasing children's interest in scientific topics from early ages on has thus become a worldwide pursuit, with specific focus on computer science, as computational thinking and programming are often regarded as main literacy skills of the 21st century [210].

There are two main, non-traditional approaches for making (STEM) learning enjoyable: designing playful learning activities in out-of-school learning spaces and gamification. Out-of-school STEM learning can take place, for example, at maker spaces[2], Fab Labs[3], coding clubs, and science museums [220, 255]. These venues typically provide children with a collaborative (work)space that enables exploring, learning, creating, and sharing. In case of maker spaces and Fab Labs, the emphasis is on making. Such settings offer a wide range of readily available tools from high-tech to no-tech. In coding clubs, the focus is on coding and robotics, whilst in science museums, a variety of scientific topics may be addressed, including making and coding. Non-curricular coding clubs can play a significant role in teaching children to program as programming is not yet an integral part of primary school curricula. The UK, Estonia, Spain, and Finland are examples from Europe where programming is already a compulsory subject in primary education. In other countries, such as the Netherlands, primary schools can decide on their own whether to teach programming to their students. Elsewhere, the opportunity to learn to program is only available through participation in out-of-school activities such as coding clubs. The overarching approach for out-of-school learning is to develop learning environments that support learners' intrinsic motivation and trigger their curiosity. Despite this worldwide inquiry, our knowledge is limited on what factors influence children's interest and learning outcomes in STEM subjects in general, and in coding or programming in specific.

1.2.4 Gamification and Digital Game-Based Learning

The other main, non-traditional approach to make learning enjoyable is gamification. Deterding defined gamification (a.k.a. gameful learning, gamified learning, edutainment, digital game-based learning (DGBL)) as *"the use of game design elements in non-game contexts"* ([68], p. 9), and Prensky described it as acquiring knowledge and skills through playing engaging computer games [229]. In general, we can say that gamification is

---

[2] http://www.makerspaceforeducation.com/makerspace.html
[3] https://fabfoundation.org/getting-started/#fablabs-full

founded upon a commonly accepted belief that games can make learning fun. However, as Iten and Petko noted, "*it is less clear what fun in serious games actually means and how is it related to cognitive, emotional and behavioural engagement*" ([127], p. 154). On top of that, despite being often assumed that digital game-based learning (DGBL) is fun, Yee [333] argued that educational games require players to do many tasks, making students feel tired and tedious. In other cases, educational games were found to take too long and were no longer fun to play after the novelty effect was gone [133]. Acknowledging that gamification often fails in the context of education, van Roy and Zaman [251] looked into the underlying issues, and provided nine gamification heuristics based on Self-Determination Theory [67] for a successful implementation of game-design elements into educational games. Furthermore, they also emphasized the need for using of background theories when designing gamification [250]. A possible background theory when designing gamification is the concept of eudaimonia [324]. The term stands for realizing one's potential, and is related to, for example, self-determination, and the balance of skills and the level of challenges one faces with. The concept of eudaimonia is often contrasted with happiness that is characterized as a hedonistic pleasure, in comparison with eudaimonia, which is related not only to positive affect, but long-term importance, need fulfillment, and feelings of meaningfulness [281]. Accordingly, both for eudaimonia and for fun challenge and need fulfillment is a key concept, but eudaimonia is different from fun as fun happens in the here and now, and it does not have to be meaningful.

Despite the inconclusive results of the effectiveness of gamification in education, in the last few decades, there has been substantial research effort invested worldwide to make learning enjoyable by using digital games for learning. Systemic reviews on gamified learning, (a.k.a. gameful learning, gamification, edutainment, digital game-based learning), and serious games, conclude that gamification allows teaching systems to improve student engagement and motivation, and lead to increased performance [54, 228, 284]. In a recent literature review Boyle et al. [40] found that science, technology, engineering, and mathematics (STEM) are the most popular subject disciplines where digital game-based learning is applied. They also found that the main aim of such games was knowledge acquisition, but other goals also appeared such as aiming for behaviour change or perceptual, affective, cognitive and physiological outcomes.

When discussing gamification, game experience - while not only pertained to educational games - is an organic part to be covered. Poels, de Kort, and IJsselsteijn defined game experience as *"participants' subjective experiences associated with digital gameplay"* ([225], p. 4). In the past two decades, there has been a considerable amount of effort invested to tackle game experience, which resulted in a number of measurement tools, mostly targeting the adult population. The most well-known ones are the Game Experience Questionnaire (GEQ [225]), the Player Experience of Need Satisfaction (PENS [259]), the User Engagement Scale (UES [204]), and the EGameFlow [91], which scales, among others, we discuss in detail in Chapter 3.

1.2.5 Play

Another related, and often confused concept to game is play. Huizinga [121] elaborated in great detail on the concept of play and concluded that play has five main aspects: it is voluntary and free, rule ordered, happens within fixed boundaries (i.e., locality and duration), it is different from ordinary life and it has no material interest. Vygotsky [321] described that play is children's voluntary activity which involves an imaginary situation and is rule-based. More recently, Sutton-Smith defined play as *"an activity that is voluntary, intrinsically motivated, fun, incorporates free will/choices, offers escape, and is fundamentally exciting"* ([287], p. 3). This latter definition also links the notion of fun to play. Despite during early childhood fun and play often concur, as children approach adolescence the hedonistic character of fun that is purely present during play gives way to challenge [72], which is a well-established dimension of fun among adults [91, 126, 259]. Therefore, we can state that play is an essential part of childhood and children's mental development, but the notion of play can be distinguished from both the notion of game and the notion of fun.

1.2.6 Affective Processes in Learning

According to Hascher *"there is rarely any learning process without emotions. (...) Despite the obvious connection between learning and emotion, still very little is known about it."* ([112], p. 13). From cognitive psychology we know that emotions influence cognitive processes and strategies, decision making and motivation, and that the aforementioned influences are reciprocal [147]. For example, emotions influence our memory (think about how differently eye-witnesses report on catastrophes), but our memory also influences our emotional reactions (think about how a positive memory can influence one's mood) [147]. In psychology research, the basic emotions that are characterized by prototypical facial expressions are happiness, sadness, surprise, disgust, anger and fear [75]. However, it has been questioned whether these six emotions play a key role in the learning process [60]. Graesser proposed that *"the ensemble of emotions that occur during learning* [i.e., boredom, confusion, frustration, curiosity, enjoyment and anxiety] *are very different from the basic emotions* [i.e., happiness, sadness, surprise, disgust, anger and fear] *that dominated psychological research for decades (...)* [as] *most of the basic emotions are not prevalent in learners and predictive of learning in contemporary learning environments"* and that "*the profile of emotions that learners experience have some commonalities but also predictable differences over task, goals, subject matter content, and population of learners"* ([101], p. 2). As no scientific consensus exists on the key emotions in learning, the different effects of positive compared to negative emotions while learning and on learning are not straightforward. Valiente et al. said that *"although researchers typically expect positive emotions to foster academic success, high-arousal positive emotions (...) may detract from achievement"* ([313], p. 130). Hence, any assumptions of a straightforward relationship between positive emotions and learning, and negative emotions and not learning would be misplaced. For example, the observational analysis of Craig et al. (2004) found significant positive correlation between confusion and measured learning gain. Additionally, in the Control-Value Theory (CVT) of achievement emotions [214] Pekrun mapped academic emotions into a two-dimensional plot based on their valence and activation, and thus he distinguished positive activating (e.g., enjoyment, curiosity), negative activating (e.g., frustration, confusion), positive deactivating (e.g., relief, relaxation), and negative deactivating (e.g., boredom) emotions. In line with this, Loderer, Pekrun, and Lester [168] concluded based on their systematic literature review that *"negative emotions like confusion, but potentially also anger or boredom, can be beneficial to learning under certain circumstances, likely only to the degree that they promote deeper engagement with contents and can be successfully resolved".*

Despite the wide variety of emotions that occur during learning, Pekrun et al. [215] argued that scientific research on academic emotions had a strong and narrow focus on test anxiety for decades. Concurring to this, Loderer et al. [168] found that research into emotions related to learning almost quadrupled in the past 20 years, but among the reviewed papers anxiety was still the most studied academic emotion, while enjoyment has become the second most frequently investigated one. Similarly, a systematic literature review of emotions in design-based learning [342] classified emotions reported in empirical studies according to the typology of emotions introduced with the Control-Value Theory. With very few exceptions, the studies reviewed sought for indications of enjoyment as a positive aspect of the learning activity, though the evidence on the expected positive impact of enjoyment or fun on learning engagement with the topic was found to be equivocal. It is noticeable that fun and enjoyment are terms often used interchangeably in design research, with fun being regularly adopted as an evaluation criterion for learning games (e.g., [234, 275]). In this work, however, we pertain exclusively to the examination of fun (which we conceptualize in Chapter 2 including its relation to related concepts), and this way we broaden our understanding on the differences and similarities between fun and other emotions.

1.2.7 Fun and Learning

The notion of fun has been gaining momentum in the past decades, especially in the context of educational technology in relation to gamified learning and child-computer interaction, and in the context of non-formal STEM education. Despite the growing interest towards the fun experience - with special regard to its relation to learning – the concept behind the term and its measurement are not clearly described, and accordingly, currently there is a lack of commonly accepted theoretical framework for defining the concept of fun and for supporting its measurement.

Bisson and Luckner [32] were among the first to discuss the positive effects of fun in the learning environment. In their view fun functions as a vehicle for evoking intrinsic motivation, reducing stress and social boundaries, and creating a safe learning environment. Papert coined the notion of *hard fun* [109] in relation to constructivist learning (which today we would call design-based learning) to describe fun in terms of learning challenges (i.e., picking own challenge and getting fun from the autonomy and competence that come from it, thus experiencing fun *because* something is difficult, and not *despite* that). Other authors argued that fun facilitates engagement [232], enhances learning [51, 171, 232, 295, 317, 327], improves programming skills [169], has a significant effect on the learning effort [169], fosters curiosity [127, 317], contributes to high-quality learning experience [295], promotes collaborative learning [51], is a predictor for learning success [127], and has an effect on gaining motivation [127]. However, it is important to note, that in the academic literature, and hence in the aforementioned studies, the terms *fun* and *enjoyment* are often used interchangeably, and in the recent years it has been usual that writings discussing fun simply do not elaborate on what their understanding of the concept of fun is [127, 150, 307]. Handling fun as a common-sense term does not contribute to a commonly accepted theoretical framework for the definition, and without that, previous research results are difficult to compare, and explaining the role of fun in learning becomes complicated.

Therefore, to be able to investigate the role of fun in learning, this research starts with the conceptualization of fun (Chapter 2), and with the design and validation of the related measurement tool called FunQ (Chapter 3).

1.2.8 Attitude and Learning

According to the American Psychological Association's dictionary[4], attitude refers to a *relatively enduring and general evaluation of an object, person, group, issue, or concept on a dimension ranging from negative to positive.* In this thesis we discuss attitude within the educational context, and hence we refer to children's attitude towards a learning activity or subject. Researchers from the 1950s have already identified links between students' attitude and their academic achievement [84]. However, recent studies are not aligned regarding the effect of attitude on learning. While some researchers presented supportive evidence towards attitude having a positive effect on students' academic achievements (i.e., learning outcomes) [19, 31, 199], others have found no direct relationship [82, 129, 130, 162, 335]. Within the specific field of STEM education, recent studies have shown that a positive correlation existed between students' attitude and their learning outcomes [104, 199, 260], and further studies have indicated that students' attitude towards programming not only influence their academic achievement [50, 300], but also affect their career choices [50]. Despite the growing body of knowledge, our understanding is still limited on how exactly students' attitude is related to their learning outcomes, what is the direction of this relationship, and what underlying factors might have an influence on it. In this thesis, among the possible influential factors we investigate the role of fun experienced while

[4] https://dictionary.apa.org/attitude

learning as a situational effect, students' self-regulation as a personal factor, and students' socioeconomic background, as an environmental impact.

1.2.9 Self-Regulation and Learning

Zimmerman defined self-regulation as *"self-generated thoughts, feelings and actions that are planned and cyclically adapted to the attainment of personal [learning] goals"* ([346], p. 14). According to Pintrich [221], self-regulated learning is an active and constructive process, during which students define the learning goals for themselves, and parallel, they also regulate, monitor, and control their cognitive and motivational processes to attain their self-set goals. From this definition it is apparent that attitude and intrinsic motivation are inherent to self-regulation, and thus they play a crucial role on the relationship between self-regulation and learning outcomes. According to Pintrich [221], highly motivated and strongly self-regulated students become the academically most successful ones, and Flavell [86] explained this by means of cognitive and metacognitive strategy use. He suggested that more motivated students utilize a wider range of strategies, leading to more efficient learning.

Regarding the relationship between emotions and self-regulated learning, recent studies [15, 215] found that academic emotions (e.g., enjoyment, hope, pride, anger, anxiety, boredom) in general, and enjoyment in specific, are significantly correlated with self-regulated learning behaviors. However, these studies did not investigate the direction of the relationship, nor did they investigate learning outcomes. Results regarding the relationship between emotions, self-regulation and learning outcomes are limited, and appear to be inconclusive, as some studies found a positive association between those (e.g., [8]), while others found no association at all (e.g., [318]). According to Villavicencio and Bernardo, *"there has been very little evidence about the moderating effect of academic emotions on the relationship between cognitive–motivational variables on the one hand and learning and achievement on the other"* ([318], p. 337). In this thesis we considered self-regulation as an element that possibly moderates the relationship between fun and learning, this way broadening our understanding on how self-regulation contributes to students' learning outcomes. Our related findings are described in detail in chapters 8 and 10.

1.2.10 Socioeconomic Status and Learning

To define our understanding of socioeconomic status (SES) we adopt the definition of Sirin, who considers SES to be *"an individual's or a family's ranking on a hierarchy according to access to or control over some combination of valued commodities such as wealth, power, and social status"* ([276], p. 418).

Investigating the relationship between students' socioeconomic background and their academic achievement has a long tradition. The meta-analysis of White [326], already back in the 80s, concluded that the way SES is defined and measured (i.e., unit of the analysis) influences the strength of the relationship between SES and academic achievement. A more recent meta-analysis in the US context [276] investigating the relationship between SES and academic achievement in general found further supportive evidence that there is an overall positive correlation between those two concepts, but the correlation between SES and academic achievement has become weaker over time, recent studies indicating a weaker relationship then before. Most recent meta-analyses taking into account different contexts and different age-groups found a positive correlation between SES and academic achievement, however, the strength of the correlation appeared to be weak to moderate [165, 166, 244]. These findings suggest that despite worldwide efforts to increase educational opportunities, the applied measures do not seem to reduce inequalities in students' academic outcomes between low- and high-SES students. Despite as just shown the relationship between SES and academic performance in general is well studied, we know much less about the relationship between SES and STEM education in specific. The few existing studies suggest that there is an even wider gap between low-SES and high-SES students' academic achievement when it comes to STEM subjects [35, 201], however, our knowledge is very limited about the underlying reasons. In this thesis, SES has been considered as a factor that potentially influences how students learn to program, and in relation to the fun experienced while learning. The studies associated are described in detail in Chapter 9 and 10.

A related concept to SES is the concept of science capital [13], which encapsulates *"all science related knowledge, attitudes, experiences and social contacts that an individual may have"* ([99], p. 5). The concept of science capital has gained a lot of traction in the past two decades in STEM education research, practice, and policy [202] as research indicates that the higher one's science capital the more likely one is to engage with science-related activities and to have a 'science identity', the latter indicating an increased likeliness to continue with science related studies after age 16 [13]. Science capital is found to be related to cultural capital, gender, ethnicity and set track in science, in other words, to one's socioeconomical background [13].

1.3 Scope of the Thesis

The main purpose of this thesis is to investigate the role of fun in learning, and to provide quantifiable supportive evidence for it. To this end, we explore with quantitative research methods, quasi experiments, and case studies in various learning settings (i.e., formal-, non-formal-, game-based-, online-, and offline learning) in Western-Europe, with children between age 8 and 16 how fun influences their learning outcomes within the field of STEM education. The evaluation of fun in learning with young children (below age 8) and with adults, outside of STEM subjects, and in non-Western-European culture is beyond the scope of this research.

1.4 Research Questions

The driving research question of this thesis was *(RQ1) What is the role of fun in learning?* In order to be able to investigate this question, first we had to define fun and develop a reliable measurement tool for the assessment of it. Therefore, we asked the following sub-questions as well: *(SQ1) What is fun?*, and *(SQ2) How can we measure fun?* Finally, we were also interested whether personal (i.e., self-regulation) and environmental (i.e., socioeconomic status) factors have an effect on the relationship between fun and learning. Hence, the final research question of this thesis was formulated as follows: *(RQ2) Is the relationship between fun and learning affected by personal factors such as self-regulation, and environmental factors such as socioeconomic status?* A number of studies introduced in this thesis did not contribute directly to responding these research questions, nevertheless, they had an important role in the better understanding of the role of fun in learning, resulting in the final answer for the research questions. Therefore, Chapter 4 and 5 describe case studies that led to the original model on the role of fun in Learning (Chapter 6); and Chapter 9 investigates the relationship between fun and learning in reflection of physiological data with the aim of triangulating research results, this way deepening our understanding on the topic. The chapters' relation to the research questions are summarized in Table 1.2 below.

Table 1.2 Summary of research questions.

	Research Questions	Chapter(s)
RQ1	What is the role of fun in learning?	4, 5, 6, 9, 10
SQ1	What is fun?	2
SQ2	How can we measure fun?	3
RQ2	Is the relationship between fun and learning affected by personal factors such as self-regulation, and environmental factors such as socioeconomic background?	7, 8, 10

Besides the research questions, gender differences were investigated throughout the thesis for multiple reasons. The main reason was that there is a large and persistent gender gap in STEM engagement [178], involving learning, related activities and careers, with girls and women being generally underrepresented. This gap is especially striking in relation to computer science, which has received great attention in the last decade (e.g., [178, 197, 209, 337]), however, the underlying reasons are still not well understood. By investigating the gender differences in our studies we aimed to contribute with possible explanations. In relation to this, reporting gender differences, in general, can inform future research and literature reviews, contributing ultimately to a better understanding of girls' STEM engagement. And last, when designing the studies, we wanted them to be equally attractive for both boys and girls, which endeavor was to be checked. However, the question of gender differences remained a side aspect of this research, and therefore, it did not become a research question.

1.5 Methodology

Throughout this thesis to address the research questions, we utilized a number of quantitative research methods. The main pillars of the thesis methodology are discussed briefly below.

1.5.1 Deductive Scale Development

First and foremost, to answer the research questions *(SQ1) What is fun?* and *(SQ2) How can we measure fun?* we applied a deductive scale developmental approach [3]. This approach allowed us to develop a theoretically grounded and empirically validated definition for fun (Chapter 2), and to develop the related measurement tool, FunQ (Chapter 3). It had four main phases: i) theory driven item construction, ii) test of initial item pool, iii) user study, iv) validation of the final item pool.

1.5.2 Case Studies

We conducted multiple case studies [3] to understand in depth the relationship between fun and learning across various study settings, such as playful coding workshops (Chapter 4 and 8) or Digital Game-Based Learning (DGBL; Chapters 5 and 7) classes. These case studies provided the necessary insights to formulate the core of this work that later translated into a model that describes *the role of fun in learning (RQ1).* Additionally, the case studies helped us to answer *(RQ2): Is the relationship between fun and learning affected by personal factors such as self-regulation, and environmental factors such as socioeconomic status?*

1.5.3 Structural Equation Modeling (SEM) and Path Analysis

To put together the pieces of information gained from the in-depth case studies, and to test whether the pieces of the puzzle fit correctly, we utilized Structural Equation Modeling (SEM) and Path Analysis [149]. These methods allowed us the simultaneous investigation of multiple relationships, leading to the development and validation of the desired quantifiable evidence for the role of fun in learning in a form of models, and hence, to answer *(RQ1): What is the role of fun in learning?* (Chapters 6 and 10).

1.5.4 Multimodal Learning Analytics

The upcoming approach of multimodal learning analytics (MMLA) combines several sources of data to serve as a virtual observer and analyst of learning activities [33, 272]. MMLA provides an unprecedented opportunity to understand students' behavior and performance during and after the learning sessions by understanding their relations with cognitive processes and affective mechanisms [62]. Therefore, MMLA can shed light to learning processes that may be invisible to the human eye and that students cannot self-report on [63, 161, 210, 269]. Accordingly, MMLA can complement our understanding on how children learn, providing more information on children's affective aspects during learning activities. Utilizing these favorable properties of MMLA, we applied this approach to widen our understanding on the relationship between fun and learning in the specific context of learning to program by combining data from physiological measures and from self-reports (Chapter 9). Thereby, this chapter contributes to answering *(RQ1): What is the role of fun in learning?*.

1.6 Thesis Outline

This thesis consists of eleven chapters which are arranged in seven parts. Firstly, The Conceptualization and Measurement of fun (Part II) focuses on laying the theoretical groundwork for defining and measuring fun and includes studies that aim to define fun and to create a reliable measurement tool for it. Secondly, The Role of fun in Learning (Part III) delves into case studies in various settings for a better understanding of the role of fun in learning by the use of FunQ. In this section we also introduce the initial FiL model that describes how fun influences learning considering children's attitude about the topic. Thirdly, the Personal and Environmental Influential Factors in Learning (Part IV) goes deeper into the understanding of the relationship between fun and learning by investigating possible factors that can nuance the aforementioned relationship. Fourthly, the fun and Learning in Reflection of Physiological Data (Part V) investigates whether fun and learning can be linked to certain physiological markers. An Extended Model (Part VI) introduces a final study in which all aforementioned factors are investigated simultaneously aiming to create an extended model on the role of fun in learning, taking into account personal (i.e., self-regulation) and environmental (i.e., socioeconomic status) factors as well. Finally, the Conclusion (Part VII) provides an overview of the research contribution of this thesis and lays the groundwork for future research. The thesis outline is visualized on Figure 1.1.

1.6.1 Part II: The Conceptualization and Measurement of Fun

Chapter 2 – Conceptualizing of fun This chapter presents the theoretical groundwork for defining fun, thereby it addresses SQ1 directly. It also investigates the notion of fun in earlier research, involving related definitions where they exist, and neighboring concepts and the distinction between those and fun. The chapter concludes with our multi-dimensional definition for fun, which definition was created based on theoretical grounds, and was validated with empirical studies. This chapter contributes with a) a review of literature regarding the concept of fun, and b) a conception of fun as a multi-dimensional theoretically motivated concept.

Chapter 3 – Measurement of fun – Introducing FunQ Chapter 3 describes three consecutive studies that aimed to a) validate the theoretically grounded definition for fun, and b) create and validate a related measurement tool – hence, it directly addresses SQ2. In accordance, this chapter introduces FunQ, a theoretically grounded and empirically validated questionnaire, designed for adolescents for the assessment of fun. FunQ consists of 18 items, and measures the experienced fun across six dimensions (Autonomy, Challenge, Delight, Immersion, Loss of Social Barriers and Stress), and bears with the appropriate validity and reliability measures. This chapter contributes with a) a multi-dimensional instrument for assessing experienced fun - the FunQ, b) a psychometric evaluation of the proposed instrument, and c) a list of suggestions regarding best practices when using self-reports with child respondents.

1.6.2 Part III: The Role of Fun in Learning

Chapter 4 – Fun in Coding – A Case Study In Chapter 4 we report on a case study that investigates children's topic-related attitudes, their state-level emotions, the fun they experienced, and the initial- and final knowledge on the subject in relation to a playful coding workshop. This study sheds light on relationships that later became the foundation for the models describing the role of fun in learning, therefore it paves the way to answering RQ1.

Chapter 5 – Fun in DGBL – A Case Study Chapter 5 presents a following case study, in which we examine fun in the setting of Digital Game-Based Learning (DGBL). Accordingly, this study focuses on how the perceived fun while playing with an educational game has an impact on children's measured- and perceived learning, motivation, attitude, self-efficacy and intention to play similar games. This chapter contributes with a deeper understanding on the effect of fun in the DGBL environment, and hence, indirectly adds to the answer of RQ1.

Chapter 6 – The Fun in Learning (FiL) Model This chapter builds on the previously introduced case studies and presents the FiL model that describes the role of fun in learning within the context of programming. The FiL model states that children's learning is significantly influenced by the experienced fun while learning to code across its positive influence on children's attitude towards coding. Therefore, this chapter contributes with quantifiable evidence regarding the role of fun in learning (RQ1), and hence provides supportive evidence for the efforts of educational researchers and practitioners who try to make learning activities more fun for students.

1.6.3 Part IV: Personal and Environmental Influential Factors in Learning

Chapter 7 – Self-Regulation and Fun in Learning – A Case Study Chapter 7 describes a case study in which we examine how students' self-regulation influences their learning in the context of a self-regulated training game. The chapter contributes with further evidence for supporting efforts making learning more enjoyable; however, it challenges the long-standing general belief on the positive relationship between self-regulation and learning. The chapter also adds to the answer for RQ2.

Chapter 8 – SES and Fun in Learning In this chapter we discuss a study that aims to understand whether students' socioeconomic status has an impact on their learning outcomes, topic-related attitudes and perceived fun while learning to code. The theoretical perspective of science capital suggests that children in high income families will hold more positive attitudes towards science and technology and will perform better in coding than children from lower income areas based on a generally higher exposure to computing technology. The chapter contributes with practical implications, as the study finds that children from the middle- and low-income school profited the most from the playful coding workshop. The chapter also contributes to answering RQ2.

1.6.4 Part V: Fun and Learning in Reflection of Physiological Data

Chapter 9 – Fun in Learning in Reflection of Physiological Data Chapter 9 introduces a multimodal data analysis study, in which we examine how fun impacts learning during a coding activity, combining continuous physiological response data from wristbands and facial expressions from facial camera videos, along with self-reported measures for learning and for the experienced fun. By using multimodal data, this study goes a step further than earlier research, which either pertained to surveys or to physiological response data only. Using the combination of the two allowed us a deeper understanding on how fun occurs during learning to program, and which physio-affective states can be used as a predictor of fun. This chapter thus contributes with a deeper understanding on the relationship between fun and learning by linking fun and learning to certain physiological markers derived from the physiological response data. Accordingly, the chapter also adds to the answer for RQ1.

1.6.5 Part VI: An Extended Model

Chapter 10 – Fun, Self-Regulation, and SES in Learning This chapter aims to compile all knowledge gathered during the course of this work by simultaneously investigating the role of fun in learning, taking into account personal (i.e., self-regulation) and environmental (i.e., SES) factors as well. Accordingly, the chapter contributes with an extended model that describes the role of fun in learning, and hence, it directly addresses RQ1 and RQ2.

1.6.6 Part VII: Conclusion

Chapter 11 – Discussion and Conclusion In the final chapter, we answer the research questions and provide a detailed summary about the research contributions of this thesis. Further, we point out the limitations of this dissertation research and discuss potential areas of future work.

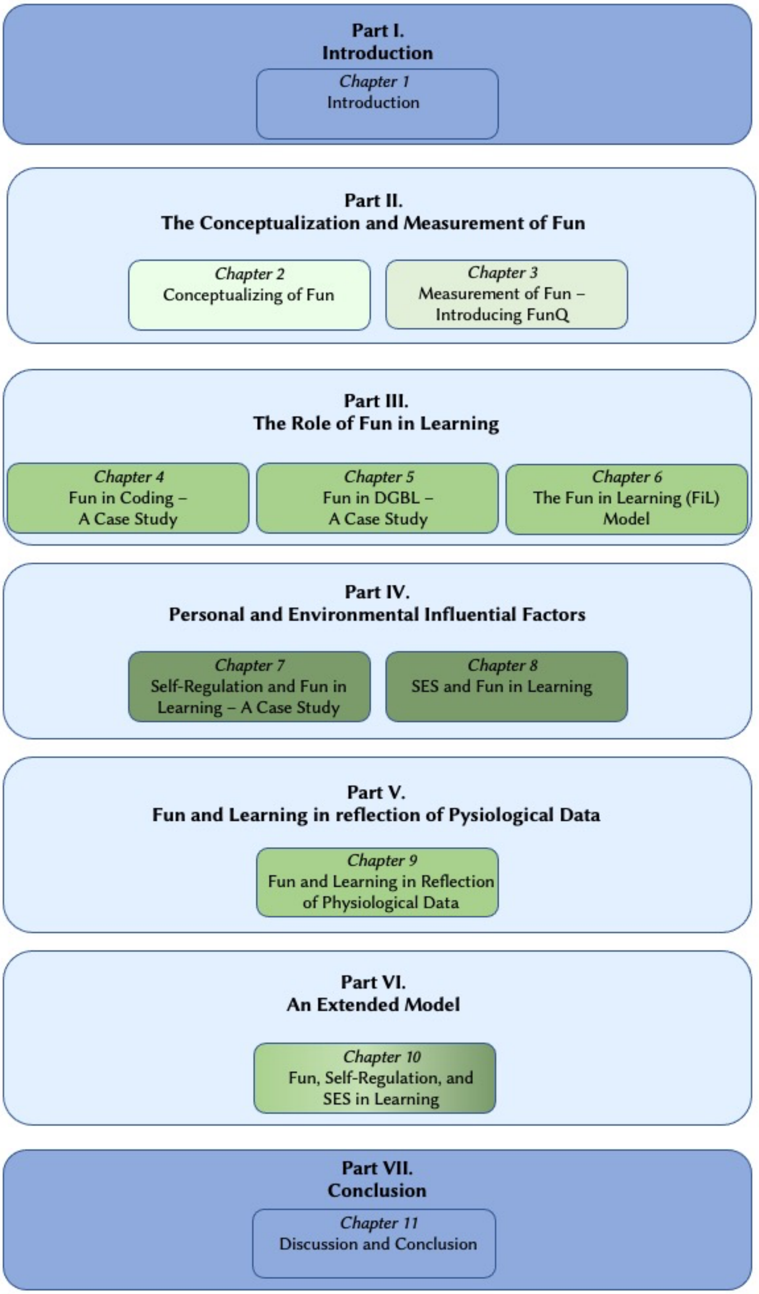


Figure 1.1 Thesis outline. Color coding of the related research question: RQ1, SQ1, SQ2, RQ2.

Part II.

THE CONCEPTUALIZATION AND MEASUREMENT OF FUN

2 Conceptualizing Fun[5]

Summary

Researchers and practitioners in learning sciences, educational technology and child-computer interaction often argue that fun is an essential element of learning. Therefore, researchers in the above fields aim to explore how learning activities could be made more enjoyable in order to facilitate engagement in the learning process, and to improve learning outcomes. Despite such wide interest, there has been little systematic effort to define and measure fun. This chapter aims to conceptualize the notion of fun based on earlier research, and hence, it contributes with a) a review of literature regarding the concept of fun, and b) a conception of fun as a multi-dimensional theoretically motivated concept.

2.1 Introduction

Fun is important in several aspects of life. Next to being a defining element of leisure and play, it is increasingly understood to play a key role in learning, work, and social interactions. Researchers and developers in the fields of educational technology and child-computer interaction often inject fun elements in their systems just like educationalists who usually aspire to make learning activities enjoyable. Research has demonstrated that the promise of fun has an inviting effect [169], that fun increases engagement with learning technologies and with learning activities [127, 169, 232, 317] and having fun has a positive effect on learning outcomes[51, 78, 169, 171, 232]. Additionally, neuroscience provides evidence at the biochemical level for the positive effects of fun in the learning environment and on learning [327].

Despite that fun is often mentioned, the concept behind the term and its measurement are not always clearly described. The Cambridge dictionary[6] defines fun as being an informal expression for *pleasure, enjoyment, or entertainment*, and as *behaviour or activities that are not serious; games or jokes*. In the academic literature the terms fun and enjoyment are frequently used interchangeably [78, 89, 127, 169, 186, 246]. In other cases, these notions stand next to each other as close relatives, however, complementary to some degree [72]. This also invites a scientific debate concerning the distinction between the two concepts [72]. Nonetheless, in the recent years it is usual that papers discussing fun simply do not elaborate on what their understanding of the concept of fun is [127, 150, 307]. In our conceptualisation we understand enjoyment as a term that describes a positive emotion,

[5] This chapter is based on the following publication: Tisza, G., & Markopoulos, P. (2021). FunQ: Assessing the fun experience of a learning activity with adolescents. Current Psychology. https://doi.org/10.1007/s12144-021-01484-2

[6] https://dictionary.cambridge.org/dictionary/english/fun, Retrieved: 18 March 2019

while we consider fun as an experience or state, which is a more extensive, nuanced, and complex notion, hence is yet more difficult to grasp and define.

Despite the growing interest towards the fun experience - with special regard to its relation to learning - currently there is a lack of a commonly accepted conceptual framework for the notion of fun. The aim of the current chapter is therefore to investigate the notion and the meaning behind the term fun.

2.2 Background

To understand what fun actually is, in the following sections we review literature related to fun, covering various contexts.

2.2.1 Earlier Attempts to Define Fun

In the academic literature, only a handful of authors have taken the time to suggest a definition for the concept of fun. Rather, we often see a disclaimer like *"defining fun is illusive at best"* ([245], p. 23). Even in a recent book on the topic of gamified learning right at a beginning of the introduction we are confronted with the following: *"What makes a game fun? And what is "fun" actually? Unfortunately, providing an exhaustive and objective answer to these legitimate questions is likely to be an impossible task since having fun is a very personal activity that can be completely different from individual to individual"* ([71], p. xv). Others, in the entertainment industry simply say that *fun is doing something that you want to do without purpose* [183]. However, it is not unusual, especially within the field of child-computer interaction in the recent years that fun is handled as a common sense term and therefore, authors simply do not elaborate on what their understanding of the concept of fun is [127, 150, 307]. For example, the authors of the fun Toolkit [234] did not even describe their understanding of the term fun despite building a measurement instrument around it.

Among the few who tried to define fun we see a tendency to describe fun in terms of its core attributes. Bisson and Luckner [32] synthesized earlier scholarly attempts to define fun into four characteristics that were inherent to it. They argued that fun is a *relative*, *situational*, *voluntary experience* and *natural/essential* to all human beings. With relative and situational they mean that fun depends on many factors, e.g., what one finds fun is not necessarily fun for another, nor is it certainly fun on another day. Fun is voluntary as *"to experience fun one must consciously or unconsciously accept to feel good, to relax, to let go and to let the situation be perceived as enjoyable"* ([32], p. 109). Fun cannot be forced, so, for an activity to be perceived as fun, the participation must be intrinsically motivated.

Glasser [98] argued that fun is one the five most essential human needs, and emphasized its importance while learning and especially for child development. He stated that fun *"is like a catalyst that makes anything we do better and worth doing again and again"* ([98], p. 28). Accordingly, Read, MacFarlane and Casey [236] discussed *returnance* as a facet of fun in child-computer interaction. In this context, returnance meant the desire to do an enjoyable activity again and again.

Based on a three-year-long study on attitudes towards physical education Dismore and Bailey [72] argued that the meaning attributed to the concept of fun changed in merit as children approached teenage years. While for younger children (7-11 years) fun was a critical factor for an activity to be enjoyable, teenagers (11-14 years) described fun in terms of a learning challenge rather than in relation to hedonic responses while playing games. Along with the aforementioned properties of fun, its stress-reducing effect has to be mentioned as well [47].

2.2.2 Fun and Learning

Research into understanding the role of fun in the learning process is found mainly in the areas of serious games and gamification in education [53], thus concerns mainly the non-formal and informal learning environment. In these fields, there is growing evidence in support of the importance of fun for learning.

Bisson and Luckner [32] elaborated on the pedagogical benefits of fun. They saw fun as a powerful tool to enhance motivation and create a safe learning environment. They summarized that fun is beneficial as a) it evokes intrinsic motivation, b) it facilitates the suspension of the social reality, c) it reduces stress, and d) it creates a state of *relaxed alertness* where *"learners feel safe to take risks, be creative, make mistakes, and most importantly, keep trying"* ([32], p. 111).

Rambli, Matcham and Sulaiman [232] stated that fun and interactive learning are one of the most powerful pedagogical factors, which could yield to create the interactive and engaged learning environment. They also added that this environment facilitates the memorization procedure of learners while keeping their attention, which ultimately enhances learning. Based on the PISA test (N > 400,000 15-year-old students) Ainley and Ainley [5] concluded that the sense of fun and excitement has a huge importance for science learning.

While investigating the learning outcomes of an educational game, Long [169] found that 87.5% of the participants joined the activity in the first place because of the promise of fun (*fun in programming games*) and it improved the skills of 79.8% of the participants. Moreover, she found that computer *games "lead to positive results in long-term learner retention by improving learning interest and more focused attention because the students enjoy the approach"* ([169], p. 280) - although she used the terms fun and enjoyment interchangeably. She also demonstrated that fun has a significant effect on the learning effort.

In a recent systematic review on emotions in design-based learning (DBL) Zhang et al. [342] found that design-based learning had an overall positive effect on students' interest and motivation to learn. Accordingly, enjoyment was among the most frequently mentioned emotions in DBL. However, this review also reflected that fun and enjoyment are barely distinguished or measured separately.

Vieira and da Silva stated that *“fun is an important element of life because it satisfies curiosity and fosters learning”* ([317], p. 130). They encouraged designers to ‘make their artefacts fun’ to stimulate users to use them. In their understanding fun consisted in attention, flow, immersion, and emotion. Tews and Noe even went further and said that *“fun is an important component of high quality learning experiences”* ([296], p. 226).

Chan, Wan and Ko [51] investigated the role of perceived fun in a collaborative learning scenario and the learning performance while using personal response systems (PRSs). Their results suggested that *“the level of fun students experienced using PRSs was found to promote collaborative learning and learning performance”* ([51], p. 99).

Iten and Petko [127] studied whether fun playing an educational game was a predictor for learning success. They used the terms fun and enjoyment interchangeably. They found that the experienced enjoyment and flow during the game had a significant effect on gaining motivation, increasing interest in the subject matter, and upon choosing to play the game again. However, their study could not demonstrate any association between the experienced enjoyment and the learning gains, which is in contrast with previous findings. Additionally, they questioned *“whether fun and enjoyment are adequate constructs to grasp meaningful motivational processes in serious game experiences”* ([127], p. 161). They referred to other authors who proposed instead ‘student engagement’ to analyze positive emotions when learning with serious games. Similarly, Sim, MacFarlane and Read [275] did not find significant correlation neither between the observed nor the reported fun and the learning outcomes.

Controversially, Tews, Michel and Noe [295] found that fun had a significant impact on informal learning in the working environment (i.e., learning from others and learning from non-interpersonal sources). They stressed that *“researchers should not necessarily focus on fun as a unidimensional construct”* ([295], p. 52). Additionally, their findings suggested that the managers’ support for fun had a significant influence on learning (learning from oneself) as well.

Similarly with adults, but in the learning environment, Lucardie in her qualitative research found that *“both adult learners and their teachers also believed that fun and enjoyment impacted on adults learning and they were able to articulate the role that fun plays in adult learning programs”* ([171], p. 445).

Elton-Chalcraft and Mills summarized their study results as follows: *“Learning which is enjoyable (fun) and self-motivating is more effective than sterile (boring) solely teacher-directed learning”* ([78], p. 482). This finding is supported by Aoki et al. [12] who investigated how the education of children with type-1 diabetes could be improved. They developed three edutainment tools and tested them. Their findings suggested that children patients found the games fun (compared to the researchers’ previous study on traditional learning methods), 91.4% of the respondents showed more interest toward the edutainment method, and more than 60% of them found that this approach would be useful as an initial education for type-1 diabetes children. They thus concluded that *"edutainment systems could have a significant potential for healthcare education especially for children"* ([12], p. 859).

Willis wrote about the neuroscience of joyful education. *"When students are engaged and motivated and feel minimal stress, information flows freely through the effective filter in the amygdala and they achieve higher levels of cognition, make connections, and experience 'aha' moments. Such learning comes not from quiet classrooms and directed lectures, but from classrooms with an atmosphere of exuberant discovery"* ([327], p. 1). She added that *"when classroom activities are pleasurable, the brain releases dopamine, a neurotransmitter that stimulates the memory centers and promotes the release of acetylcholinem, which increases focused attention"* ([327], p. 2). Additionally, she claimed that despite *"some schools have unspoken mandates against these valuable components of the classroom experience"* ([327], p. 3), no neuroimaging or brain wave analysis data exist that would demonstrate any downshifting effect of joy - a term she used interchangeably with fun - in the classroom. A summary of the benefits of fun while learning is provided in Table 2.1.

In sum, previous research provides a growing support for the view that fun has positive effects on learning. However, the above introduced studies do not define and measure fun, and moreover, none of these studies is a controlled experiment which would compare the effects of introducing fun elements versus not. The only one study - of Iten and Petko [127] - that comes close by studying effects of enjoyment but in that study no distinction is made between fun and enjoyment. To be able to make such claims precise regarding the effect of fun on learning, or to evaluate such activities, we need to define clearly what fun is, and have a reliable instrument for the measurement of it.

Table 2.1 Summary of the benefits of fun in the learning context.

Benefit	Source
Evokes intrinsic motivation	[32, 127, 169]
Helps keeping/increasing learners' attention	[169, 232, 327]
Increases interest/curiosity	[12, 127, 317]
Creates an interactive and engaged environment	[232]
Creates a state of relaxed alertedness	[32]
Facilitates suspension of the social reality	[32]
Reduces stress	[32]
Enhances learning	[51, 78, 169, 171, 232, 295, 317, 327]
Effects learning effort	[169]
Promotes collaborative learning	[51]

2.2.3 Fun and Play

During childhood, fun and play often concur. Despite that the two concepts are closely related, they should be distinguished from each other. Huizinga [121] elaborated in great detail on the concept of play and concluded that play has five main aspects: it is voluntary and free, rule ordered, happens within fixed boundaries (i.e., locality and duration), it is different from ordinary life and it has no material interest. Sutton-Smith defined play as *"an activity that is voluntary, intrinsically motivated, fun, incorporates free will/choices, offers escape, and is fundamentally exciting"* ([287], p. 3). Gajadjar, de Kort & IJsselsteijn considered play as an *"intrinsically motivated, physical or mental leisure activity that is undertaken only for enjoyment or amusement and has no other objective"* ([92], p. 105). Especially in early childhood, fun and play are overlapping notions. However, as indicated already [72], with adolescence the hedonistic character of fun that is purely present during play gives way to challenge, which is a well-established dimension of fun among adults [91, 126, 259].

2.2.4 Intrinsic Motivation and Social Aspect

Bisson and Luckner argued that the promise of fun *"can motivate learners to engage in activities with which they have little or no previous experience"* ([32], p. 110). Therefore, fun is not only an experience, but it can be itself a strong, intrinsic motivating factor to encourage children to try new challenges. Already back in the 80s, Malone and Lepper [172, 173] studied (educational) computer games to understand better what makes them interesting and exciting for children. They found that intrinsic motivation was a key factor, which, according to their theory, could be evoked by the optimal level of challenge, curiosity, and fantasy. Bisson and Luckner [32] also suggested that the combination of fun and play can act as a catalyst to eliminate inhibiting factors inherent to our socialisation. In their opinion, *"the more genuine and intense the fun is, the greater the suspension of reality will be. Consequently, fun can transform social insecurity into trust and camaraderie, and a restrictive self-image into the freedom of expression"* ([32], p. 110). This property of fun is closely related to the concept of *Flow*.

2.2.5 Flow

The concept, or rather the experience of Flow was defined by Csikszentmihalyi [61] as an optimal experience of any sort of activity, where the following characteristics were present: a) an intrinsically rewarding experience, b) a loss of reflective self-consciousness, c) a distorted experience of time, d) an intense and focused concentration on the present, e) a merge of actions and awareness, f) an optimal balance between challenge and skills, g) a sense of control over the situation, and h) clear goals and immediate feedback. Considering the definitions above it appears that experiencing fun and being in the psychological state of Flow overlap substantially.

Abbasi et al. [1] argued based on structural modelling that the theoretical constructs experience (a.k.a. Flow) and engagement should not be used interchangeably to investigate the subjective experience of video game play. Rather, they proposed a model of playful-consumption experience, which consisted of different types of experience (emotional and sensory) and different types of engagement (cognitive, affective, and behavioral), and they discussed enjoyment as one of the emotional experience factors, which, in their terminology, used interchangeably with the term fun. Thus, they considered enjoyment or fun as a part of the emotional experience.

As Tasci and Ko described, *"a distorted sense of time, in general, is taken as an indicator of engagement, desire, enjoyment, excitement and thus, having fun; when it feels as if time went more quickly than it actually did, this is a sign of fun and vice versa"* ([293], p.167). Rodriguez-Ardura and Meseguer-Artola [243] also stated that when one feels that the activity is going smoothly and is fun, then time flies, and one undergoes a distortion of the temporal experience of time.

2.2.6 Challenge

Caine and Caine [47] showed how learning is maximised when combining fun and challenge, which they called a state of relaxed alertness, and suggested that a major goal for educators should be to challenge students in a natural way so conceptual mapping (i.e., intellectual connections) could happen without evoking a downshifting response. Mellecker, Lyons, and Baranowski have also found while evaluating video game design with children that *"an engrossing story in which a player faces increasing challenges and can increase skills quickly enough to overcome the challenges, but not so quickly as to get bored by the challenges, appears to provide an important game design structure for enhancing fun or enjoyment"* ([186], p. 144). Chu, Angello, Saenz, and Quek [53] described how during a curriculum-based making activity children had the most positive feelings and got engaged the most when the level of challenge matched their skills. It has been mentioned already how adolescents attributed experiencing fun to challenge [72] which corresponds to the challenge aspect of the Flow experience. Rodriguez-Ardura and Meseguer-Artola [243] explained the effect of challenge on Flow in terms of the cognitive evaluation theory [258]. They suggested that as long as individuals have a psychological need to feel competent, activities that trigger positive challenge can lead to experiencing optimum experience and intrinsic motivation - because they satisfy the individual's need for competence.

2.3 Conclusion

Based on the reviewed theories we can assume that fun is a multidimensional construct. The importance and the interrelation of the concerning concepts such as control over the activity ([61]; identified from the learning sciences literature), challenge ([53, 61, 72, 91, 126, 186, 259]; identified from Flow Theory [61], Intrinsic Motivation Theory [32], and game literature), enjoyment ([32, 98, 236], identified from game literature), engagement and immersion ([47, 61, 293]; identified from Flow Theory [61] and game literature), intrinsic motivation ([32, 61, 98, 172, 173]; identified from Intrinsic Motivation Theory [32] and learning sciences literature), social connectivity ([32, 61]; identified from learning sciences literature), and stress ([47]; identified from learning sciences literature) has been argued in the earlier sections. Based on this reasoning the following dimensions were identified that cover all the referred theories and were expected to be defining for the experienced fun: *Autonomy*, *Challenge*, *Delight*, *Fear of Damage*, *Immersion, Loss of Social Barriers*, *Pressure* and *Stress*. These initial dimensions are defined as follows.

The *Autonomy* dimension assesses whether one experiences control over the activity. The *Challenge* dimension assesses whether one feels challenged during the activity. The *Delight* dimension describes the positive emotions experienced during the activity. The *Fear of Damage* dimension aims to assess whether one experiences fear of hurting someone or causing damage. The *Immersion* dimension intends to indicate whether one immerses in the activity by losing the sense of time and space. The *Loss of Social Barriers* dimension aims to monitor one's social connectivity. The *Pressure* dimension focuses on whether one experiences their own participation as voluntary or as obligatory. And the *Stress* dimension describes the negative emotions experienced during the activity, and is contra-indicative for having fun.

To empirically validate this theoretically grounded definition, we followed a deductive scale development approach [3], which is described in the following chapter along with the comparison of our definition to other, related constructs (see Chapter 3). After the empirical validation, we ended up with our final definition for fun as follows:

Fun is an emotional experience during which one feels in control over the activity and is intrinsically motivated for participation, one experiences an optimal level of challenge matching their level of skills, one feels 'well' during the activity and does not feel 'bad', one is immersed in the activity losing the perception of time and space, while letting go of social inhibitions.

3 Measurement of Fun – Introducing FunQ[7]

In the previous chapter we have reviewed the body of literature and established our initial definition for the notion of fun on theoretical grounds. In this chapter we discuss three studies that aimed to contribute to the empirical validation of the notion of fun, and ultimately, to create a tool for the reliable measurement of fun, which we named FunQ. Therefore, this chapter contributes with the empirical validation of our previously established definition for fun, and that of the related measurement tool, FunQ.

Summary

In this chapter we introduce three studies with the goal of a) empirically validating the previously established definition for fun, and b) designing and validating the related measurement tool, FunQ. In the first study we tested the initial questionnaire item pool with 75 students (M_{age} = 11.78). In the second, think-aloud study we tested the comprehensibility of the items with six 11-year-old students. In the third and final study 128 students (M_{age} = 12.15) participated, and their data contributed to the validation of FunQ. For the scale development, we applied a deductive scale development approach. For the model testing CFA was used and second-order latent variable models were fitted. In this chapter we thus introduce and discuss the final 18-item version of the FunQ that consists of six dimensions (Autonomy, Challenge, Delight, Immersion, Loss of Social Barriers and Stress) and bears with the appropriate validity and reliability measures ($\omega_{overall}$ = 0.875 and $\omega_{partial}$ = 0.864; *RMSEA* = 0.052 and *SRMR* = 0.072). We end the chapter with suggestions for best practices for using self-reported scales in general, and FunQ in specific with adolescents. This chapter contributes with a) a multi-dimensional instrument for assessing experienced fun - the FunQ, b) a psychometric evaluation of the proposed instrument, and c) suggestions for best practices for using self-reported measures with child respondents.

3.1 Introduction

Despite the growing interest towards assessing the fun experience - with special regard to the relation to learning - currently there is a lack of reliable measurement tools. Where they exist (mostly in the field of human-computer interaction), they rather measure product liking (acceptance or preference) with young, preliterate children (e.g., Fun Toolkit [234]; Fun Semantic Differential Scales [338]; This or That [340]). Or in case of adults, the most widely known instruments (e.g., EGameFlow [91]; UES [204]; GEQ [226]) measure game enjoyment and engagement or the gaming experience along several dimensions.

[7] This chapter is based on the following publication: Tisza, G., & Markopoulos, P. (2021). FunQ: Assessing the fun experience of a learning activity with adolescents. Current Psychology. https://doi.org/10.1007/s12144-021-01484-2

Remarkably, there is a gap in research with adolescents (age 11-18). The aim of this chapter is to introduce our research that aimed to create a tool for the measurement of the experienced fun, which is psychometrically and theoretically sound, comprehensive yet parsimonious, practical and child appropriate, specially developed for adolescents and can be used in the learning environment across various fields of research.

The chapter is structured in three main parts. In the first part, we discuss the methodological challenges of developing questionnaires for children and adolescents, and we introduce existing tools for measuring fun, including their suitability for our target audience and scope of research (adolescents in the learning environment), and their possible pitfalls. In the second part, we describe the development of the FunQ. We start with the construction of the initial item pool, we follow with the think-aloud study results, then we finish with introducing the final version of the instrument, including its psychometric properties. The final version of the FunQ consists of 18 items along six dimensions (*Autonomy, Challenge, Delight, Immersion, Loss of Social Barriers and Stress)* and bears with the appropriate validity and reliability measures ($\omega_{overall} = 0.875$ and $\omega_{partial} = 0.864$; RMSEA = 0.052 and SRMR = 0.072). In the third and last part, we discuss our research findings, the possible applications of the developed measurement tool, we summarize our scientific contribution, and propose best practices for using self-reported measured with child respondents.

3.2 Background

3.2.1 Distinction Between Young Children and Adolescents

Approaching the childhood from the perspective of cognitive and psychological development, both Piaget and Erikson account for the shift that occurs on the edge of the adolescence. According to Piaget [218] the *formal operational stage* begins approximately at the age of 11-12, and lasts into adulthood when children develop the ability to think about abstract objects and to logically test hypotheses. In the theory of Erikson [81] stage 5 is approximately between age 12-18, during which children search for a sense of self and personal identity exploring their own personal values, beliefs, and goals. Without further delving into the characteristics of the teenage, it can be summarized, that during this age children's identity is formed and the way they understand the world changes. They begin to shape their opinion and they learn how to express it as well. Their cognitive abilities - such as memory capacity, language skills, concentration span, etc. - are approaching quickly the level of an adult's. De Leeuw argued that the *"age of 11 is seen as a turning point in memory capacity when children appear to function as well as adults"* ([163], p. 15). From the beginning of this age, therefore they are less prone to the typical response biases that are common for younger children (see section 3.2.3 Attention Span). Ultimately this means that when they are asked about their opinion, the answers will be more differentiated than in younger ages (this is also shown by [233–235]) and are generally more valid [187].

Moreover, Dismore and Bailey [72] showed that the meaning and the content of the concept of fun altered during childhood. This shift was found to be around age 11 as well, which is in synchrony with the psychological and cognitive changes while approaching adolescence. Moreover, the World Health Organization (WHO) also defines adolescents being between age 10 and 19[8]. On the basis of these arguments, the definition and measurement of fun will address children over the age of 10.

3.2.2 Surveying Children and Adolescents

In designing measurement tools specifically for child and adolescent respondents, their competencies and differences to adults have to be taken into account. Hall, Hume and Tazzyman [108] emphasized that generally, children preferred Likert-type scales over similar simple response items and that free-recall questions were useful especially in spoken surveys [235]. De Leeuw [163] found that in general, the older the child the more reliable the answers will be, and that children were better informants on topics directly related to them such as their feelings and other subjective phenomena. Further, de Leeuw stated that *"below the age of 7 children do not have sufficient cognitive skills to be effectively and systematically questioned"* ([163], p. 6) and added that individual (semi-) structured interviews were more suitable than questionnaires for children between 7 and 12 (see also [27]). Mellor and More [187] found that children below the age of 12 had difficulties in answering questions about abstract concepts such as their own behaviours, bodily states or emotional states. They related it to the theory of Piaget about the formal operational stage of development. Regarding adolescents, de Leeuw [163] suggested that questionnaires could be used as for adults, however, there should be special attention devoted against ambiguity in item wording. Therefore, using simple language and formulating items as exactly as possible is a must and ensuring language appropriateness by readability testing is highly recommended.

Read [234] and de Leeuw [163] discussed the challenges of designing measurement tools for children and adolescent, and Mellor and Moore [187] wrote specifically about the use of Likert-type scales with young respondents. The main caveats which can jeopardize the reliability of surveys involving children and adolescents are discussed below.

3.2.3 Attention Span

Children's and adolescents' attention span (or sustained attention) is crucial for directly measuring them (i.e., not observing their behaviour). Attention span is defined as the time a person is able to selectively attend to relevant information, such as listening to a teacher and persisting on a task [181]. Scientists have been studying attention span for a long time. In a literature review dating back to the 50s Moyer and von Haller Gilmer [196] reviewed nineteen studies which found that the attention span of young children ranged from one

---

[8] https://www.who.int/health-topics/adolescent-health#tab=tab_1

to twenty-five minutes. Additionally, they found that the attention span was lower in a group situation and that there was a difference *"between work tasks, such as reading, and the activities in which a child engages in playing with a toy"* ([196], p. 466). Moreover, Sousa [278] suggested that motivation had an effect on the attention span, and Bradbury [41] even called this effect crucial.

Within the field of educational psychology, a generally accepted and referred rule of thumb for the length of the average student's attention span is 10 to 15 minutes during lectures [66, 100, 182, 329]. Although formulas are available for approximating the length of the attention span per age, no supporting empirical evidence exists. Lin, Hsiao, and Chen [164] showed that the sustained attention developed between ages 6 and 15. Nonetheless, since they have measured the sustained attention by the Continuous Performance Test (CPT), they only report on the hit- and false alarm rates for the evaluation of sustained attention and did not provide the length of the attention span in minutes.

Controversially to the generally accepted 10-15 minutes, a study conducted at Microsoft [190] claimed that the average attention span was only 12 seconds in 2000, 8 seconds in 2013 and it is ever decreasing. Although the validity of these numbers has been contested [41], it appears that people nowadays, and especially the younger generation, get distracted easily and hold their attention for less time than was the case in the past.

To safeguard the reliability of answers, it is important to consider children's and adolescents' attention span adjusting the length of the inspection to their attention span. Based on the above, a survey - that is not a particularly engaging task for a child - should not require sustained attention by adolescents for more than 10-15 minutes.

3.2.4 Bias

When working with children and adolescents, the risk of introducing bias is high and different types of bias can be manifested compared to those concerning adults. The most common bias types concerning children and adolescents are discussed below.

Suggestibility pertains to the influence of the researcher on the way the respondent encodes, stores, retrieves, and reports events. This effect is due to a range of social and psychological factors. *Social desirability bias* is when the respondent provides the answer that they think the examiner asking the question wants to hear. *Satisficing* is a tendency of the respondent to select a good enough option, instead of the very best one. In the case of surveys, this phenomenon could be manifested as giving a superficial response that appears to be reasonable, but without thoroughly considering all answer possibilities. *Acquiescence bias* is the tendency of the respondent to agree or respond positively. *Extreme responding* is the type of response bias when the respondent mainly selects the most extreme options/answers available. *Straight-lining* is the tendency of the respondent to provide answers in a way that the responses form a line, or rather a visual pattern. Extreme responding is a sort of straight lining, however, meanwhile in extreme responding it can be assumed that the respondent reads the question and considers the responses, in case of straight-lining this assumption cannot be made.

Around the age of 11, the suggestibility of children decreases while the importance of peers increases. Therefore, *peer pressure* can be a serious issue with early adolescents (12-16 years) [163]. However, contrary to adults, the *item non-response* appears not to be a problem with children and adolescents [27]. That is, the error size in responses by children and adolescents is approximately stable across different conditions and not dependent on the content of the question. This is assumed to be in relation to their cognitive abilities, namely that they cannot fully apply an optimizing strategy [37], therefore they will not skip difficult questions. It has, however, a downside. The difficult, or vague questions will not be indicated by a missing value pattern, but the quality of those responses remains doubtful. Therefore, the importance of simple and short questionnaire items is stressed and the application of think-aloud interviews to check whether any an item is problematic is highly recommended.

Earlier research has shown that children are particularly prone to the above-described bias types [27, 108, 163, 234, 235] to very different degrees for different ages, which has to be considered when developing measurement instruments for any specific age-group.

3.2.5 Existing Measurement Tools

Methods for evaluating fun derive primarily from the domain of human-computer interaction and especially child-computer interaction, where fun is seen as an essential component of children's interactive experience whether they approach technology as users (e.g., of an application or consumer device), learners or players [176].

There are a few existing measurement tools that have been designed to gather opinion on the 'funness' of an experience or product. Where these exist, they target either young children or adults. Some studies report the use of a survey or a list of questions to be asked from children within the narrow scope of the study, however, they are not intended for further use nor are they validated. A list including the most-known tools for measuring (aspects of) fun is shown in Table 3.1.

The This or That [339] method examines preference. Despite being a validated measurement tool, it is constrained by its comparative structure: it is only suitable when measuring the preference of one product/experience over the other, which is particularly suitable for its targeted age group of 2 to 7.

The Fun Semantic Differential Scales [338] is a measurement tool for evaluating games with nursery-aged children based on choosing between photos of a child expressing different emotions (*love - don't know - hate*). While it has been shown to work well for the target age group, it has not been psychometrically validated, it addresses fun as a unidimensional construct, and is not sufficiently refined for teenagers (see Table 3.1).

The Fun Toolkit [234] is a set of tools that targets a wide age range up until teenage to measure the 'funness' (Smileyometer) and preference of products (Fun sorter and Again-

again table). However, it handles fun as a unidimensional construct, and it faces a problem that younger children tend to use mainly the higher values of the Smileyometer. Despite being widely used, it has not been psychometrically validated (see Table 3.1).

The Five Degrees of Happiness [108] was introduced to address the extreme response bias of the Smileyometer discussed above. Its target audience is children between age 9 and 11, and like its predecessor handles fun as a unidimensional construct and has not been validated psychometrically. Further, the emphasis on positive emotions makes it less suitable for assessing less pleasant experiences that might include frustration or disappointment (see Figure 3.1).

The study of Iten and Petko proposed a list of Likert scales for the evaluation of a game and attitudes towards learning games [127], which however, has not been validated. Despite being suitable for teenagers and measuring multiple dimensions, those dimensions are not linked to fun and refer to the serious game rather than to the personal experience (i.e., How is the game instead of How do one feels while playing the game).

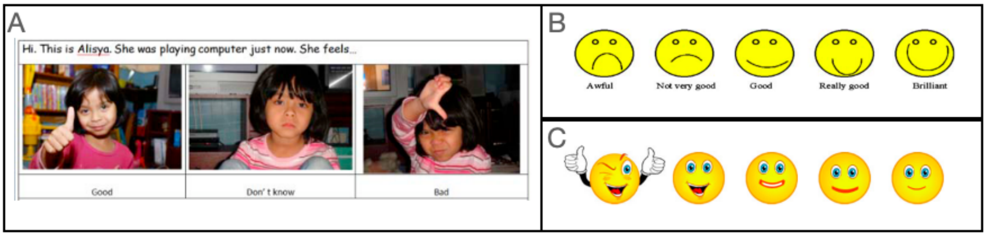


Figure 3.1 Examples of existing scales for measuring children's preference. (A) fun Semantic differential Scales [338]. **(B) The fun Toolkit** [234]. **(C) Five Degrees of Happiness** [108].

This limitation is avoided in the Physical Activity Enjoyment Scale [145], which measures the personal experience of a physical activity, rather than the activity as such. However, it does not help conceptualize enjoyment which is measured as a unidimensional construct across bipolar scales.

The PENS [259], UES, and UES-SF scales [204] are validated measurement tools, made for adults, and have a strong focus on the evaluation of games (usability, aesthetics, novelty, intuitive controls, in-game competence etc.) rather than on the personal experience (flow, intrinsic motivation etc.). This limits their applicability in different contexts and does not contribute to our purpose of defining and measuring fun as a psychological construct.

The GEQ [225] has the focus on how one feels while playing a game and measures enjoyment as a multidimensional construct, however, it has been validated only with adults in a gaming environment, which limits its applicability in different contexts. Besides, given that the scale is designed for adults, its vocabulary is quite advanced, so it is questionable whether young respondents would be able to comprehend and rate the scale reliably.

The Player Experience Inventory (PXI) [2] measures player experience at two levels: the level of functional and psychosocial consequences. As its name suggest, it focuses on the player experience, and while most of the items focus on how one feels while interacting with a game, it does not measure enjoyment or fun, and the scale has been designed for and validated with adults.

In the same domain a study [1] proposed a Playful-consumption experience questionnaire for the assessment of consumer video game engagement for adolescents. However, they measured enjoyment as a unidimensional construct and as a subdimension of emotional experience, and the questionnaire has not yet been validated.

The EGameFlow [91] measures the enjoyment of an e-learning game across the dimensions of Csikszentmihalyi's Flow theory, thus it equates enjoyment to the Flow experience, and it is validated for adults. Given the nature of the scale, its usability in different contexts with different ages is limited.

The FUN scale [293] handles fun as a multidimensional construct, however it is validated to measure the fun value of a touristic destination as a product among adults, which is reflected in its vocabulary. Additionally, the focus of the scale is to evaluate whether a place, a hotel or a restaurant is fun and not to assess the personal experience.

The EmoForm [344] is a tool for the assessment of emotions during Design Based Learning. While the instrument is designed for adolescents it examines various emotions, it does not examine fun, only enjoyment across a single item, and it has not yet been validated.

Table 3.1 Measurement tools for evaluating preference/engagement/experienced fun

Tool's name	Year	Age	Psychometr. validated	Internal consist. †
This or That [339]	2013	2-7 yrs	yes	yes
Fun Semantic Differential Scales [338]	2011	3-5 yrs	no	no
Fun Toolkit [234]	2008	4 yrs - teenagers inclusive	no	no
Five degrees of happiness [108]	2016	9-11 yrs	no	no
List of statements for the Evaluation of- and attitudes towards- learning games [127]	2016	10-13 yrs	no	yes
Physical Activity Enjoyment Scale [145]	1991	Undergrad. students	yes	yes
Player Experience of Need Satisfaction (PENS) [259]	2006	Undergrad. students	yes [135]	yes
User Experience Scale (UES) and UES-SF [204]	2018	Adults	yes	yes
Game Experience Questionnaire (GEQ) [225]	2007	Undergrad. students	yes	yes
Player Experience Inventory (PXI) [2]	2022	Adults	yes	yes
Playful-Consumption Experience Questionnaire [1]	2019	16-19 yrs	no	yes
EGameFlow [91]	2009	Undergrad. students	yes	yes
FUN scale for understanding the hedonic value of a product [293]	2016	Adults	yes	yes
EmoForm [344]	2019	13-14 yrs	no	yes

† whether any internal consistency measures (e.g., Cronbach's alpha or Omega) are published

From this review of earlier work, we can conclude that there is currently no psychometrically validated inventory in general, and targeting adolescent respondents in specific, that is theoretically grounded and that treats fun as a multi-dimensional construct. The necessity of having multiple dimensions was not only shown by the theoretical review in the previous chapter, but it helps to conceptualize and define fun rather than treating it as an opaque descriptor or an umbrella-term. The present chapter introduces an instrument designed to fill this gap by providing a tool that is psychometrically and theoretically sound, comprehensive yet parsimonious, practical and adolescent-appropriate, and can be used in the learning environment across various fields of research.

For the development of the FunQ only a fraction of the referred scales were relevant. It is important to mention, that several items of different measurement tools overlap with each other (e.g., items measuring the Flow experience, the perceived competence, the enjoyment etc). Selected questionnaire items from the EGameFlow [91], the Evaluation of- and attitudes towards- learning games list [127], the FUN scale [293], the GEQ [225], the Physical Activity Enjoyment Scale [145], and the Intrinsic Motivation Inventory [256] have been included in the initial item pool of FunQ, albeit, rephrased to be adolescent appropriate and to reflect a personal experience instead of evaluating the activity (e.g., *I had fun* instead of *It's [the activity] a lots of fun*). Additionally, the pool of items was extended by further ones that reflect the underlying factors, and the adopted items were organized in the FunQ by the factorial structure proposed in the previous chapter and not by the dimensions of the original instrument.

3.3 Development of the FunQ

The development of the FunQ consisted of four main phases applying a deductive scale development approach [3]. During the first phase, initially, the dimensions of the instrument were constructed based on theoretical grounds (see Chapter 2). Then, a pool of possible items was created according to the previously defined dimensions, based on the cited theories and existing measurement tools. Second, the initial item pool of the instrument was tested with 75 students. Consequently, a comprehensive yet parsimonious model was created. Third, we conducted think-aloud interviews with six students to assess possible pitfalls relating to the questionnaire items. Fourth, the questionnaire was administered to another 150 students and the validity and reliability measures were calculated.

3.3.1 Methodology of the Survey Design

In the development of the FunQ, we paid attention to a number of key issues as follows.

First, to gauge relevant abstract concepts - such as challenge, fun, flow, stress, and autonomy - questions were formulated by asking the emotional and behavioural reflections of these concepts. Then statements were derived as responses to the questions based on the underlying theories. For example, *How do I feel when I am enjoying an activity? - I feel* *delighted;* and *How do I feel when I'm in Flow? - I feel that time flies.* Then the emotional and behavioural reflections were filtered, and the wording was adapted to be youth appropriate, e.g., *I feel delighted* was transformed into *I feel good,* and *I feel happy.* This approach aligns with the recommendation by de Leeuw [163] that the items should be worded to focus on how one felt during the activity. It is not only easier for children to identify themselves in such items but also, this phrasing avoids asking them to judge the activity, thus reducing the risk of social desirability bias, which could arise if they had to evaluate an activity designed by the teacher or another adult. Additionally, we can expect that the respondents will be more likely to use the full range of the scale rather than the extremes of the scale. This latter issue has been extensively addressed by previous studies [108, 233–235] on scales developed specifically for children.

Second, the questionnaire uses contra-indicative statements to measure factors that could indicate one is not having fun. The sentences within these factors are phrased in a non-negating, thus positive way, however, their content is anticipated to be contra-indicative for experiencing fun (e.g., During the activity I felt bad). Such factors are the experienced stress and tension, which are presented in the questionnaire by the initial factors *Fear of Damage, Pressure and Stress.* Having contra-indicative items among the statements of the questionnaire is intended to prevent - or highlight - *acquiescence bias* and *straight-lining* (see section 3.2.4 Bias) and thus serve as a control for the reliability of the answers.

Third, the questionnaire items were phrased very briefly and in simple language, so that young respondents would find them easy to comprehend and evaluate. Language appropriateness was checked by several measures, which from the Flesch Reading Ease [87] score is 77.9 and the Flesch-Kincaid grade level [148] is 3.7 for the initial - 50 item - version of FunQ, and 84.7 and 2.7 respectively for the 18-item final version. This indicates that the text of the 50-item questionnaire is respectively *fairly easy to read* and is understandable for an average end-of-third-grade student (age 9); and the text of the 18-item final version is *easy to read* and is understandable for an average end-of-second-grade student (age 8). This is in agreement with de Leeuw's suggestion [163] that the readability level of items should be about two grades lower than the target group.

Fourth, the appearance of the survey was created by considering the specialties of the target group. Therefore, based on previous findings [29] the text was presented in Comic Sans type with 12 pt size, which was found to be the most preferred font type and size among 9-11 years old respondents. Besides, the items were highlighted with alternating colours - so that it makes easier to keep track of the responses -, and additionally a colourful design was created to make the questionnaire inviting for children and adolescents (see Appendix A). The idea of adding cartoons to the design was considered but was abandoned for fear of the questionnaire appearing too childish for adolescents.

Fifth, there was special attention devoted to the type of response format. Based on the findings of Mellor and Moore [187] and de Leeuw [163] it was decided to use a 5-point

Likert scale where the points are based on words that reflect the frequency of behaviour/thoughts (i.e., never/rarely/sometimes/often/all the time).

Sixth, the questionnaire was designed so that the response time should stay within the anticipated average concentration span of 10-15 minutes [66, 100, 182, 329]. The questionnaire items once finalized, were randomly mixed.

3.3.2 Construction of the FunQ

The factorial structure of the FunQ was created by adopting a deductive scale development approach [3] following similar steps as in previous research [175, 227]. That is, the factors were established strictly on the previously referred theories. Based on those theories we made the assumption that the experienced fun is a multidimensional construct. The importance and the interrelation of the concerning concepts such as control over the activity [61], challenge [53, 61, 72, 91, 126, 186, 259], enjoyment [32, 98, 236], engagement and immersion [47, 61, 293], intrinsic motivation [32, 61, 98, 172, 173], social connectivity [32, 61], and stress [47] has been argued in the previous chapter (see Chapter 2). Based on this reasoning the following factors were established that covered all the referred theories and were expected to be defining for the experienced fun: *Autonomy*, *Challenge*, *Delight*, *Fear of Damage*, *Immersion, Loss of Social Barriers*, *Pressure* and *Stress*.

Then, the referred frequently used measurement tools were scrutinized whether they measure any of the dimensions of the FunQ. Consequently, some items from other measurement tools [91, 127, 145, 225, 256, 293] were considered to be taken into the FunQ. For this, the above-detailed protocol for rephrasing and adjustment was followed. Thereafter, the number of questionnaire items were further expanded by items based on the emotional and behavioral reflections of the factor defining concepts, while keeping the number of items limited with regards to the attention span of the target respondent population. The item pool was evaluated and adjusted in several consecutive steps according to topic experts' recommendations.

In the initial version of the FunQ, the *Experienced fun* is measured across eight dimensions. The *Autonomy* factor (4 items) measures whether the child experienced control over the activity. The *Challenge* factor (10 items) assesses whether the child felt challenged during the activity. The *Delight* factor (9 items) targets the positive emotions experienced during the activity. The *Fear of Damage* factor (4 items) aims to control whether the child experienced fear of hurting someone or causing damage. The *Immersion* factor (8 items) intends to indicate whether the child immersed in the activity by losing the sense of time and space. The *Loss of Social Barriers* factor (4 items) aims to monitor the social connectivity of the child. The *Pressure* factor (5 items) investigates whether the child experiences his/her own participation as voluntary or as obligatory. And the *Stress* factor (5 items) measures the negative emotions experienced during the activity. In total, the initial version of the questionnaire consisted of 50 items which from 16 were reverse statements (see Appendix B). The items were evaluated by the students on a 5-step Likert-type scale.

3.3.3 Analysing the Structure of Initial FunQ Item Pool

To assess the fit of the FunQ for its aimed purpose, the initial 50-item version of the FunQ was administered to students after they visited an interactive exhibition about the Dutch Delta Works. Based on the statistical analyses a comprehensive yet parsimonious model was created which contains 18 items across six factors. Some items of the final model were slightly adjusted according to the following think-aloud study (see section 3.4.2 Think-Aloud Evaluation of Initial Item Pool), which slightly modified 18-item version of the FunQ was used for the third study.

3.3.4 Think-Aloud Evaluation of Initial Item Pool

Besides statistically testing the fitness of the FunQ, we conducted think-aloud interviews [80] with six students, for which we used the initial 50-item version of the instrument. The think-aloud interview is a commonly used method for assessing participant's thought processes especially when confronted with a new situation or artefact. During the interview, the interviewee is asked to verbalize their thoughts on the subject of testing while being actively engaged with it. With the think-aloud interviews we aimed to get an insight a) whether are problematic or misunderstandable items, and b) whether the respondents have the same understanding of the questionnaire items as it was intended by the researchers. In our case, the procedure was as follows. To start with, the interviewer explained the method and gave examples to the student what is expected from them. Thereafter, the student was asked to read the FunQ items aloud and verbalize any thoughts that came in their mind. While administering the think-aloud interviews we followed the recommendations of Markopoulos et al. [176] for conducting think-alouds with children. Accordingly, the role of the interviewer in this situation was mainly to observe, but if needed, to facilitate the verbalization and eventually to ask for clarification. Therefore, in contrast with the recommendations of Erikson and Simon (i.e., staying in background (not helping or explaining) and encouragement in a neutral manner), we applied a more relaxed approach to think-aloud (i.e., a more dialogical conversation, with more social interactions and an encouraging way of facilitation) to create a friendly and safe environment for the young participants.

3.3.5 Psychometric Properties of FunQ

As a final step in the herein described study, we collected further data from 150 students visiting a museum with their school to test the validity and reliability of the final 18-item version of FunQ on a new data set.

3.3.6 Statistical Analyses and Measures

The think-aloud interviews were analyzed qualitatively (see section 3.4.2 Think-Aloud Evaluation of Initial Item Pool). For the assessment of the psychometric properties of the instrument the statistical analyses and measures are detailed below.

We applied confirmatory factor analysis (CFA) and second-order hierarchical latent variable modelling. Our choice for CFA is supported by the deductive scale development methodology we followed. That is, the FunQ factors were established strictly on the above-referred theories, and according to previous papers [118, 285], CFA is the appropriate choice when *"the researcher uses knowledge of the theory, empirical research, or both, postulates the relationship pattern a priori and then tests the hypothesis statistically"* ([285], p. 1).

3.3.6.1 Internal Consistency

For measuring internal consistency, the Cronbach's alpha (α) and the Omega (ω) coefficients were calculated. These statistics indicate whether the items measure the same underlying construct. Despite Cronbach's alpha is the most widely known internal consistency measure, it has been the subject of considerable criticism [73, 195, 217, 241]. Thus, we adopted the Omega coefficient as the main indicator for internal consistency as it has been proven to be more reliable [217, 241]. Given that there is a wide variety for the acceptable internal consistency values (starting from 0.45 [290]) for the reported study we regarded the Omega values above 0.6 acceptable [107, 293].

3.3.6.2 Model Fit

For assessing model fit (whether the factorial structure found in the data is in agreement with the proposed one) we considered a variety of model fit indexes and addressed parameter estimates and their magnitudes. Hu and Bentler [117] suggested to rely on a combination of indexes that have different measurement properties (e.g., CFI and SRMR). We used the selected indexes listed below based on the recommendations of Jackson, Gillapsy and Purc-Stephenson and Kline [128, 149].

- $\chi 2$ value. The $\chi 2$ value is a general, commonly used maximum likelihood approximation for the overall model fit. It tests whether the model implied covariance matrix differs significantly from the measured values. However, it is known that the $\chi 2$ value is affected by the sample size and is mostly significant when $N > 75$.
- Comparative Fit Index. The Comparative Fit Index (CFI) is an incremental fit index which ranges from 0 and 1, with higher values indicating better model fit. The cut-off value for the CFI proposed by Hu and Bentler [117] is 0.95 or higher, which indicates a good model fit.
- Root Mean Squared Error of Approximation. The Root Mean Squared Error of Approximation (RMSEA) is a parsimony-based index. The index value typically ranges between 0 and 1, but higher values than 1 are also possible. An index value of 0 indicates a perfect model fit. When the p-value is ≥ 0.05, then the hypothesis of

close fit is justified. The cut-off value for the RMSEA proposed by Hu and Bentler [117] is 0.06 or lower, which indicates a good model fit.

- Standardized Root Mean Squared Residual. The Standardized Root Mean Squared Residual (SRMR) is an absolute fit index. The index value ranges between 0 and 1, where 0 marks perfect fit. Thus, the lower the value, the better the fit is. The cut-off value for the SRMR proposed by Hu and Bentler [117] is 0.08 or lower, which indicates a good model fit.

For modelling, second-order hierarchical latent variable models were fitted. For the data analysis we used the RStudio 1.1.453 [252] software, for modelling the lavaan package [248], and the semTools package [138] to calculate the Cronbach's alpha and Omega coefficients.

3.4 Results

3.4.1 Analyzing the Structure of Initial FunQ Item Pool

3.4.1.1 Data

Data were collected in English (original language of the FunQ) at the beginning of October 2018 from 75 students from the first year of a Dutch secondary school with English speaking specialisation (39 boys, 33 girls, 3 not given, $M_{age} = 11.78$, $SD = 0.45$). Consent was attained according to the Dutch regulations: both parents and the child had to sign the informed consent form. The questionnaire was administered on paper after the students attended to Deltapark Neeltje Jans, an interactive exhibition and information center about the Delta Works, and designed and built their own dams. The data is coming from three groups.

3.4.1.2 Missing Data

The proportion of missing values both for the whole sample and the questionnaire items is 1.8%. Since the normality of the data cannot be assumed, we used the non-parametric test of homoscedasticity to check whether data is missing completely at random (MCAR). The test resulted in a nonsignificant p-value ($p = 0.259$), thus it was assumed that the values are missing completely at random. Hence, to handle missing data, full information maximum likelihood (FIML) estimation was used.

3.4.1.3 The Initial 50-Item Pool

To begin with, the construct validity of the proposed factorial structure was assessed by means of confirmatory factor analysis (CFA) and second-order latent variable modelling. The contra-indicative items and factors that are marked with an (R) (see Appendix B) were reversed for the data analysis and for reporting the results. For details about the calculation of the coefficients see the referred package [138]. The computation for overall internal consistency at the first-order level ($\omega_{overall}$ = 0.916), and for the second-order factor *Experienced fun* ($\omega_{partial}$ = 0.909) revealed a high value of Omega coefficient.

3.4.1.4 The Final 18-Item Model

Since we aimed to create a comprehensive and parsimonious model, the factor loadings of the first-order factors and the second-order factor *Experienced fun* were examined. First, it was investigated whether all proposed factors contribute equally well to the second-order factor *Experienced fun*. The analysis revealed that the *Fear of Damage* factor had no significant effect on the *Experienced fun* (*standardized factor loading* = 0.027, p = 0.891) therefore it was removed from the model. We elaborate on this decision in the following sections of this chapter (i.e., section .3.4.2.2 and 3.5). Then, the number of the questionnaire items was reduced based on the factor loadings. In case the standardized factor loading of an item was < 0.3, it was considered not to be substantial [107] for the given factor and in several consecutive steps the non-substantial items were removed. Additionally, based on the modification indexes and factor covariances provided by the lavaan package, the *Autonomy* and the *Pressure* factors were merged. During the model fitting process, the internal consistency of the modified factors and the model fit was continuously monitored.

Comparing the final 18-item model to the initial 50-item pool, the *Fear of Damage* factor was completely removed as it appeared not to be related to the second-order factor *Experienced fun*. Additionally, the *Pressure* and *Autonomy* factors were combined into one factor that measures the free choice/voluntary participation of the child, named *Autonomy*. The final 18-item version of FunQ is presented in the Appendix B.

Table 3.2 summarizes the statistics of the final 18-item model. The internal consistency of the majority of the remaining factors is above the cut-off value (ω > 0.6), which indicates that the items are measuring the same underlying constructs. Despite the internal consistency of the *Challenge* (ω = 0.477) and *Immersion* (ω = 0.488) factors is below the cut-off value, they both appear to have a significant effect on the *Experienced fun* ($p_{Challenge}$ < 0.001, $p_{Immersion}$ < 0.001) and a standardized factor loading well above the 0.3 margin (0.719 and 1.022 respectively), therefore the factors were kept in the model. In fact, the *Immersion* factor has a standardized factor loading above 1, which is unusual, however, acceptable, suggesting high correlations among the factors [137] that is desirable for second-order latent variable modelling.

The standardized factor loadings of the first-order variables on the second-order variable *Experienced fun* in the final model are all above the cut-off value of 0.3, therefore they are considered substantial.

Table 3.2 Statistics of final 18-item model on the first data set. The internal consistency coefficients and the standardized factor loadings of the factors on *Experienced fun*.

Factor	Cronbach's alpha	Omega	Standardized factor loading on *Experienced fun*	P-value
Autonomy	0.755	0.770	0.763	0.007
Challenge	0.525	0.477	0.719	< 0.001
Delight	0.801	0.808	0.996	< 0.001
Immersion	0.408	0.488	1.022	< 0.001
Loss of Social Barriers	0.632	0.647	0.577	0.001
Stress (R)	0.862	0.863	0.804	< 0.001
Overall	0.867	0.896	-	-

The internal consistency of the second-order factor *Experienced fun* is presented in Table 3.3. The Omega values are found to be above the cut-off value (ω > 0.6). This finding suggests that the *Autonomy, Challenge, Delight, Immersion, Loss of Social Barriers* and *Stress* factors measure with high reliability the same underlying construct, the *Experienced fun.*

Table 3.3 Statistics of the final 18-item model on the first data set. The internal consistency coefficients of *Experienced fun* as second-order factor.

	Omega at level 1†	Omega at level 2‡	Partial Omega§
Experienced fun	0.822	0.924	0.888

† the proportion of the second-order factor explaining the total score

‡ the proportion of the second-order factor explaining the variance at first-order factor level

§ the proportion of observed variance explained by the second-order factor after partialling the uniqueness from the first-order factor

The model fit indexes of the final 18-item model are introduced in Table 3.4. Despite the borderline values, given the relatively small sample size, the sufficient factor loadings, p-values, and internal consistency values, and to prevent hard fitting to the data, we decided to adhere to this model and test it on a new, larger data set.

Table 3.4 Fit indexes of the final 18-item model on the first data set

	χ^2 value	CFI	RMSEA	SRMR
Reference value	-	> 0.95	< 0.06	< 0.08
Final 18-item model	207.507; $df = 129$, $p < 0.01$	0.857	0.090	0.086

3.4.2 Think-Aloud Evaluation of Initial Item Pool

3.4.2.1 Data

The think-aloud interviews were conducted in English on 14 February 2019 at an international school in the Netherlands with six 11-year-old students after participating in a playful learning activity during which they prototyped a robot. Consent was attained according to the Dutch regulations: both parents and the child had to sign the informed consent form. Strengthening further the voluntary character of the participation in the interview, from the fifteen students who delivered the consent form complete, six who were willing were invited for the think aloud interviews.

3.4.2.2 Analysis

For the think-aloud interviews, the initial 50-item pool of the FunQ was used for the sake of completeness. For the evaluation of the interviews the following aspects were considered:

- Misreading of words
- Difficulty with reading the item (when no reading difficulty was observed in general)
- Asking clarification about the item
- Adding comments that imply that the item is not relevant
- Interpreting the item in a way that does not align with the intended meaning

In general, a good usability of FunQ was found. Specifically, the suitability of the design, the language, and the scale labelling was justified by the interviews. It appeared that students – even being at the lowest range of the target user age group – went through most of the questions smoothly. Implying that FunQ is user friendly, the appearance supports the evaluation of the items, which are readable (font type and size) and understandable (language). Also, the used labelling of the steps of the scale (chosen based on the suggestions of de Leeuw [163]) appeared to be adequate as it helped students to think back and identify themselves with the statements.

According to the above established criteria, based on the interviews, eight items emerged as problematic. Those are detailed in Table 3.5.

Table 3.5 FunQ items and the discovered issues during the think-aloud interviews.

Item	Issue
[E4] I want to do the activity again.	"Well, I think the answer would be never because why would I do the same thing, make the same robot again? But maybe I could change it. So the answer would be rarely, but if the activity was the same but I got to make some different things like a different robot, then I would take it as often"
[S6] This was an activity that I couldn't do very well.	'couldn't' was continuously misread as 'could'
[P3] I did this activity because I had no choice.	"Yes, I had to, but if I had a choice I'd still do it. But I'll just have to put all the time"
[A4] I could make some choices about the activity.	Some asked for clarification, some interpreted is as follows: "There could be some things which could be improved"
[I7] I forgot about troubles.	"I don't really have any troubles" "I couldn't forget about them because I didn't have any – so I could say never"
[P1] I felt it was not my choice to do the activity.	'not' was continuously skipped while reading aloud and interpreting the item
[I8] I forgot about my daily routine.	"I don't have a daily routine" "I don't have a daily routine – so I'll say all the time"
[I6] I forgot about my homework.	"we don't have any homework" "no, 'cause I don't have homework" "yes, because we don't really have homework in this school"

The items indicated by the interviews and the statistical analyses of the initial item pool were in approximate overlap, except for the items of the *Fear of damage* factor. None of those items appeared to be problematic during the interviews. Therefore, we can conclude that the items of the *Fear of Damage* factor are understandable and comprehensible to the respondents. Hence, the non-significant effect on *Experienced fun* is not due to the quality of the items.

According to the results of the think-aloud interviews, slight modifications were applied in general to the questionnaire and in specific to item E4. Namely, item E4 was slightly modified: from *I want to do the activity again* to *I want to do something like this again.* Additionally, since most of the students misread *the* activity as *this* activity, it was corrected accordingly in the whole questionnaire. Consequently, the questionnaire was used in this modified way during the next data collection and thus in the final validation step of the FunQ.

3.4.3 Psychometric Properties of FunQ

3.4.3.1 Data

Data were collected between 8 and 17 March 2019 during the British Science Week at the Science Museum. The reason for collecting the second data set in the UK was to collect responses from native English-speaking adolescents to ensure the quality of the instrument. The questionnaire responses were obtained from eight school classes who attained during this period the interactive Wonderlab: The Equinor Gallery program. In total, 150 responses were collected, however, the quality of 22 responses was questionable. They showed signs of typical response bias (straight lining; see section 3.2.4 Bias) or many responses were missing (e.g., the second half of the questionnaire) questioning the reliability of those responses. For the sake of data quality, the data of those 22 respondents were completely removed before the analysis started. The validation was conducted on the data of the remaining 128 respondents (64 boys, 45 girls, 19 not given, M_{age} =12.15, SD = 1.079).

The questionnaire was administered on paper after the students had participated in the activity. Consent was attained across the class teachers according to the British regulations. For the descriptive statistics and the test of normality see Appendix C.

3.4.3.2 *Missing Data*

The proportion of missing values for the whole sample is 2% and for the questionnaire items is 1.7%. Since the normality of the data cannot be assumed, we used the non-parametric test of homoscedasticity to check whether data is missing completely at random (MCAR). The test resulted in a nonsignificant p-value (p = 0.093), thus it was assumed that the values are missing completely at random. Hence, to handle missing data, full information maximum likelihood (FIML) estimation was used.

3.4.3.3 *The Model Fit of the 18-Item Model*

To assess the validity of the previously established final 18-item model, we fitted it to the second data set. Table 3.6 introduces the statistics of the 18-item model on the second dataset, Figure 3.2 depicts the model with the related factor loadings. The internal consistency of the majority of the remaining factors is above the cut-off value (ω > 0.6) suggesting that the items are measuring the same underlying constructs. Despite that the internal consistency of the factor *Challenge* (ω = 0.425) is below the cut-off value, it appears to have a strong significant effect (*std. factor loading* = 0.990, p = 0.002) on the *Experienced fun*, therefore keeping the factor in the model is justified. The standardized factor loadings of the first-order variables on the second-order variable *Experienced fun* of the final model on the validation data set are all above the cut-off value of 0.3, therefore they are considered substantial.

Table 3.6 Statistics of the 18-item model on the second data set. The internal consistency coefficients and the standardized factor loadings of the factors on *Experienced fun*.

Factor	Cronbach's alpha	Omega	Standardized factor loading on *Experienced fun*	P-value
Autonomy	0.578	0.653	0.734	0.006
Challenge	0.434	0.425	0.990	0.002
Delight	0.751	0.766	0.993	< 0.001
Immersion	0.640	0.608	0.885	< 0.001
Loss of Social Barriers	0.611	0.633	0.516	0.006
Stress (R)	0.790	0.791	0.649	< 0.001
Overall	0.843	0.875	-	-

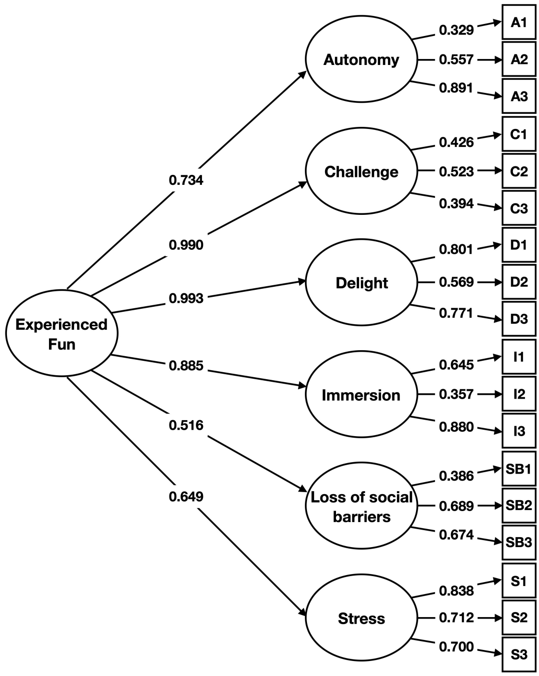


Figure 3.2 The second-order hierarchical model results of the final 18-item model on the second data set. Standardized factor loadings are shown. All of the related p-values are below the 0.05 margin.

The internal consistency of the second-order factor *Experienced fun* is presented in Table 3.7. The Omega values are found to be above the cut-off value (ω > 0.6). This finding suggests that the *Autonomy, Challenge, Delight, Immersion, Loss of Social Barriers* and *Stress* factors measure with high reliability the same underlying construct, the *Experienced fun.*

Table 3.7 Statistics of the 18-item model on the second data set. The internal consistency coefficients of *Experienced fun* as second-order factor.

	Omega at level 1†	Omega at level 2‡	Partial Omega§
Experienced fun	0.794	0.928	0.864

† the proportion of the second-order factor explaining the total score

‡ the proportion of the second-order factor explaining the variance at first-order factor level

§ the proportion of observed variance explained by the second-order factor after partialling the uniqueness from the first-order factor

The model fit indices of the 18-item model on the second data set are introduced in Table 3.8. Evaluating the model fit indices and considering the limitations of the χ^2 test, we can conclude that based on the RMSEA and SRMR values the model fit is sufficient.

Table 3.8 Fit indexes of the 18-item model on the second data set

	χ^2 value	CFI	RMSEA	SRMR
Reference value	-	> 0.95	< 0.06	< 0.08
18-item model	173.632; df = 129, p = 0.005	0.933	0.052	0.072

3.5 Discussion

In recent years it has become common practice to address fun as a *common sense* notion instead of precisely defining the meaning of the concept [127, 150, 307] making its measurement difficult. Therefore, after we defined our understanding on the notion of fun in the previous chapter (Chapter 2) based on theoretical grounds, the aim of this chapter was to create a tool for the multidimensional measurement of the experienced fun that is psychometrically and theoretically sound, comprehensive yet parsimonious, practical and child appropriate, specially developed for adolescents and that can be used in the learning environment across various fields of research. To this end we have adopted a deductive scale development approach, which is widely used in the field of industrial and organisational psychology [294]. Accordingly, the conceptualisation of the construct of fun was theory driven based on a thorough review of literature related to fun. We examined a network of related concepts contributing a theoretically founded conception of fun for our targeted demographic. We concluded that for adolescents to experience an activity as fun they need a) to feel in control of the activity and be intrinsically motivated for participation (*Autonomy*); b) to experience an optimal level of challenge matching their level of skills

(*Challenge*); c) to feel *well* during the activity (*Delight*) and d) to not feel *bad* (*Stress*, contra-indicative); e) to be immersed in the activity losing one's perception of time and space (*Immersion*) and f) to let go of social inhibitions (*Loss of Social Barriers*). The FunQ is put forward as a tool for testing how a learning activity maps on its different dimensions.

Our conception of fun was tested by the statistical analysis of the created instrument. The final model consisting of 18 items across six dimensions. Besides statistically testing the instrument, the comprehensibility and appropriateness for the youngest members of the target age group was checked by the think-aloud interviews, thus implying suitability for older teens as well, and the questionnaire was adapted accordingly. The final version of the FunQ has been shown to have reliable internal consistency both at the first- and second-order level. Since the two data sets (for testing the initial item pool and for validating the final 18-item version) were collected at two different countries (the Netherlands and the UK), from eleven groups of adolescents who participated in three different kinds of learning activities, it is assumed that the revealed model is not activity specific. Additionally, given that the FunQ items are phrased in a general way, we anticipate that the instrument will be applicable in a broad range of different contexts to assess the experienced fun of an activity among adolescents.

The data analysis confirms our initial expectations that fun is a multi-dimensional construct. Among the dimensions of fun examined it seems that the *Fear of Damage* has no significant effect on whether adolescents experience an activity as fun, however, the existence of the rest of the proposed factors is confirmed with the note that the *Autonomy* and *Pressure* factors were merged as they appeared to measure the two extremes of the same dimension.

With the largest standardized factor loading, the *Delight* factor has the greatest contribution to the *Experienced fun*. This factor focuses on the positive emotions and the related desires. It sounds natural that fun is a positive experience, and as such, it implies the desire for repetition [32, 98, 236]. This aspect is captured by the *Delight* factor, which our findings indicate as an organic part of the *Experienced fun*.

To maintain the engagement and therefore to stay in the activity while experiencing it continuously as fun, however, the optimal level of challenge is required. While previous research with children investigated challenge and fun as separate constructs [47, 53], our model considers that challenge is a facet of the experienced fun. This idea appears in measurement tools designed for adults [91, 225], however, for adolescents, the association has only been highlighted in relation with physical education activities by the qualitative study of Dismore and Bailey [72]. Our findings suggest that *Challenge* is the second most important factor of the *Experienced fun*, though it is left for future investigations to establish the suitability of challenge as a dimension of fun for children of different ages.

The *Immersion* factor measures the loss of time and space. When one is deeply engaged the immersion in the activity happens that leads to the loss of sense of time and space [61, 243, 293]. In the FunQ, this aspect is mapped by the *Immersion* factor which was found to have the third highest factor loading on the *Experienced fun.*

The *Autonomy* factor bears the fourth highest standardized loading, and it assesses whether the child feels control over their participation as well as the activity itself. As it was summarized above, fun is a voluntary experience [32] therefore intrinsic motivation is seen a key factor for participation [186]. Additionally, applying the Flow theory [61] it was expected that feeling in control over the situation is related to the motivation as well. Our finding supports this theory as the *Autonomy* factor that refers to the experienced control over the participation and the activity itself appeared to have a significant effect on the *Experienced fun.*

Compared to the research instruments currently used to measure fun, e.g., in the context of evaluations of interactive systems and educational (serious) games, this study also includes contra-indicative items and factors to the construct of *Experienced fun* to enhance the validity of the tool but also to allow the assessment whether the activity that is intended to be fun causes unintentionally any distress to the participants. The antagonism between stress and fun has previously been taken as obvious. Caine and Caine [47] mentioned the stress-reducing effect of fun, however, without statistically testing it. Our findings provide supportive evidence that negative emotions are contra-indicative for experiencing fun as the effect of *Stress* factor was found to be significant.

According to Flow theory [61] immersion should result in social barriers to be largely removed. That is, while experiencing fun and immersing in the activity, the suspension of reality is triggered, which, in turn, leads to loss of self-consciousness. Once the person is less self-conscious, they are becoming less engaged with themself, is less afraid of rejection, and more open for others, which ultimately results in the breakdown of social barriers. Bisson and Luckner [32] indicated that in the case of children the combination of fun and play could act as a catalyst to eliminate inhibitions inherent to our socialisation. Our findings support this theory as the *Loss of Social Barriers* factor had a significant contribution to the *Experienced fun.* In other words, while having fun, children could connect to each other easier than usual.

Regarding the psychometric properties, the internal consistency measures (Cronbach's alpha and Omega) for the second-order factor *Experienced fun* provide evidence that the questionnaire measures reliably the underlying construct. And the model fit indices suggest a sufficient model fit. It is therefore proposed that the FunQ is suitable and valid to measure the experienced fun with adolescents. However, the role of challenge on the experienced fun among adolescents is proposed to be further investigated, especially as the internal consistency of the *Challenge* factor did not meet the criterion level ($\omega > 0.6$).

Since in working with children and adolescents it is preferable to address them in their mother tongue, or at least in a language they are comfortable with, on top of the original English version a Dutch adaptation of the FunQ has been created. Developing measurement instruments for different languages is not purely a matter of translation, as many concepts do not translate well, while the translated items should reflect coherently the underlying theories. For this reason, it was important to carefully assess the extent to which the adapted instrument measures the same concept as the original one. For the adaptation, we followed the cultural adaptation protocol of self-report measures [26], hence we have undertaken the following steps: i) forward translation from English to Dutch, ii) backward translation from Dutch to English, iii) comparison of the original and the backward translated English text, assessing discrepancies, iv) testing the questionnaire. During phase iii), we ensured semantic-, idiomatic-, experiential-, and conceptual equivalence between the source and the target version. For the details about the Dutch adaptation of FunQ see the work of Tisza, Gollerizo and Markopoulos [298]. The Dutch version of FunQ can be found in Appendix D.

Comparing FunQ to other instruments, FunQ covers a similar ground to the This or That [340], the fun Semantic Differential Scales [338], the fun Toolkit [234] and the Five Degrees of Happiness [108] instruments, however, FunQ is a theoretically founded instrument, which handles fun as a multidimensional construct instead of being unidimensional, and it is designed for-, and validated with adolescents instead of young children. Regarding the PENS [259], the UES, and UES-SF [204], the GEQ [225], the Playful-consumption experience questionnaire [1], and the EGameFlow [91] scales and the list of Likert scales for the evaluation of a game and attitudes towards learning games [127], they all designed for the gaming environment, hence, their usability is limited in the learning environment for which FunQ has been created. Additionally, the aforementioned scales mainly target adults and mostly focus on the evaluation of a product or game, in comparison with FunQ, which is designed for adolescents and focuses on the personal experience while being engaged with a learning activity. The FUN scale [293] is validated to measure the fun value of a touristic destination as a product while the Physical Activity Enjoyment Scale [145], as its name suggests, evaluates a physical activity, thus they both target a different field than FunQ. Comparing FunQ to the EmoForm [344], they are both designed for adolescents and for the learning environment, but the former focuses on the experienced fun as a multidimensional construct, while the latter investigates a broader range of emotions, handles enjoyment unidimensionally, and has not been validated yet. Therefore, we conclude that FunQ is a much-needed instrument, which measured fun as a multidimensional construct covering playful learning activities involving adolescents.

3.6 Limitations and Future Work

The herein introduced study is limited to the general population of adolescents in the learning environment. Additionally, as mentioned above, particular attention should be paid to investigate further the role of challenge on the experienced fun for different ages and settings. To further expand the potential of the questionnaire, follow up studies shall investigate the psychometric properties of the questionnaire for different ages, and examine its scope of application: whether the FunQ can be applied to evaluate fun not only in relation to learning, but in other activities in which fun can play a useful role, such as participation in experimental studies, child-computer interaction, playful activities and experiences.

3.7 Conclusion

This chapter aimed to empirically validate the previously established definition of fun (Chapter 2), and hence, to provide a reliable measurement tool for the assessment of the fun experience. To this end, we followed a deductive scale development approach, and validated FunQ in four consecutive steps, involving three studies. This chapter contributes a) a multi-dimensional instrument – FunQ - for assessing the experienced fun, which targets specifically adolescents both in the design, the content, the response format, and b) a psychometric evaluation and validation of the proposed instrument. We conclude that FunQ is a reliable, and much needed addition to the current palette of available measurement tools for the assessment of fun.

Part III.

THE ROLE OF FUN IN LEARNING

4 Fun in Coding – a Case Study[9]

In Part II of this thesis, we established our definition for fun, which was then empirically tested and validated. In this part we introduce two case studies (Chapter 4 and 5) in which we investigate the relationship between fun and learning using FunQ. These case studies lead to the initial model on the role of fun in learning (Chapter 6). The current chapter contributes by extending our understanding on the relationship between fun and learning.

Summary

There is a worldwide pursuit to increase children's interest in STEM (Science, Technology, Engineering, Mathematics) especially in computer science through extra-curricular activities such as coding workshops, hackathons, and FabLab initiatives. However, the underlying reasons for children's willingness for participation in such activities, and the effect of participation on children's topic-related knowledge are still not well understood. To understand the factors influencing children's attitude towards programming and to investigate what affects children's learning during such activities, we designed a workshop for introducing primary school students to programming and implemented it for a Dutch primary school class as an exploratory case study. The workshop was held during school hours but as an extracurricular activity. We recorded students' attitudes towards programming, their state-level emotions, the fun they experienced, and the initial- and final knowledge on the topic. Our findings indicate that the coding workshop had a positive effect on students' state-level emotions, as they felt significantly happier, more excited, and more in control at the end of the workshop than at the beginning of it. We also found that students' attitude toward programming changed significantly and positively during the workshop, and that students' attitude about programming is influenced by the experienced fun while learning to code regardless their gender. Additionally, we found that the workshop was successful in terms of knowledge acquisition: both the measured and the reported learning indicate that students learned during the activity. Our findings also indicate that students reported learning has a positive association with their state-level emotion feeling *in control* and that the measured learning is negatively influenced by high levels of stress. Accordingly, our results draw attention to the downshifting effect of high arousal emotions on the measured learning. Throughout the chapter we discuss gender differences along the study findings and elaborate on further practical implications.

4.1 Introduction

Coding or programming is often seen as an excellent way to nurture 21st century skills, and coding is widely considered as the literacy skill of the 21st century [210], however, the underlying reasons for children's and adolescents' willingness for participation in such

[9] This chapter is based on the following publication: Tisza, G., Markopoulos, P., & Bekker, T. (2022) Learning to code: interplay of attitude, emotions, and fun. Manuscript submitted for publication.

activities is still not well understood. Emotions may be a key factor of participation in coding activities. Accordingly, there is a rise in research to better understand the role of emotions in technology-based learning environments [168], however Graesser argued that *"the fields of psychology, education, and computer science have individually been much too slow in investigating the intersection of emotions, learning, and technology until the last two decades"* ([101], p. 1). While previous research indicated the importance of motivation and attitudes on the willingness to participate in coding activities, there is no consensus as to which emotions are experienced and what role they play in a technology-based learning environment [101]. To complicate things, assessments obtained with different tools, such as self-report measures, judge's reports, physiological data, behavioural data, and non-instructive multichannel sensing measures, are little to moderately correlated [101].

Along with emotions and attitudes, gender might play an important role on children's and adolescents' participation in coding activities. It is well known that girls (and females in general) are underrepresented in STEM fields (Science, Technology, Engineering and Mathematics) [178] and so there is a world-wide pursuit to increase the involvement of girls in science. However, based on a recent report of Girls Who Code [97], the applied policies[10] to increase access to Computer Science (CS), and to increase the volume of CS classrooms in the United States are currently unable to increase the participation of girls in programming: *"The data shows that existing policies to bring more girls into computer science aren't just missing the mark, they may actually be doing more harm than good"* ([97], p. 11). Based on the report, girls participation rate in CS classes from school years 2016-2017 to 2017-2018 in states with access policies decreased the most in Arkansas with 4.1%, and increased the most in Utah with 3.4%, and it was overall way below 50%. European researchers also stated that at the moment researchers are lacking of evidence for designing effective and engaging coding experiences for children [210] and that gender differences are relatively understudied in coding and making activities [209]. The same holds true for the reasons underlying these observations.

The herein introduced study aims to broaden our knowledge on the possible factors that influence students' attitudes toward coding, hence, indirectly affect their willingness to take part in coding activities. This chapter is set to investigate not only students' attitude but also their state-level emotions (emotions at a given moment) and possible interactions between those and the reported and measured learning while investigating gender differences as well. In the context of a two-hour long coding workshop we examined students' emotions, attitude towards programming, the fun experienced, and their learning. Our findings indicate that the workshop had a significant and positive effect on students' emotions and attitude towards programming, and that the workshop was successful in terms of learning. Additionally, we provide evidence on the possible influential factors

[10] "Girls Who Code qualifies policies designed to increase access to computer science as those that fall within at least one of the three categories: Computer Science Standards, funding, or High School Computer Science Offering." ([97], p. 6)

associated with the attitude change and learning gain and discuss gender differences along the study findings.

The current study brings novel insights on the influential factors that play a key role on attitudes toward programming and knowledge acquisition during playful coding learning experiences, and the role of experiencing fun in this setting. In the following sections of the chapter we examine related findings and theories, before we introduce and discuss our study findings, and elaborate on future research.

4.2 Background

4.2.1 Emotions and Learning

As Hascher stated, *"there is rarely any learning process without emotions. (...) Despite the obvious connection between learning and emotion, still very little is known about it. For decades, learning was mainly analyzed in terms of cognitive or motivational aspects. As a consequence, learning theories ignored affective processes for a long period of time"* ([112], p. 13). This is in line with Graesser ([101]), who claimed that researchers started to focus on the relationship between emotions, learning and technology only in the past two decades. Moreover, Pekrun et al. [215] adds that scientific research on academic emotions had a strong and narrow focus on test anxiety for decades. A recent systematic review [168] investigating emotions in the technology-based learning environment supported Pekrun's argument as the study found that anxiety was the most frequently investigated academic emotion (which was investigated in approximately the half of the cases); followed by enjoyment, which was investigated in approximately in one fourth of all reviewed cases, and boredom (approx. in 15% of the cases). They compared additionally the reported emotions in the technology-based environment and non-technology-based environments and found slightly less anxiety and more enjoyment in the technology-based environment. However, the effect varied across different learning settings. They added that their study findings indicated possible nonlinear relationship between emotions and other factors, which has to be further investigated. Accordingly, Hascher [112] urged future research in the field in order to broaden our knowledge and understanding on a wider spectrum of emotions in the academic environment.

From cognitive psychology we know that emotions influence cognitive processes and strategies, decision making and motivation, and that the aforementioned influences are reciprocal [147]. In psychology research, the basic emotions that are characterized by prototypical facial expressions are happiness, sadness, surprise, disgust, anger and fear [75]. However, it has been questioned whether these six emotions play a key role in the learning process [60]. Graesser clearly stated that *"the ensemble of emotions that occur during learning* [i.e., boredom, confusion, frustration, curiosity, enjoyment and anxiety] *are very different from the basic emotions* [i.e., happiness, sadness, surprise, disgust, anger and fear] *that dominated psychological research for decades (...)* [as] *most of the basic emotions* *are not prevalent in learners and predictive of learning in contemporary learning environments"* and that *"the profile of emotions that learners experience have some commonalities but also predictable differences over task, goals, subject matter content, and population of learners"* ([101], p. 2). As no scientific consensus exists on the key emotions in learning, the different effects of positive compared to negative emotions while learning and on learning are not straightforward, hence, any assumptions of a straightforward relationship between positive emotions and learning, and negative emotions and not learning would be misplaced [60, 168].

Mayer [180] pointed out three main research challenges when investigating emotions and learning. The first main challenge is the identification of the key emotions in e-learning, the second challenge is their appropriate measurement, and the third is the explanation of the findings, with special regards to the causes and the consequences of emotional states while learning.

4.2.2 Fun and Learning

Bisson and Luckner [32] were among the first to discuss the positive effects of fun in the learning environment. In their view fun functions as a vehicle for evoking intrinsic motivation, reducing stress and social boundaries, and creating a safe learning environment. Other authors argued that fun facilitates engagement [232], enhances learning [51, 171, 232, 295, 317, 327], improves programming skills [169], has a significant effect on the learning effort [169], fosters curiosity [127, 317], contributes to high-quality learning experience [295], promotes collaborative learning [51], is a predictor for learning success [127], and has an effect on gaining motivation [127]. Regarding Elton-Chalcraft and Mills *"learning which is enjoyable (fun) and self-motivating is more effective than sterile (boring) solely teacher-directed learning"* ([78], p. 482). However, other studies failed to demonstrate significant positive associations between the experienced fun and the learning outcomes [127, 275]. As Hascher stated, *"enjoyment in school is one of six constitutive dimensions of student well-being (...) so far, academic enjoyment has been investigated in terms of different events of enjoyment or as enjoyment in specific subjects. Rarely, enjoyment was addressed to the learning activity itself"* ([112], p. 21-22).

According to Nandi and Mandernach *"there is tremendous amount of interest in making education more engaging and interesting for students"* ([198], p. 346), and within the informal (STEM) learning context, game jams, hackathons and game creation events *"have been acknowledged by academics and policy makers as a viable alternative to traditional approaches"* ([88], p. 38). Game jams *"provide participants with the opportunity to create a game within a specific constrain or limitation (time, technology, theme, or mode of transport)"* ([88], p. 39). According to Fowler [88], over one-third of the participants report on willing to attend the next game jam for fun, while approximately one-fourth willing to attend to learn new things, including new skills, too. Hackathon are *"events that have been described as a problem-focused computer programming event"* ([88], p. 39), however, compared with the game jams, hackathons do not necessarily have the focus on game development. According to Nandi and Mandernach [198], hackathons provide students with a fun and engaging format to learn about programming in the non-formal learning environment. They also reported on an observation that students who participated on hackathons had slightly higher GPAs (grade point averages) compared with non-participating students. While investigating reasons for participating in hackathons, learning and skill improvement were among the most frequently mentioned ones, and among the other reasons having fun was often found [43, 88, 158].

In sum, while some studies discuss the coding activity in terms of students having fun, and that the coding activity increases students' attitude toward coding and their learning outcomes [210, 260] and possibly contribute to a higher GPA [198], based on our best knowledge, the direct link between the experienced fun and the attitude change, and the effect of fun on learning has not been scientifically studied yet.

4.2.3 Gender Differences in Coding: Attitudes and Skills

Master et al. stated that *"the gender gap in science, technology, engineering, and math (STEM) engagement is large and persistent"* ([178], p. 92). It is known that students' success in computer science courses and their career choices could be affected by their attitudes toward programming [50] and previous findings indicated more positive attitudes toward programming among boys than girls [25, 153, 197, 253]. Gender differences on beliefs about programming are present as early as in the age of six [178], however, those beliefs can be influenced by experience. While findings suggested that girls attitudes could be influenced and the change could be measured immediately, such a change cannot be assumed sufficient for changing girls' stereotypes about programming or robotics, as *"changing stereotypes is difficult, even among children"* ([178], p. 101). Yucel and Rızvanoğlu [337] found gender differences in all of the nine attributes they examined: perceived competence, perceived coding difficulty, identification, perceived game difficulty, perceived success, level of enjoyment, level of anxiety, likelihood of playing another time and likelihood of trying new features. Overall, girls found Code Combat - a code learning game - and programming less attractive than boys did. This aligns with Master et al. [178] where after a short intervention which generated a positive experience, the technology motivation of girls and boys were statistically at the same level, while in the control group (without the positive robotic experience) the technology motivation of girls was especially low compared to that of the boys. However, other research indicated, that interacting with visual programming environments (such as Scratch) had a positive effect on children's attitude toward programming [104, 260], and after the interaction no significant gender difference could be found in children's attitude toward programming [104, 139, 140, 349]. Gunbatar and Karalar conclude that *"when teaching programming through visual programming environments, the gap between gender differences can be closed in terms of many variables"* ([104], p. 931). Hackathons might also serve as an event to make programming more attractive for young females. In their study, Ruiz-Garcia et al. [254] designed a hackathon specially for girls, including a mentorship program. They found that 100 from the 111 participating girls were absolute novices to hackathons, and their results indicated that the participating girls *"will continue exploring on their own the technologies they learned, as well as explaining them to their friends. The most successful feedback is that they are now more interested in studying engineering degrees"* ([254], p. 255). Another study [242] reported on a hardware hackathon, in which they used the LillyPad Arduino to design wearables, this way aiming to attract more female participants. They concluded that their specific focus on wearable design was successful in diversifying participation, in other words, in attracting more than usual females to the hackathon.

A recent study [197] investigating Canadian children's attitude toward coding found that 72% of the surveyed boys, and 57% of the surveyed girls were very- or extremely interested in careers that involve use of digital technologies. However, 50% of the boys and 27% of the girls said to be very- or extremely interested in having a career that involves coding or programming. Regarding children's attitudes about programming the research findings indicated that *"boys were 15 percentage points more likely than girls to describe coding as interesting; 13 percentage points more likely to describe it as cool; and 14 percentage points more likely to describe it as important. Girls were 14 percentage points more likely than boys to characterize coding as difficult"* ([197], p. 9). They found not only a difference in the attitudes toward programming, but in the programming-related self-confidence as well: *"while 41 per cent of boys say that they are somewhat or totally confident in their coding and programming abilities, only 28 per cent of girls exhibits these levels of confidence"* ([197], p. 10). They also noted that *"boys' higher self-reported confidence in their coding abilities is not necessarily evidence that they are more skilled than girls"* ([197], p. 11). This is supported by the findings by Papavlasopoulou et al. [209] who concluded that girls do not lack in related skills and competences compared to boys.

Regarding the learning outcomes while learning to code with Scratch, studies suggested no gender difference [209, 283]. On the other hand, the use of other methods based on physical computing principles, decreased the gender difference in the learning outcomes [253]. Papavlasopoulou, Sharma and Giannakos reported that *"children with higher levels of excitement had the same characteristics as those who reported high learning"* ([210], p. 57), indicating that the reported learning scores might be biased by the experienced level of excitement. Another study using eye-tracking to assess engagement found that *"children's level of engagement during coding activities moderates the relationships between their intention to participate in the activity and [perceived] learning"* and *"children's level of engagement during coding activities moderates the relationship between their intention to participate in the activity and enjoyment"* ([273], p. 71).

Concerning emotions and gender differences in the technology-based learning environments Loderer et al. [168] found weak relationships between both positive and negative emotions and gender. Regarding designing gender specific coding activities

Master et al. [178] stated that they may backlash with unintended consequences as dividing children by gender can lead to increased stereotyping and making STEM superficially appealing for girls can lead to later disappointments. Additionally, they noted that huge individual gender differences exist among boys and girls, hence there are less technology inclined boys, and more technology inclined among girls as well. Yücel and Rızvanoğlu [337] also emphasised the importance of developing genderless or gender-neutral activities and code-learning environments for children.

In sum, previous findings indicate a gender difference in attitude towards-, and participation in coding activities while scientific evidence supports that boys and girls are cognitively equally skilled. It has also been shown that with positive interventions children's and adolescents' attitude can be positively shaped. However, the current state of the art is equivocal as to whether designing gender-specific coding activities are necessary or useful for increasing girls' participation in coding, neither is there a clear view on what factors influence children's and adolescents' attitudes toward programming and knowledge acquisition.

4.2.4 Research Aim

In the remainder of this chapter we present an exploratory case study in which we set out to investigate what factors influence students' attitude towards programming and their knowledge acquisition while learning to code, with special regards to the role of fun and state-level emotions such as happiness, excitement and control play, taking into account possible gender differences. Accordingly, we formulated the following research questions:

- How does the workshop influence students' emotional state (happiness, excitement, and control)?
- What factors influence students' attitude about programming?
- What factors influence students' reported and measured learning?
- What is the role of the experienced fun on students' attitude about programming, the reported learning and the measured learning?
- Is there any gender difference present in the investigated relationships?

4.3 Method

4.3.1 The Activity

We designed a playful coding workshop in collaboration with SkillsDojo, an open-source company that develops and disseminates technologies and applications for children aged 6-14. The workshop was a non-formal activity, building on participants' intrinsic motivation for participation, and applied a learning-by-doing approach [142], hence, playfulness was by nature inherent to it. Furthermore, participating students were invited to follow a video guide, in which playfulness was reflected in the tone and introduction of the task and the visual design. Additionally, the workshop also evoked students' creativity by encouraging them to use their own ideas to solve the tasks. For the workshop we used three of the SkillsDojo videos (www.kidzcourse.com/workshop) to introduce coding with micro:bits[11] to participating students in a fun and playful way. The first video introduces the basics by teaching students how to display their names on the LED panel of the micro:bit. The second video shows students how to make a rock-paper-scissors game from the micro:bits. Finally, the third video shows students how to create their own micropet which reacts to kinetic stimuli. When selecting the videos we considered the followings:

- Suitable for novices (the first video introduces the micro:bit and basic programming terms).
- Difficulty increases gradually (the second video is more complex than the first, and the third one is more challenging than the second).
- Equally suitable for boys and girls (we speculated that there is no gender difference in the liking of the stone-paper-scissors game, and that making a micropet is interesting for both boys and girls - especially given that participating students could develop their own design besides the pre-printed monkey, cat and bunny templates).

The workshop was designed as a single-occasion, two-hour long activity with the following structure:

1. Introduction of the topic and the structure of the workshop (~ 5 min)
2. Pre-activity data collection (~ 10 min)
3. Creative coding with micro:bits supported by the three videos (~ 90 min)
4. Post-activity data collection (~ 10 min)

After the introduction the researcher handed out the pre-workshop questionnaire to the participating students. Each child was equipped with a Chromebook and a micro:bit. Once the questionnaires were collected, students were asked to explore the micro:bits, assemble and plug them in the Chromebooks. Then, the first video was played for all on the whiteboard and was paused according to the instructions so that each child could understand the procedure and the way the videos work. After watching the first video together, students were asked to follow the second video at their own speed, each on their own Chromebook. They could work alone or together with their classmates, depending on their own preference. The teacher and the researcher were walking around in the class, helping students and facilitating interaction among classmates. Help was mainly asked when students encountered difficulties for example when their code was not working, so the researcher helped them debug their code; or when they needed technical assistance

[11] Micro:bits are pocket-sized, programmable microcomputers with a display of 25 LED lights, two programmable buttons, an accelerometer and a magnetometer sensors and a Bluetooth and USB connectivity. It can be programmed in several programming languages, including Python, Matlab and C++.

(e.g., how to save the code on the micro:bit). After the second video, students could follow the third video or they could create their own code. For making the body of the micropets students were allowed to use color pencils, glue, scissors and (pre-printed) paper. When the time was up, students were asked to tidy up their table and the post-workshop questionnaire was handed to them.

4.3.2 Participants

For the workshop and hence for participation in this study, Dutch teachers could sign up their classes. Despite that the workshop was held in a classroom setting during school hours, students' participation was voluntary. Accordingly, informed consent was obtained from both the students and their parents/caretakers. The herein described study was conducted in June 2019 in a Dutch primary school with a group of 23 students between age 10 and 12 (M_{age} = 10.96, *SD* = 0.767; 10 boys, 13 girls).

Prior to the workshop we asked students about their previous experience with coding across two 5-step Likert-type questions: 'Do you have any idea about programming?' (1-Not at all; 5-I am a pro) and 'How many coding workshops have you participated before?' (1-None; 5-Six or more). 26.1% (6 students) report on not having participated before in any coding-related workshop and 30.4% (7 students) report on having no previous knowledge on programming.

4.3.3 Measures

To address participating students' emotions, attitude, learning and the fun they have experienced during the workshop, we used a number of previously validated measurement tools (see summary at Table 4.1).

At the beginning and at the end of the workshop for the measurement of students' state-level emotions we used 5-step bi-polar scales (happy - unhappy, calm - excited, controlled - in control). These pairs were selected from the Semantic Differential Scale [185] and highly correlated with the dimensions of the Self-Assessment Manikin [42], which are both developed for the assessment of affective reactions. The internal consistency of the three emotional bipolar scales is acceptable (*Cronbach's* $\alpha > 0.6$ [107]; $\alpha_{pre\text{-}workshop} = 0.772$, $\alpha_{post\text{-}workshop} = 0.939$).

To assess students' attitude (i.e., feeling or opinion about something) towards programming we used a single item measure that addresses students' general attitude about the topic: 'Programming is my thing'. This item we have consciously selected from earlier research [300] based on its simplicity, general nature and validity provided by cross-validation. The general reliability of single-item measures in comparison with multiple item measures was proved by earlier research [28].

For the assessment of fun, at the end of the workshop we recorded FunQ [299]. The questionnaire is evaluated across eighteen 5-step Likert-type questions along six dimensions. The internal consistency of the scale in the current sample is acceptable (*Cronbach's* α = 0.784).

For the assessment of learning, we utilized two measures that address different levels of learning according to Bloom's taxonomy [34]. The self-reported measure (i.e., reported learning; linked to the *Evaluation* level) measures students' perceived learning. Considering that knowledge tests can never cover every single detail of a learning process, hence usually fail to capture *all* learning that has taken place, self-report measures can be a good indication for learning given that they provide the respondents the freedom to take aspects into account that have not been investigated by the knowledge test. Reported and measured learning have thus a complementary nature, as reported learning has the potential to capture learning that is not examined by the knowledge assessment test. For the measurement of the reported learning we used a single-item measure, adopted from earlier research [210, 300].

The knowledge assessment test (linked to the *Knowledge* level) addresses the factual knowledge students gained as a result of the workshop (see the knowledge test and its descriptive statistics in Appendix E and F). We administered the knowledge assessment test both before and after the workshop. The knowledge assessment test was developed by the researchers in agreement with SkillsDojo. It contained seven questions with four response options. Four out of the seven questions were about terms related to programming, which are introduced and explained during the videos (e.g., 'What/Who is a variable?'). Three questions were on programming scripts that are the foundation of the workshop and their way of working is explained thorough in the videos (see example Figure 4.1). Accordingly, the knowledge test aligns well with the learning objectives of the video content. The measured learning was calculated by subtracting the pre-workshop knowledge assessment scores form the post-workshop scores as suggested by previous research in the field [275]. Using difference scores is a commonly accepted way among educators for addressing learning gain, and its reliability has been proven by various authors previously [30, 231, 297, 348]. The internal consistency of both the pre- and post-workshop test is acceptable ($Cronbach's\ \alpha_{pre\text{-}workshop} = 0.708, Cronbach's\ \alpha_{post\text{-}workshop} = 0.818$).

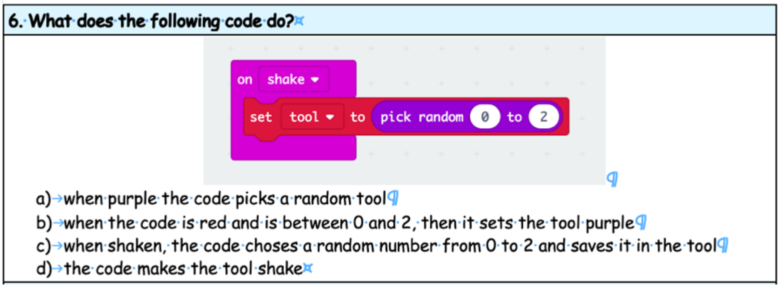


Figure 4.1 Example from the knowledge assessment test.

The completion time for both the pre- and post-workshop questionnaires was approximately 10 minutes.

Table 4.1 The investigated dimensions, their operational definitions, and their respective measures.

Dimension	Operational definition	Measure	Source
Attitude	The degree to which students indicate their attitude towards the subject.	'I think programming is my thing' (1) Not at all – (5) Absolutely	[300]
Emotions	The degree to which students indicate their state-level emotions.	Three five-step bi-polar scales: 'How do you feel now?' Happy - Unhappy; Calm - Excited; Controlled - In control	[185]
Fun	The degree to which students experienced fun during the activity.	FunQ	[299]
Reported learning	The degree to which students indicate their learning during the activity.	'Have you learned something new today about programming?' (1) Not at all – (5) A whole lot	[210]
Measured learning	The difference between the post-workshop and pre-workshop knowledge assessment test score.	Seven multiple choice questions addressing students' programming related knowledge (see Appendix E).	n.a.

4.3.4 Data analysis

For the data analysis SPSS Statistics version 25 software was used. To address the research questions, we applied correlation analysis, t-test, ANOVA and linear regression.

4.4 Results

4.4.1 Emotional State and Attitude Toward Programming

We aimed to assess whether the workshop affected students' attitudes towards programming. Therefore, we asked students to indicate on a 5-point smiley-face scale whether they think that programming was their thing at the beginning and at the end of the workshop. Paired sample t-test indicates that students' attitude toward programming changed significantly ($p = 0.012$, $t = 2.732$, *Cohen's d* = 0.570). In other words, students found programming at the beginning of the workshop less of *their thing* ($M = 3.39$, $SD = 1.118$) than at the end of the workshop ($M = 3.96$, $SD = 1.107$) (see Figure 4.2).

We also investigated whether students' state-level emotions changed in the course of the workshop. Hence, we asked students to indicate their emotional state on three bipolar scales at the beginning and at the end of the workshop. We found that students felt significantly happier ($p = 0.035$, $t = 2.297$, *Cohen's d* = 0.541), more excited ($p = 0.003$, $t = 3.543$, *Cohen's d* = 0.859) and more in control ($p = 0.041$, $t = 2.236$, *Cohen's d* = 0.559) at the end of the activity than at the beginning of it.

To assess whether is a gender difference in the above described tendencies, we tested the effect of gender applying repeated measures ANOVA. The results suggest no significant gender difference in the attitude change ($p = 0.253$, $F = 1.384$, *partial* $\eta^2 = 0.062$) or in the attitude scores ($p_{pre\text{-}workshop} = 0.123$, $t = 1.608$, *Cohen's d* = 0.669; $p_{post\text{-}workshop} = 0.597$, $t = 0.536$, *Cohen's d* = 0.226) and (see Figure 5.2). Additionally, we found no significant gender difference on the reported emotional states: neither happiness ($p = 0.755$, $F = 0.101$, *partial* $\eta^2 = 0.006$), nor excitement ($p = 0.381$, $F = 0.815$, *partial* $\eta^2 = 0.052$), nor feeling in control ($p = 0.200$, $F = 1.806$, *partial* $\eta^2 = 0.114$) appears to be gender dependent. Thus, the workshop had the same effect on both girls and boys.

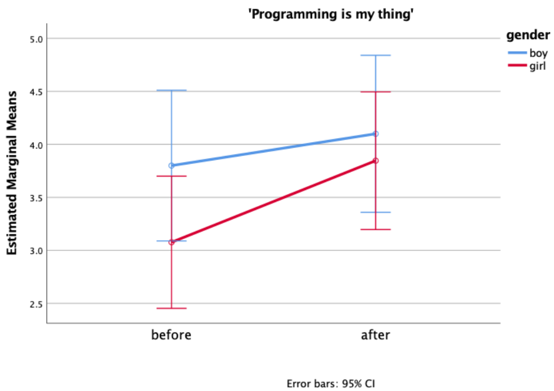


Figure 4.2 Attitude change: *'Programming is my thing'* scores before and after the workshop.

Regarding the participants' previous knowledge, we found that at the beginning of the workshop boys reported significantly higher values than girls for the question whether they have any idea about programming ($p = 0.005$, $F = 9.964$, $\eta^2 = 0.333$). This significant difference does not hold true for the number of coding workshops boys and girls had participated in ($p = 0.057$, $F = 4.124$, $\eta^2 = 0.186$).

4.4.2 Influential Factors on Children's Attitude About Programming

To start with, we applied correlation analysis to investigate the relationship between students' post-workshop attitude about programming and their pre-workshop attitude, pre- and post-workshop emotional states, the level of fun (FunQ), happiness (FunQ Delight) and stress (FunQ Stress) they have experienced during the workshop and their learning outcomes. The analysis reveals an association between students' post-workshop attitude and the pre-workshop attitude score (r = 0.602, p = 0.002), the pre-workshop emotional state *excited* (r = 0.501, p = 0.025), the *Delight* dimension score of FunQ (r = 0.436, p = 0.048), the *Stress* dimension score of FunQ (r = -0.478, p = 0.038), and the *reported-* (r = 0.524, p = 0.012) and *measured learning* (r = 0.432, p = 0.040) scores.

To determine the direction of the relationships found by the correlation analysis, we applied a stepwise regression analysis to model students' post-workshop attitude scores. We added the pre-workshop attitude, pre- and post-workshop emotional states, the level of fun (FunQ), happiness (FunQ Delight) and stress (FunQ Stress) they have experienced during the workshop and their learning outcomes as possible predictors. The analysis resulted in three consecutive, nested models. The final model, *Model C* explains the 87.2% of the variance (R^2 = 0.872). The significant predictors in the model are the measured learning (p = 0.001, t = 4.877, β_{std} = 0.618), the pre-workshop attitude score (p = 0.006, t = 3.661, β_{std} = 0.482) and the perceived learning (p = 0.017, t = 2.995, β_{std} = 0.394)

4.4.3 Learning Outcomes

To assess whether students learned during the workshop we used a knowledge assessment test and a self-report measure. In this section we introduce the study findings in relation to learning.

4.4.3.1 Reported Learning

To assess the perceived learning, at the end of the workshop we asked students to indicate on a 5 step Likert-type scale whether they learned something new during the workshop (see Table 4.2). On average, students reported to have learned much (M = 4.05, SD = 0.899). Independent sample t-test indicates no significant gender difference in the *reported learning* scores (p = 0.510, t = -0.671, *Cohen's* d = -0.291). The results indicate that we did not encounter a ceiling effect.

Table 4.2 'Have you learned something new today about programming?' response rates. 1 response (4.2%) is missing.

Not at all	A bit	Something	Much	A whole lot
0	4.3% (1 student)	21.7% (5 students)	34.8% (8 students)	34.8% (8 students)

4.4.3.2 Measured Learning

We compared how students scored on the knowledge assessment test at the beginning and at the end of the workshop to assess whether they gained factual knowledge. The average score on the pre-workshop test is 4.09 (SD = 1.98), while the average score on the post-workshop test is 4.96 (SD = 2.18). We conclude that we did not encounter a ceiling effect. Paired sample t-test shows that in general, students scored significantly higher on the post-workshop knowledge assessment test on the pre-workshop test (p = 0.016, t = -2.600, *Cohen's d* = -0.542). We found no significant difference between genders (p = 0.667, t = -0.436, *Cohen's d* = -0.184). Both the pre- and post-workshop knowledge test scores were higher for boys, but not significantly ($p_{pre-workshop}$ = 0.285, t = 1.098, *Cohen's d* = 0.462; $p_{post-workshop}$ = 0.521, t = 0.653, *Cohen's d* = 0.275; see Figure 4.3).

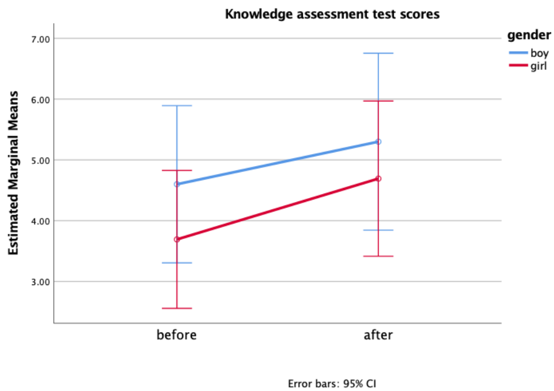


Figure 4.3 Knowledge assessment test scores before and after the workshop.

4.4.4 Influential Factors on Children's Learning

In order to assess how different factors influence learning, we applied linear regression analysis with stepwise selection, using both the reported- and the measured learning as outcome variables. As predictor variables we used the pre-workshop programming experience, the pre- and post-workshop attitude, pre- and post-workshop emotional states, and the level of fun (FunQ), happiness (FunQ Delight) and stress (FunQ Stress) students have experienced during the workshop.

4.4.4.1 Reported Learning

When modelling the possible influential factors of the reported learning (i.e., 'Have you learned something new today about programming?') the regression analysis finds that the post-workshop *in control* emotional state is the only significant predictor (p = 0.024, t = 32.714 β_{std} = 0.671), explaining 45.0% of the variance of the measured learning scores (R^2 =

0.450). From these results we conclude that students with high *reported*, thus perceived learning were the ones who felt in control at the end of the workshop.

4.4.4.2 Measured learning

When modelling the possible influencing factors of the *measured learning* (pre-workshop learning assessment score subtracted from post-workshop score) the regression analysis results in a single significant predictor. The *FunQ Stress* score explains the 63.9% of the variance of the measured learning scores (R^2 = 0.639; p = 0.003, t = -3.993, β_{std} = -0.799). Explaining the findings we conclude that students with high learning gain were the ones who experienced low levels of stress during the workshop.

4.4.5 Fun

For assessing the fun value of the workshop, we recorded the FunQ [299] with the participating students at the end of the workshop. For the statistical testing we reverse coded the *Stress* dimension, summed the scores of each dimension, and we calculated the grand total FunQ score as well (possible minimum score is 18 and possible maximum score is 90). The calculated grand total FunQ score ranges between 49 and 84 (M = 70.06, SD= 10.18) of which we can conclude that approximately covers the higher half of the possible score range.

Regarding the average scores on the separate dimensions (possible minimum score is 3 and possible maximum score is 15 on each dimension), we can conclude that the *Stress* factor (negative emotions) has the lowest average score (M = 4.11, SD = 2.35) while the *Delight* factor (positive emotions) has the highest (M = 13.14 SD = 1.90). Based on the separate dimension scores and the grand FunQ score we conclude that the workshop was stressful for students, it evoked positive emotions and students experienced it as fun. This finding is further supported by the spontaneous positive feedback by the teacher the day after the workshop: *"This morning I asked my class about yesterday's lesson. All of them were very enthusiastic. Group 8 said that they will be jealous if we could do this lesson again next year!!"*

We investigated whether is a gender difference in the grand total FunQ scores between boys and girls. The results of the independent sample t-test suggest that girls and boys experienced the workshop equally fun as no significant difference found between them (p = 0.932, t = 0.086, *Cohen's d* = 0.047). Furthermore, there was no significant difference found between boys and girls for the separate dimension scores of the FunQ ($p_{autonomy}$ = 0.645, t = 0.468, *Cohen's d* = 0.214; $p_{challenge}$ = 0.950, t = -0.063, *Cohen's d* = -0.028; $p_{delight}$ = 0.461, t = -0.752, *Cohen's d* =-0.332; $p_{immersion}$ = 0.955, t = -0.057, *Cohen's d* = -0.026; $p_{lossofsocialbarriers}$ = 0.190, t = -1.364, *Cohen's d* = -0.634; p_{stress} = 0.807, t = 0.248, *Cohen's d* = 0.118).

For modelling fun - as measured by the FunQ sum score - by linear regression we used the pre-workshop programming experience, the pre- and post-workshop attitude, the pre-and post-workshop emotional states, and the reported and measured learning scores as possible predictors. The analysis resulted in two nested models. The more complex model, model B explains the 78.1% of the variance ($R^2 = 0.781$) and has the *number of workshops* the child previously participated in ($p = 0.001$, $t = 5.317$, $\beta_{std} = 1.058$) and the pre-workshop *attitude* score ($p = 0.036$, $t = -2.524$ $\beta_{std} = -0.502$) as significant predictors. In other words, students' positive attitude at the beginning of the workshop had a negative effect on the experienced fun during the workshop, while the previous experience with coding (measured by the number of workshop students participated before) had a positive effect on the experienced fun. This previous finding, we speculate, could be due to expectation-management, but we propose further examination.

4.5 Discussion

While children's and adolescents', and especially girls', engagement in STEM fields and computer science has been in researchers' focus in the past decade, previous research indicated that the applied policies fail to increase girls' participation [197]. Among the underlying reasons the importance of emotions has been argued, but there is as yet no scientific consensus on which emotions and in what ways play a key role in the technology-based learning environment [101]. As indicated by Mayer [180], there is a need for broadening our knowledge on emotions that play a key role on learning and for the understanding the causes and consequences of those. While the effect of some emotions in the academic environment has widely been studied in the last decades, our understanding on the key influential factors on the willingness to learn programming is more limited.

The herein introduced research aimed to expand our knowledge on the possible factors that potentially influence students' attitude toward coding, hence, indirectly affect their willingness for participation in coding activities. Along with students' programming-related attitude we investigated their state-level emotions and the experienced fun, and possible interactions between those and the reported and measured learning. For this purpose, we collected data from Dutch primary school students before and after participating in a playful coding workshop. The results showed that students' attitude and state-level emotions positively changed during the workshop, that students' attitude is greatly influenced by the experienced stress and fun, and that the state-level emotions and the experienced stress play a key role on the measured learning accordingly.

Evaluating the workshop in general, we conclude that students found it fun and it had a positive influence on their emotional state. Children felt happier, more excited and more in control at the end of the workshop than at the beginning of it regardless their gender. A noteworthy finding is that at the beginning of the workshop boys reported higher values for the question whether they have any idea about programming than girls. While we cannot verify the validity of these claims, we emphasize that boys programming-related self-confidence can play a role just as found at Canadian children [197]. Such an explanation is supported by the fact that boys' attitude about programming was more positive at the beginning of the workshop than that of the girls. Nevertheless, participating students' attitude about programming – regardless of their gender – changed significantly and positively during the workshop. This finding aligns with previous work indicating that interacting with visual programming environments influences positively children's attitude toward programming [104, 260] and that after the interaction no significant gender difference is present in children's attitude toward programming [104, 139, 140, 349]. However, previous studies did not examine possible underlying effects.

Addressing the research question on possible influential factors on students' attitude about programming, our research indicates that experiencing excitement at the beginning of the workshop and having a sense of learning during the workshop has a positive impact on students' attitude about programming let them be boys or girls. However, we also found that the experienced fun was affected by students' initial attitude about programming and the number of coding activities they have participated previously. This reflects the reciprocal relationship - known from cognitive psychology [147] - between emotions, cognitive processes and strategies, decision making and motivation.

Concluding the learning section, we found that when students felt more in control of their participation then they felt like they have learned a lot, but in fact, they learned more when the level of perceived stress/negative emotions was low.

Regarding the learning gain, both the reported and the measured learning indicated that students learned during the course of the workshop regardless their gender. While some of the above-referred studies have examined the learning outcomes in terms of self-reported measures [273], our study results highlights the need for examining various levels of learning given that they are complementary in nature. As a knowledge assessment test can never capture all learning that has taken place, the reported learning provides students with the freedom to consider additional elements of learning (e.g., soft skills) that are not scrutinized by the knowledge test. In our study we found that students' reported learning has a positive association with their state-level emotions - also found by [210] -, and that the measured learning is negatively influenced by high levels of stress. These latter results are in synchrony with Pekrun's Control-Value Theory [213] with regards to negative deactivating emotions. We conclude that high arousal negative emotions interfere with active engagement with the task, therefore with the learning process as well.

Our results on the influential role of fun in learning are in contrast with the previous findings of Sim et al. [275] who investigated learning with a learning game with students between age 7 and 8, and that of Iten and Petko [127] who investigated learning with a learning game with students between age 10 to 13, and who did not find significant correlation neither between the observed nor the reported fun and the learning outcomes. They are in line with the work of others [198, 210, 260], who discussed the coding activity in terms of students having fun while the coding activity increased students' attitude toward coding and their learning outcomes and observed a higher GPA (grade point average) among hackathon participants compared with non-participating student -

although in those studies researchers did not examine the relationship, just report on the co-existence. We propose that the discrepancy in earlier studies regarding the effect of fun on learning might be rooted in the sort of relationship investigated (correlation vs causation and direct vs indirect effect), and hence, future studies should focus more on the type of relationship examined to increase our understanding on the role of fun in learning.

4.6 Limitations and Future Work

The herein introduced exploratory case study has been conducted with one school class. As a consequence, our findings are not representative for all. Replicating the study in different contexts eventually with more students would be beneficial for the assessment of the generalizability of the results. Additionally, by the application of quantitative methods, future research could investigate in-depth and explain the stated relationships. Moreover, future research should address the long-term effects of such interventions on participants' attitudes to cover the existing research gap noted by Master et al. [178].

4.7 Conclusion

In sum, previous findings indicated a gender difference in attitude toward coding that could be positively shaped and hence equalized by providing positive coding experiences for children and adolescents, e.g., by introducing coding with a visual programming environment. Our study results provide further support for such findings. However, the current state of the art has no clear view on what influences children's and adolescents' attitude toward programming, and hence their willingness for participation in coding activities. Our research contributes greatly to a better understanding of children's and adolescents' programming-related attitude by investigating possible underlying factors, and the interplay of those with the learning outcomes. Based on the herein introduced results we conclude that children's attitude about programming is greatly influenced by their learning experience and is in relation with the experienced fun while learning to code, which has further impact on knowledge acquisition. Crucially, our study suggests that related research needs to attend to the difference due to the complementary nature of the reported and the measured learning, which can explain contradictory or surprising results of earlier studies. Further, our results draw attention to the downshifting effect of high arousal negative emotions on the measured learning. Given the exploratory nature of this study, the next steps require investigations in different contexts to be able to draw generalized guidance and a more detailed picture on the possible influential factors and key-emotions in the technology-based learning environment.

5 Fun in DGBL – a Case Study[12]

In the previous chapter we introduced a case study in the context of a playful programming workshop with primary school students, which aimed to investigate emotions, attitude, and fun while learning to code. In the current chapter we present a following case study, which was conducted with secondary school students in the context of digital game-based learning (DGBL), and which focused on the experienced fun, motivation, attitude, self-efficacy and intention to play while learning with a digital game. The chapter, thus, contributes with a deeper understanding on how fun influences learning in the context of DGBL, and with secondary school students.

Summary

Digital Game-Based Learning (DGBL) has been attracting increasing attention from researchers and educators, especially as related studies suggest it can enhance learning and positively affect students' motivation, attitude, self-efficacy, and intention to play similar games. However, research into DGBL is not explicit about the role of fun in DGBL. In this study we hypothesized that the perceived fun while playing with an educational game has a positive impact on students' measured and perceived learning, motivation, attitude, self-efficacy and intention to play similar games. We conducted an online survey study with 28 secondary school students (M_{age} = 13.54) before and after playing with an online educational game on the topic of biology. The activity took place during school hours, in the formal-learning context. The results indicate that the fun they experience while playing the game has a significant and positive effect on the perceived learning, the change in students' motivation and self-efficacy, and the intention to play similar games. However, no significant effect was found on the measured learning and the change in attitude towards the subject. Our findings partially support the contention that making DGBL more fun improves learning. Future research should seek further empirical evidence in other topic areas and for different ages, and to explore how DGBL can help improve attitudes towards the topic as well.

5.1 Introduction

Digital learning games have gained popularity [79] due to the increased availability of computer and multimedia technologies at schools. A recent systematic review by Hainey et al. [106] indicated that digital game-based learning (DGBL) has been widely applied to various topics in science, mathematics, languages, social issues, history, and music. There is a growing body of empirical evidence showing that DGBL enhances students' learning

[12] This chapter is based on the following publication: Tisza, G., Zhu, S., Markopoulos, P. (2021). Fun to Enhance Learning, Motivation, Self-efficacy, and Intention to Play in DGBL. In: Baalsrud Hauge, J., C. S. Cardoso, J., Roque, L., Gonzalez-Calero, P.A. (eds) Entertainment Computing – ICEC 2021. ICEC 2021. Lecture Notes in Computer Science, vol 13056. Springer, Cham. https://doi.org/10.1007/978-3-030-89394-1_3

[40, 58] and their motivation to learn [44, 70, 111, 122]. Additionally, DGBL is associated with increased self-efficacy [4, 122, 188, 286, 322], better attitude toward the subject [6, 286] and increased intention to play learning games in the future [247].

Teaching with digital games builds on the idea that learning with digital games resembles the leisure time that people spend playing video games: it is fun and intrinsically motivating [205]. Accordingly, in DGBL literature the assumption that educational games are fun is rarely doubted or verified through measurement. As follows, empirical research with measured entities is scarce on the possible effects on fun in DGBL, and often contradictory. For example, Sim, MacFarlane and Read [275] found that the fun seven and eight years-old children experienced while learning with a digital game was not correlated with learning. Iten and Petko [127] reported similar results with primary school students: their regression analysis revealed that the enjoyment of the game is not associated with the self-reported learning, nor to the measured learning. However, a meta-analysis on computer games as learning tools [144] and a literature review on game-based learning suggested a direct link between learning and enjoyment [58]. Additional to the aforementioned findings, previous research on serious games pointed out that children's interest to engage with similar subjects increased according to the perceived enjoyment, and that children's intention to use similar games was significantly influenced by their attitude (anticipated simplicity and usefulness) toward the game [127]. Given the wide interest in this topic, the empirical research on the role fun plays in DGBL is quite limited and the results are inconclusive.

In this study, we set out to investigate how the experienced fun while playing a digital game influences students' learning, motivation, attitude, self-efficacy, and intention to play similar games. We recruited secondary school students to play a biology educational game - Code Fred: Survival Mode. Before and after playing the game, we administered a knowledge assessment test, students' motivation, attitude, and self-efficacy, and after playing the game we additionally measured students' perceived learning, intention to play similar games and we asked them to report on the fun they have experienced while playing. In the following sections of the chapter we introduce related research, followed by the study methods, results and the discussion of the herein introduced findings.

5.2 Background

5.2.1 Digital Game-Based Learning

Digital game-based learning (DGBL) is a recently emerged term that refers to acquiring knowledge and skills through playing engaging computer games [229]. It is frequently used interchangeably with other terms such as educational games, learning games, serious games, and edutainment. Throughout the chapter we use the term DGBL, and we refer to games as educational games.

Nowadays, educational games are being used across a wide range of subjects. In a recent literature review Boyle et al. [40] found that science, technology, engineering, and mathematics (STEM) are the most popular subject disciplines where DGBL is applied. This confirms earlier studies [58, 93, 106] and is of particular relevance as STEM abilities are seen as crucial for future development of oneself as well as the society [261].

5.2.2 Fun in DGBL

In this study we operationalize fun as defined in Chapter 2. Accordingly, any learning activity, including game-based learning is fun when one is intrinsically motivated for participation, feels in control of the activity, immersed in the experience by losing sense of time and space, the level of skills meets the level of challenge, the activity evokes positive and not negative emotions, and it supports the abandonment of social inhibitions.

It is often assumed that DGBL is fun. However, Yee [333] argued that educational games require players to do many tasks, making students feel tired and tedious. In other cases educational games were found to take too long and were no longer fun to play after the novelty effect was gone [133].

The effect of fun on learning with digital games is more controversial. Some studies found a positive association [58] and others no relation at all [127, 275]. Long [169] reported on a game for learning to program from which 80% of the study participants learnt something, while Sim, MacFarlane and Read [275] examined three educational applications for young children and reported no significant correlation between the observed or the reported fun and students' learning. When an educational game reaches its purpose and it is indeed fun to play with then the perceived enjoyment of the game motivates students to continue learning about the subject taught in the game [127] and increases students' attitude about the subject [127]. Additionally, previous research found that a higher level of game enjoyment is correlated with a higher motivation to learn [127]. However, the same article [127] reported that the enjoyment of the game had no influence on students' intention to play again, despite other studies suggesting the opposite [275]. There appears to have been no prior studies examining the possible relationship between fun in DGBL and self-efficacy.

5.2.3 Learning in DGBL

As mobile technology advances, digital games are no more bound to desktop computers and video consoles and can be played at different locations and times of the day [17, 40, 216]. Learners are having ample opportunities to play digital games, which has spawned the interest in their potential benefits regarding knowledge and skill acquisition, but also their affective, motivational, perceptual, physiological and cognitive outcomes [40, 58]. Research in DGBL has mostly focused on learning outcomes [40, 58, 123]. Empirical findings to date are divided. Some studies indicate that digital games are suitable for learning purposes for varied subjects including math, science, biology and psychology [40, 45, 58, 93, 124, 169], while other studies report an adverse effect of digital games on learning [110, 319, 332]. For the assessment of learning in DGBL, researchers use either knowledge tests [20, 120, 205, 249, 336] or measure learning based on the learners' self-report [127, 203].

5.2.4 Motivation, Attitude, Self-efficacy, and Intention to Play in DGBL

DGBL can enhance motivation to acquire certain knowledge [119, 125], which is a key element of successful learning [315]. Tüzün et al. [312] found that primary school students had a significantly higher intrinsic motivation and lower extrinsic motivation in a geography DGBL environment as compared to their motivation in the traditional school context. Others reported that students' motivation for mathematics was significantly higher for students learning in a mathematical game-based learning environment in comparison to a control group [122]. Such a positive relationship between students' learning motivation and DGBL is further supported by other studies [44, 70, 111]. On the other hand, Huizenga, Admiraal, Akkerman, and Dam [120] reported no significant differences in motivation between children learning in a DGBL and those in control groups.

Earlier research demonstrated that learning in DGBL can improve participants' attitude toward the subject [6, 127, 286]. Iten and Petko [127] found that the more children enjoy playing a DGBL the more they get interested in the subject matter. Akinsola and Animasahun [6] also found that students' achievement and positive attitude toward mathematics can be improved by the use of simulation-games environments. Sung and Huwang [286] developed a mindtool-integrated collaborative educational game, which was found to promote students' learning related attitudes.

Self-efficacy refers to students' perceptions and beliefs about their academic capabilities [266], and was found to be an effective predictor of academic motivation and achievements [347]. Hun, Huang, and Hwang [122] reported that students gained more self-efficacy in a mathematical game-based e-book learning environment compared with traditional instruction methods. Afari, Aldridge, Fraser, and Khine [4] measured students' academic efficacy before and after playing a mathematics game, and reported a significant improvement. Meluso, Zheng, Spires, and Lester [188] also suggested that after playing an educational game students demonstrated an increase in self-efficacy toward science.

Another study investigating serious games [127] reported that the anticipated usefulness and the anticipated simplicity of the learning game - which they labeled as attitude - are significant and positive predictors of the intention to play similar games. This is in accordance with Çankaya and Karamete [48] who found that DGBL had a positive influence on participants' intention to play the game again and with Rambli, Matcha and Sulaiman [232] who reported on children's strong willingness to play again with an augmented reality learning game.

The contradictory results between previous research on the effect of fun in DGBL may be attributed to inconsistent conceptions and measurement of fun and enjoyment in earlier studies. For many of the studies reported above, the notion of enjoyment and fun are left implicit [127, 169, 275]. Additionally, some of them used measures that may lack a theoretical basis, and consist of a single item [169, 275] where the predictive validity is questionable. Others used informal observations [275], which are not sufficient in conceptualizing and quantifying the role of fun in DGBL. The only study that is similar to the herein introduced study is that of Iten and Petko [127], however, despite using multi-item measures, they did not provide a clear definition of fun, which they use interchangeably with enjoyment. Nevertheless, they questioned the role of fun in serious games and called for research to investigate possible aspects of engagement from different angles.

5.2.5 Hypotheses

To address the apparent gap in earlier research regarding the potential benefits of DGBL we conducted a study aiming to understand how fun can impact students' learning, motivation, attitude, self-efficacy, and intention to play similar games in DGBL. Accordingly, we hypothesized that:

- **H1**: The experienced fun during DGBL positively affects students' reported learning.
- **H2**: The experienced fun during DGBL positively affects students' measured learning.
- **H3**: The experienced fun during DGBL positively affects students' motivation.
- **H4**: The experienced fun during DGBL positively affects students' attitude.
- **H5**: The experienced fun during DGBL positively affects students' self-efficacy.
- **H6**: The experienced fun during DGBL positively affects students' intention to play similar games.

5.3 Method

5.3.1 Participants

We recruited students from a Dutch bilingual secondary school. Our selection criteria were to have a good level of English comprehension and to be between age 13 and 14. The study was approved by the Ethical Board of Eindhoven University of Technology, Department of Industrial Design and informed consent was obtained from the participants and their parents before the study took place. In total, 28 of the second-year students participated in the study (12 boys, 16 girls, $M_{age} = 13.54$, $SD = 0.508$).

5.3.2 Procedure

The study took place in the spring of 2020. Since all formal educational activities took place online due to the COVID19 pandemic, participants took part in this study (i.e., played the game and responded the questionnaires) in an online classroom setting instead of the traditional physical environment. All students had access to laptops and had experience with using laptops to follow online lessons. Before the study started, the students were informed about the procedures by the teacher and informed consent was obtained. At the beginning of the lesson, students received a step-by-step guide to follow. This instructional document introduced the steps to take during the lesson and provided links to the questionnaires and the game. Throughout the lesson all participants were present in an online group meeting, so whenever a question or a problem emerged, students could ask help directly from the researcher and the teacher.

The study consisted of three sections: the pre-game data collection followed by playing the game Code Fred: Survival Mode, and the post-game data collection. Responding to the questionnaires before and after the game took approximately 7 minutes each. Students had approximately 25 minutes to complete the game. Ten minutes before the end of the lesson we asked students to stop playing the game - if they were still busy at that time - and do the post-game questionnaire.

5.3.3 Measures

To assess the interaction between the experienced fun and students' learning, motivation, attitude, self-efficacy, and intention to play similar games, we adopted validated measures and collected data from the study participants before and after playing the game. The pre-game questionnaire investigated students' motivation, attitude, and self-efficacy, along with a knowledge assessment test for the measured learning. The post-game questionnaire additionally investigated students' self-reported (i.e., perceived) learning, the fun they have experienced during playing the game and their intention to play similar games. The used measures, their operational definition, the items we used, and their respective sources are presented in Table 5.1.

For the assessment of fun we used FunQ [299]. In this study we excluded three items referring to social interactions given that participants played the game alone. Accordingly, participants rated their agreement on a 5-point scale (1 - Strongly disagree; 5 - Strongly agree) along fifteen items.

According to Bloom's taxonomy [34], six levels of learning can be distinguished. In this study we addressed two levels by collecting data on knowledge acquisition (*Knowledge*) and perceived learning (*Evaluation*). For the assessment of knowledge acquisition, we designed a test consisting of six multiple choice, single selection questions. The six test questions refer to six chapters of the game by asking about the knowledge and the related task to save Fred. For example, in chapter one the message appears on the screen 'Danger is detected. Send adrenaline to these organs to escape: eye, hearth, liver'. In the knowledge test we ask students the following: When your body detects danger, it reacts by sending adrenalin to the following body parts to escape EXCEPT: eyes / lungs / liver / brain. To gain a point, students must select the correct response (lungs). For each question participants could gain one point, resulting in a maximum of six points. We used the difference between the pre-game and the post-game scores as an indication of learning (i.e., measured learning).

The perceived learning was measured by a single item measure adopted from previous research [210, 300]. Participants were asked ‘Have you learned something new today about biology?’ and they could indicate their agreement on a 5-point scale (1- Not at all; 5 - A whole lot).

To evaluate students’ attitude toward biology, we adapted three items from the Motivated Strategies for Learning Questionnaire [222]. During adaptation we slightly changed the original items to fit better the study purpose (e.g., instead of ‘It is important for me to learn the course material in this class’ we used ‘It is important for me to learn what’s taught in the biology class.’ as we were interested in students’ general attitude toward biology classes, not the specific gamified class we conducted for the study). Students indicated their agreement with the items on a 5-point Likert scale (1 - Strongly disagree; 5 - Strongly agree).

For the assessment of students’ motivation toward biology we adapted three items from the Attitudes toward Mathematics Inventory [292]. During the adaptation we exchanged the word ‘mathematics’ into ‘biology’ to fit the study purpose (e.g., original item: ‘I am willing to take more than the required amount of mathematics’; adapted item: ‘I am willing to take more than the required amount of biology’). Students rated their agreement with the items on a 5-point scale (1 - Strongly disagree; 5 - Strongly agree).

For the measurement of students’ self-efficacy, we adapted three items from the Motivated Strategies for Learning Questionnaire [222] to have the focus on biology classes in general and not the gamified learning setup applied for the study (e.g., original item: ‘I expect to do well in this class.’; adapted item: ‘I expect to do well in biology class’). Students were asked to indicate on a 5-point scale their agreement with the items (1 - Strongly disagree; 5 - Strongly agree).

To address students’ intention to play similar games we adapted three items from the Motivated Strategies for Learning Questionnaire [222], which were evaluated on a 5-point scale (1 - Strongly disagree; 5 - Strongly agree). For this dimension, we used original items that indicate task value and extended it to become a reasoning for playing similar games again (e.g., original item: ‘I think the course material in this class is useful for me to learn.’; adapted item: ‘I want to do similar activities in my biology classes because I think these kinds of activities are useful’). Cronbach’s alpha values indicate a good internal consistency for all used scales (see values in Table 5.1).

Table 5.1 The measured factors, their sources, operational definitions, and respective questions. All items were evaluated on a 5-point scale (1 – Strongly disagree; 5 – Strongly agree)

Factor [source]	Operational definition	Item/Question
Fun [299] (α = 0.940)	The degree to which students experienced fun during the activity.	I did this activity because I had to. (reversed item) I did this activity because I wanted to. I want to do something like this again. During the activity… I knew what to do. I felt I was good at this activity. I did something new. I was curious. I had fun. I was happy. I felt that time flew. I forgot about school. I felt good. I felt bad. (reversed item) I felt angry. (reversed item) I felt sad. (reversed item)
Perceived learning [210]	The degree to which students indicate their learning during the activity.	Have you learnt something new today about biology?
Motivation [292] (α_{pre} = 0.789; α_{post} = 0.912)	The degree to which students indicate their motivation toward the subject.	I am willing to take more than the required amount of biology. I plan to take as much biology as I can during my education. The challenge of biology appeals to me.
Attitude [222] (α_{pre} = 0.822; α_{post} =0.850)	The degree to which students indicate their attitude toward the subject.	I think what we are learning in biology class is interesting. It is important for me to learn what's taught in the biology class. I think what I'm learning in the biology class is useful for me to know.
Self-efficacy [222] (α_{pre} = 0.608; α_{post} = 0.757)	The degree to which students indicate their self-efficacy toward the subject.	I'm certain I can understand the ideas taught in the biology class. I expect to do very well in biology class. I know that I will be able to learn the material for the biology class.
Intention [222] (α = 0.945)	The degree to which students indicate their willingness to play similar games.	I want to do similar activities in my biology classes because I find these kinds of activities are interesting. I want to do similar activities in my biology classes because I think these kinds of activities are useful. I want to do similar activities in my biology classes because I feel these kinds of activities are important.

Measured learning [n.a.]	The difference between the post- and pre-score on the knowledge test.	1. When your body detect danger, it reacts by sending adrenalin to the following body parts to escape EXCEPT: *eyes / lungs* / liver / brain* 2. Your body delivers oxygen from the lungs to leg muscles by the help of: *neuroglobin / cytoglobin / hemoglobin* / myoglobin* 3. When your body is losing blood, it needs to quickly heal the wound by gathering the following elements EXCEPT: *platelet / clotting factor / fibroblast / epithelium** 4. When your body is invaded by bacteria, it will inspect the bacteria in your: *blood / infected cells / lymph node* / spinal cord* 5. When your body sends an 'infection alert', the following happens: *your body warns you to take a paracetamol / your body releases antibodies to disable bacteria before they multiply and infect cells* / your body makes you thirsty, so you'll drink a lot and flush the bacteria away / your body releases antihistamine to kill bacteria before they attack your organs* 6. When your body has high blood sugar, it needs to release: *glucose from pancreas / glucose from liver / insulin from pancreas* / insulin from liver*

* Indicates the correct answer for the knowledge test

5.3.4 The Game

To test these hypotheses, we asked students to play Code Fred: Survival Mode (see Figure 5.1). In choosing the game we had several requirements:

- The subject covered is equally suited for boys and girls. Research has shown that among STEM subjects, biology is most equally appealing for both genders [304, 330].
- The length of a game session is suitable for classroom use, taking no longer than 30 minutes.
- The game covers a certain topic sufficiently, so that knowledge acquisition can be measured meaningfully.
- The game is fun and educative and aligns well with participants' knowledge and curriculum.

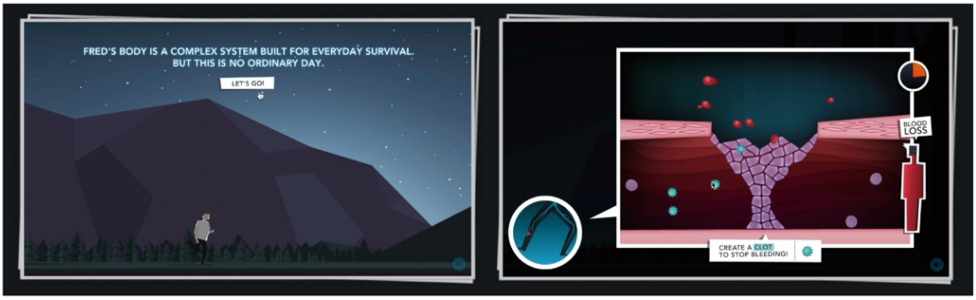


Figure 5.1 Screenshots of Code Fred: Survival Mode. *Left*: opening scene. *Right*: episode 5 – gather elements from across the body to quickly heal the wound.

The game is developed by the Museum of Science + Industry Chicago and it teaches the player about the human body. The player leads the injured Fred through twelve episodes, e.g., to get oxygen to the muscles, to stop bleeding, to heal a wound, to inspect and disable bacteria, and to maintain a stable blood sugar level. The game is completed when the player completes all twelve episodes and brings Fred safely back to his camp. The game takes about 25 minutes, it has a well-defined topic and connects well to the curriculum, and it is fun to play for both boys and girls.

5.4 Results

To assess the fun experienced during the herein introduced study we calculated the FunQ scores by summing the values after correcting for the values of the reversed items. This resulted in an average score of 51.43, (SD = 12.04) from the possible range of 15 – 75, from which we conclude that students had fun while playing the game. Correlation analysis between the FunQ and the item 'During the activity, I had fun' indicates a significant correlation ($r = 0.910$, $p < 0.001$), supporting the validity of the aforementioned claim.

Regarding the measured learning, we first calculated the scores for the pre-game and the post-game knowledge assessment tests ($M_{pre\text{-}game} = 2.39$, $SD = 0.96$; $M_{post\text{-}game} = 3.57$, $SD = 1.40$). Independent sample t-test found no gender difference in the knowledge test scores ($p_{pre\text{-}game} = 0.781$; $p_{post\text{-}game} = 0.266$). Then, we subtracted the pre-game test scores from the post-game scores. This resulted in an average score of 1.18 (SD = 1.16) for the measured learning. Paired sample t-test indicates that this difference is significant ($p < 0.001$), thus we conclude that based on the measured learning scored students have learned by playing the game.

To address participants' perception about their learning and test we asked them after playing the game whether they thought they had learned something new about biology. Students self-rated their learning on average 2.93 (SD = 1.09) on a 5-point scale, which translates to having learnt 'something'.

To compare students' reported learning with the measured learning scores we applied correlation analysis. We did not find a significant correlation between the measured and the perceived learning ($r = 0.246$; $p = 0.206$).

The reported average score for students' motivation at the beginning of the study was 2.64 ($SD = 0.81$), and it was 2.82 ($SD = 1.04$) on a 5-point scale after playing the game. We did not find any gender difference (independent sample t-test, $p_{pre\text{-}game} = 0.110$; $p_{post\text{-}game} = 0.865$). The average change in the motivation score is 0.18 ($SD = 0.67$). Paired sample t-test indicates that the change is not significant ($p = 0$ .170).

As for students' attitude toward biology, the average score at the beginning of the study was 3.60 ($SD = 0.80$), and it was 3.63 ($SD = 1.00$) on a 5-point scale after playing with the game. Independent sample t-test indicates no gender difference in the scores ($p_{pre\text{-}game} = 0.708$; $p_{post\text{-}game} = 0.930$). The average change in the attitude toward biology score is 0.04 ($SD = 0.53$). Paired sample t-test indicates that the change is not significant ($p = 0.725$).

Regarding the self-efficacy, the reported average score at the beginning of the study was 3.58 ($SD = 0.49$), and it was 3.64 ($SD = 0.70$) on a 5-point scale after playing with the game. No gender difference was found (independent sample t-test, $p_{pre\text{-}game} = 0.203$; $p_{post\text{-}game} = 0.741$). The average change in the self-efficacy score is 0.06 ($SD = 0.58$). Paired sample t-test indicates that the change is not significant ($p = 0.592$).

Regarding students' intention to play similar games, the reported average score at the end of the game was 3.43 ($SD = 1.19$) on a 5-point scale.

For a summary of the pre- and post-game scores and the related statistics see Table 5.2.

Table 5.2 Pre- and post-game mean scores of measured learning, motivation, attitude, self-efficacy, and intention to play. All items were evaluated on a 5-point scale (1 – Strongly disagree; 5 – Strongly agree).

	Pre-game	Post-game	p	Cohen's D	Gender difference
Measured learning	$M = 2.39$	$M = 3.57$	< 0.001	0.984	$p_{pre} = 0.781$
	$SD = 0.96$	$SD = 1.40$			$p_{post} = 0.266$
Motivation	$M = 2.64$	$M = 2.82$	0 .170	0.192	$p_{pre} = 0.110$
	$SD = 0.81$	$SD = 1.04$			$p_{post} = 0.865$
Attitude	$M = 3.60$	$M = 3.63$	0.725	0.039	$p_{pre} = 0.708$
	$SD = 0.80$	$SD = 1.00$			$p_{post} = 0.930$
Self-efficacy	$M = 3.58$	$M = 3.64$	0.592	0.099	$p_{pre} = 0.203$
	$SD = 0.49$	$SD = 0.70$			$p_{post} = 0.741$

5.4.1 Fun and Learning

To quantify the effect of fun on the measured learning and test hypothesis 2, we conducted a regression analysis. The analysis reveals that the experienced fun explains 7.1% of the variance in the measured learning scores ($R^2 = 0.071$; $\beta_{std} = 0.266$, $p = 0.172$). In other words, the effect of the experienced fun on the measured learning is small and non-significant.

To understand how the experienced fun affects the perceived learning and test hypothesis 1, regression analysis was applied. The analysis reveals that the experienced fun explains the 42.4% of the variance in the perceived learning scores (R^2 = 0.424; β_{std} = 0.651, p < 0.001). In other words, students who experienced more fun with the game report to have learnt more compared to those who experienced less fun.

5.4.2 Fun and Motivation

To test hypothesis 3 regarding the effect of fun on students' motivation, a regression analysis was applied. Results indicate that the experienced fun while learning with a digital game explains 31.1% of variance in the post-workshop motivation scores (R^2 = 0.311; β_{std} = 0.558, p = 0.002). Thus, having fun while learning significantly contributes to students' increased motivation toward the subject.

5.4.3 Fun and Attitude

To quantify the effect of fun on students' attitude and to test hypothesis 4, we conducted regression analysis. Results show that 42.5% of the variance in the post-workshop attitude scores is explained by the experienced fun (R^2 = 0.447; β_{std} = 0.668, p < 0.001), which effect is significant. In other words, having fun while playing digital games contributes significantly to students' attitude about the topic.

5.4.4 Fun and Self-efficacy

To assess the relationship between fun and self-efficacy and to test hypothesis 5, a regression analysis was applied. The analysis indicates that fun accounts for 51.8% of the variance in the post-workshop self-efficacy scores (R^2 = 0.518; β_{std} = 0.720, p < 0.001). Thus, having fun while playing with an educational game significantly contributes to students' increased self-efficacy.

5.4.5 Fun and Intention to Play

To test hypothesis 6 regarding the effect of fun on students' intention to play similar games we applied regression analysis. Results show that 70.3 % of the variance in the intention to play similar games scores is explained by the experienced fun (R^2 = 0.703, β_{std} = 0.838, p < 0.001). In other words, having fun while learning with a digital game has a considerable impact on students' willingness to play similar games. For a summary of all regression analysis results see Table 5.3.

Table 5.3 Summary of the regression analysis results. The effect of fun on students' learning, motivation, attitude, self-efficacy, and intention to play.

	R^2	β_{std}	p
Measured learning	0.071	0.266	0.172
Perceived learning	0.424	0.651	< 0.001
Motivation	0.311	0.558	0.002
Attitude	0.447	0.668	< 0.001
Self-efficacy	0.518	0.720	< 0.001
Intention to play	0.703	0.838	< 0.001

5.5 Discussion

Our study investigated the role that fun plays in DGBL, for which earlier research is scarce and results are contradictory. Some researchers found no relationship between fun and learning while playing with an educational game [127, 275], while others [58, 144, 169] report on a significant association between the two. While we did not find a significant relationship between the experienced fun and the measured learning (H2), we did find that the experienced fun while playing a digital learning game has a significant and positive effect on students' perceived learning (H1). Hence, hypothesis 1 is supported and hypothesis 2 is refuted. A possible explanation by Koriat and Bjork [152] is that perceived learning often does not reflect the actual learning achievement since the judgment of learning is influenced by various factors. Furthermore, we suggest that students' perception of their learning might not only refer to factual knowledge but includes other skills that are not part of the knowledge assessment test. Additionally, the two measures address different levels of learning according to Bloom's taxonomy [34], which also explains the discrepancy between the two scores.

As for the effect of fun on students' motivation, our results provide an explanation to previous findings [44, 70, 111, 119, 122, 125, 312], as we found that having fun while learning with a digital game has a significant and positive influence on students' motivation (H3). Thus, our results support hypothesis 3. While previous studies on DGBL found that DGBL enhances students' motivation, they did not investigate from where exactly this relationship derives from. Our findings suggest that experiencing fun while learning with a digital game affects significantly students' motivation, resulting in the previously often found positive effect of DGBL on students' motivation.

Further, the effect of fun on attitude toward biology was significant in our study (H4), which is similar to an earlier report of a significant effect of fun on students' subject-related attitude [127] and to our earlier findings related to learning to code reported in Chapter 4. This result is also in line with previous research [6, 127, 286], which demonstrated that DGBL can improve students' attitude toward the subject, but those earlier studies did not investigate possible underlying reasons. Therefore, hypothesis 5 is justified. However, our research goes one step further then earlier studies as it provides a possible explanation to the earlier find association between increased attitude and the use of DGBL, by suggesting that students' subject-related attitude is explained partially by the fun students experienced while learning.

Our results also indicate that having fun while learning with an education game has a significant influence on students' self-efficacy (H5). Thus hypothesis 5 is justified. We propose that this finding refines and explains earlier reports [122, 188], that DGBL enhances students' self-efficacy.

Regarding the effect of the experienced fun while learning with a digital game on students' intention to play similar games in the future (H6), our findings support previous research [275] and concur with the one of Iten and Petko [127] as we found that fun is a strong influencer on students' willingness for engaging with similar games in the future. Accordingly, hypothesis 6 is supported by the results.

5.6 Limitations and Future Work

This study has been conducted as an online class, which is different from the traditional formal learning environment. Nevertheless, since the game was online and students had access to the researcher and their class teacher during the study, we argue that student's experience was not hindered by the setup. The use of the single-item measure for the perceived learning, we believe, did not hinder the predictive validity given that the item was successfully used in previous research and it measures a concrete and simple construct. Nevertheless, future research is required focusing on the differences between the measured and perceived learning. Furthermore, while there was a clear increasing tendency in students' motivation, attitude and self-efficacy scores, the difference between the pre-game and post-game scores were not significant. We attribute this finding to the properties of the game and speculate that a different topic or a more fun game would have resulted in a stronger increase in the aforementioned scores. Accordingly, we call for future research with different games and on different subjects to investigate the herein revealed effects.

5.7 Conclusion

In sum, this study supports designers, researchers and educators making digital game-based learning fun as our research demonstrated a positive effect of fun on students' motivation, self-efficacy and intention to engage with similar games, which are considered as key factors to successful learning. Additionally, our results also shed light on the underlying effect of fun on the often-found positive association between DGBL and motivation, attitude and self-efficacy. Namely, we found that experiencing fun while learning with a digital game positively influences students' motivation, attitude and self-efficacy, which explains why in earlier studies a positive association was found between DGBL and the investigated factors. Moreover, we provided a possible explanation for the disputed effect of fun on learning. Accordingly, we encourage both designers and researchers to not take fun granted in DGBL and we call for future research for a better understanding of fun in digital game-based learning. To conclude, this study further extended our understanding on the effect of fun on learning in a different context compared with the study introduced in Chapter 4 (DGBL vs. coding workshop), and on a different age group (13-14 vs 10-12 years). Remarkably, in both context and age groups our findings indicate a positive association between fun and students' learning outcomes and their subject-related attitude.

6 The Fun in Learning (FiL) Model[13]

In Chapter 4 and Chapter 5 we presented two case studies that aimed to investigate the relationship between fun, learning, emotions, attitude, self-efficacy, and intention to play with primary and secondary school students, within a context of a programming workshop and a digital game-based learning class. The findings of those case studies provided us with insights to the possible relationship between fun and learning. In the current chapter we hypothesize a model on this relationship and test it by means of structural equation modeling. This chapter, therefore, contributes with the FiL model that quantifies the relationship between fun and learning.

Summary

As discussed in the earlier chapters, there are growing efforts amongst educators and especially researchers in gamification and maker spaces to ensure that learning environments are fun and enjoyable. Accordingly, efforts to evaluate whether students enjoyed a certain learning activity are often an important aspect of innovations. However, the role of fun in learning in the aforementioned fields is not well understood not least due to a lack of a common theoretical framework for defining the concept of fun and for supporting its measurement. Building on our earlier studies introduced in the previous chapters, in this study we set out to investigate and conceptualize the role that fun plays on students' learning to program and its impact on their attitude towards the topic. Accordingly, we designed a two-hour-long programming workshop to introduce the topic to students in a playful way. The workshop took place during school hours, however, as a non-curricular activity. 86 students with ages between 9 and 12 participated in the study. For the analysis, we used structural equation modelling and mediation analysis. Our results support efforts of educational researchers and practitioners who try to make learning activities more fun for students. While fun was not shown to have a direct effect on learning assessed through a self-report measure, it had a significant and positive indirect effect on perceived learning through student's attitude towards coding. According to these findings, this chapter introduces the Fun in Learning (FiL) model.

6.1 Introduction

One of the primary aims of educators is to get and keep learners motivated, which can improve learning effectiveness [169]. Accordingly, substantial research effort has been invested in making learning enjoyable. Gamification of learning is a popular approach [169] which is founded upon a commonly accepted belief that games can make learning fun. However, as Iten and Petko note, "*it is less clear what fun in serious games actually*

[13] This chapter is based on the following publication: Tisza, G., & Markopoulos, P. (2021). Understanding the role of fun in learning to code. International Journal of Child-Computer Interaction, 28, 100270. https://doi.org/10.1016/j.ijcci.2021.100270

means and how is it related to cognitive, emotional and behavioural engagement" ([127], p. 154). Additionally, they argued that despite that this belief is widely held, there is a lack of empirical evidence regarding the relationship between the experienced fun while learning and learning with serious games. Nevertheless, it is quite common for creators of such games to evaluate them based on whether learners experience fun (e.g., [234, 275]) or they even consider fun as a direct user benefit [263].

Another approach to making learning enjoyable is to develop learning environments that foster intrinsic motivation and engagement by triggering the learners' curiosity. Such learning environments can be found in informal and non-formal learning spaces such as science museums, maker spaces or coding clubs. A recent Europe-wide study on STEM-related (Science Technology Engineering and Mathematics) informal and non-formal learning activities reported that approximately two-thirds of the investigated activities intended to be playful, and more than one-third of the investigated activities aimed to engage their participants with scientific topics [304]. These findings reflect a worldwide pursuit of increasing children's interest in STEM fields by making learning fun [97] and particularly in computer science given that programming is frequently seen as the literacy skill of the 21st century [210].

From previous research we know that students' success in computer science courses and their career choices can be affected by their attitudes towards programming [50]. Earlier studies also indicated that children's attitude towards programming is positively affected by interacting with visual programming environments (such as Scratch) [104, 260]. In a field study, Long [169] found that the most common motivation (87.5% of their participants) for participating in a game-based coding experience was to have fun, and her study results suggested that having fun while learning results in a higher learning effort to be committed. Other studies described - either based on observations, or data collected with a 6-item unidimensional enjoyment scale - the coding activity in terms of students having fun, and reported an improvement in their attitudes towards programming and positive change in learning outcomes [210, 260], without however investigating the link between those. An important reason for this limitation in this body of research is that despite the apparent importance of fun during learning, there is a lack of measurement instruments to assess fun, and especially so in adolescents. These issues we have addressed in the earlier chapters (Chapter 2 and 3).

In order to investigate how fun affects learning to program, we report on a study where a playful learning to code activity is evaluated in three primary school classes. In this context we investigated how the experienced fun, measured by the FunQ, influences students' learning to code, taking into account students' attitude about the subject. The activity was presented in the form of a 2-hours-long workshop and aimed to introduce programming for primary school students with the use of micro:bits (pocket-sized programmable microcomputers). Before and after the session students filled in a questionnaire on their attitude towards coding. Additionally, after the workshop students were asked to evaluate the activity with the use of the FunQ and report on their perceived learning.

The current study is the first to assess the direct and indirect effects of fun on learning to program while taking into account students' attitude about the topic. In the following sections, we introduce earlier research and related theories and provide our theoretical framework for the study. Then, we present our method, the data collected and the data analysis and results, followed by the discussion of research findings, limitations, future research and conclusion.

6.2 Background

6.2.1 Learning, Attitudes, and Emotions

Traditionally, the aim of learning has been knowledge acquisition. Based on the work of Schunk [265], learning assessment can be supported by a) direct observations, b) written responses, c) oral responses, d) rating by others and e) self-reports. Under self-reports, Schunk includes questionnaires, interviews, stimulated recall, think-aloud and dialogue. We adopt the definition of Schunk and understand the self-report measure as *"people assessments of and statements about themselves"* ([265], p. 16). Several design research studies in education include informal assessments of fun and learning, which, however, do not allow quantifying the relationship between them (e.g., [103]). Therefore, in the herein introduced study we assessed students' knowledge by a self-report measure, which we adopted from previous research [210, 301].

Studies on the influence of attitudes on learning carried out as early as the 1950s (e.g., [84]), suggested an association between attitudes and academic achievement. Recent studies on the association between attitudes towards learning different subjects and the learning outcomes (a.k.a. academic achievement) [19, 31, 199] suggested that a positive attitude towards learning or an academic subject is associated with an inclination to learn. However, others found no direct relationship between the attitude about school and academic achievement [82, 129, 130, 162, 335]. Lee [162] investigating the PISA (Programme for International Student Assessment) 2003, 2009 and 2012 data sets - involving nearly a half a million students from 65 countries - found no significant correlation between mathematics and reading achievement and students' general attitude towards school. Thus, the question of the role of attitude on learning appears to be controversial.

Although enjoyment has traditionally been more extensively investigated than fun, we need to note that the notion of fun and the importance of the fun experience is getting widely acknowledged [184]. Accordingly, fun is often an evaluation criteria for example for learning games (e.g., [234, 275]) or even considered as a direct user benefit [263]. Nevertheless, the concept of fun has been hugely neglected in the field of psychology which is reflected in the fact that *"no psychology textbook has fun in its index"* ([184], p. 160).

Accordingly, fun is often used interchangeably with other positive emotions such as enjoyment. In a recent literature review, investigating emotions and learning in the technology-based learning environment in the past 50 years, Loderer, Pekrun and Lester [168] found that enjoyment was the second most frequently investigated academic emotion, accounting for approximately a quarter (24.1%) of all reviewed papers. They found a positive association between enjoyment and control and positive valuation (i.e., attitude) of learning/technology, engagement, strategy use, curiosity/interest and the learning outcomes.

Addressing the relationship between emotions and learning from the perspective of cognitive psychology and neuroscience, Willis claimed that *"when classroom activities are pleasurable, the brain releases dopamine, a neurotransmitter that stimulates the memory centers and promotes the release of acetylcholine, which increases focused attention"* ([327], p. 3). Hence, *"superior learning takes place when classroom experiences are enjoyable and relevant to students lives, interests, and experiences"* ([327], p. 1).

6.2.2 Conceptualizing Fun and its Assessment

Whatever we do, we have to make it fun has become a modern cliché, especially in the United States, applied to several fields of life starting from playing sports, throughout teaching children modern languages or encouraging people to eat more fruit and vegetables. [184]. Yet, as stated by various authors [32, 53, 72, 127, 186], defining fun is neither easy nor straightforward and accordingly, its measurement is complicated [275]. As a result, it is only occasionally defined in the academic literature, pinpointing the lack of underlying conceptual framework - and raising concerns about the clarity of measurements and unity of previous research [94, 184, 299]. This problem is well introduced in the work of McManus and Furnham [184] who present a huge variety of related concepts that are frequently used interchangeably with fun in academic literature, while they also note that *"only in a very occasional set of studies is there a direct confrontation with the nature of fun and its definition"* (p. 160) and accordingly, there is a *"lack of conceptual clarity in the literature concerning the nature of fun"* (p. 160). As already noted above, in academic literature, enjoyment is frequently used interchangeably with and as a synonym to fun [78, 89, 127, 169, 184, 186, 246]. However, other researchers [72] argued that meanwhile the two concepts are related, they are not the same as fun has a more nuanced interpretation.

Table 6.1 Related measurement tools and the dimensions they investigate compared with the FunQ dimensions.

Instrument	Dimensions						Target age
FunQ [299]	***Autonomy***	***Challenge***	***Delight***	***Immersion***	***Loss of Social Barriers***	***Stress***	adoles-cents
IMI [257]	Perceived choice	-	Interest / Enjoyment	-	Relatedness	Pressure / Tension	adults
PENS [135, 259]	Autonomy	-	-	Presence / Immersion	Relatedness	-	adults
Flow state scale [131]	Sense of control	Challenge-skill	Autotelic experience	Loss of self-consciousness & Transformation of time	-	-	adults
EGame Flow [91]	Autonomy	Challenge	-	Immersion	Social interaction	-	adults
GEQ [135, 225]	-	Challenge	Positive affect	Flow	-	Negative affect & Tension	adults

For this study we conceptualized fun as described in Chapter 2. While there are clear links between the FunQ and Self-Determination Theory (SDT) and the Intrinsic Motivation Inventory (IMI), as noted by Grosshandler and Grosshandler ([103], p. 228) *"self-determination is a crucial factor in the construction of fun and learning"*, however, we can make a clear distinction between those. The two questionnaires measure along different dimensions according to their underlying theories, and FunQ targets adolescents while IMI is designed for adults. The comparison of the dimensions of the FunQ and some frequently used measurement tools for closely related concepts are displayed in Table 6.1. In the herein introduced study for the assessment of fun we used the FunQ.

6.2.3 Fun in Relation to Attitudes and Learning

Most research on the relation between fun, attitudes and learning is found in the field of digital game-based learning, however, earlier research appears to be controversial. Chan, Wan and Ko [51] investigated whether fun has a moderating role in the relationship between interactivity, active collaborative learning and university students' learning performance while using personal response systems. Their findings indicated that fun students experience while using personal response systems could promote collaborative learning and learning performance. However, they did not measure learning as such but rather students' self-efficacy in using personal response systems (PRS), which included items such as *I have mastered the use of PRSs.* Iten and Petko [127] investigated empirically the relationship between enjoyment and willingness to play, as well as between the learning gains and knowledge assessment test results in students aged 10-13 playing serious games. Their results indicated that the experienced fun had a significant effect on the motivation to learn and to engage again with the learning game. However, they did not find a significant association between neither the self-reported nor the measured learning and the experienced fun. In accordance with these results, Sim, MacFarlane and Read [275] found no significant correlation between the observed- or the reported fun and the learning in 7 and 8 years-old-children using educational software. However, they find that the software children selected as the most fun was the one they would like to use again. In contrast with the aforementioned findings, several previous studies [51, 171, 232, 295, 317, 327] suggested that fun somehow has a positive effect on learning, but this effect is rarely quantified or measured directly, which is very likely to be due to the lack of reliable measurement tools and the lack of a common theoretical framework for fun.

6.2.4 Fun in Coding Activities

Coding is frequently seen as the literacy of the 21st century [210], and accordingly, there is a worldwide pursuit to increasing children's and adolescents' interest by making learning fun [97].

Long, in her field study with adults on programming games for educational purposes, found that having fun while learning was a significant contributor to the learning effort [169]. She examined across a survey whether the game promoted self-motivated learning and self-reported learning effectiveness. Cetin and Ozden [50] found in their large-scale study with university students that students' success in computer science courses - and their career choices - could be affected by their attitudes towards programming. Other studies focusing on younger students indicated that children's attitude towards programming is positively affected by interacting with visual programming environments (such as Scratch) [104, 260]. Additionally, further studies described the coding activity in terms of students having fun, and reported an improvement in their attitudes towards programming and positive change in learning outcomes [210, 260], without however investigating the link between those. Despite the aforementioned studies suggest that fun has some kind of influence on learning to code, based on our best knowledge no systematic examination exists on the relationship of those, nor a theoretical framework exists yet that would focus especially on the role fun plays on learning to program.

6.2.5 Hypotheses

Accordingly, based on the aforementioned studies, with special regards to previous research done in the field of learning to program, we set up the following hypotheses:

- **H1**: The experienced fun has a direct and positive effect on students' learning to code.
- **H2**: The experienced fun has a direct and positive effect on students' attitude about coding.
- **H3**: Students' attitude about coding has a direct and positive effect on their learning to code.

- **H4**: The experienced fun has an indirect and positive effect on students' learning to code.

The hypothesized relationships between fun, attitude and learning to code are depicted in Figure 6.1.

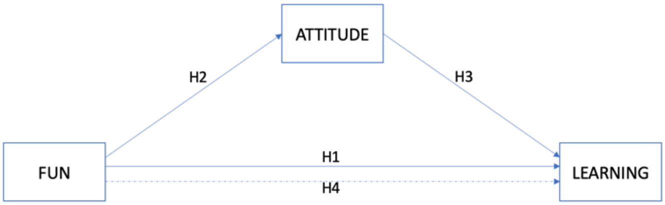


Figure 6.1 The role of fun on learning: hypothesized model. Dashed line indicates indirect effect.

6.3 Method

6.3.1 The Activity

To test our hypotheses, we collected data from participating students before and after three creative coding workshops. The main aim of the workshop was to introduce programming to students in a playful way. The workshop was designed as a non-formal learning activity, during which no specific learning goals were declared. Nevertheless, as it is frequent during such activities, it followed a learning-by-doing approach. The workshop lasted two-hours and was given to 9 to 12 years old primary school students during school hours, however, as an extracurricular activity. It was based on interactive instructional videos that introduce the basics of programming, defining and explaining terms, giving tasks to students, and providing instructions on the steps to take. We used the *Inventing with Microbit* video series of SkillsDojo[14], from which we selected three specific videos to play during each workshop. Accordingly, during the workshop students learned about basic programming terms (e.g., microcomputer, editor) and wrote code to display their names (first video), made a stone-paper-scissors game (second video; see Figure 6.2) or created micropets that react to kinetic stimuli (third video).

Playfulness was by nature inherent to the workshop as it was a non-formal activity, which built on students' voluntary participation, used digital technology and followed a learning-by-doing approach [142]. Students were invited to follow the video guide, in which the playfulness was reflected in the design, the tone and the introduction of the tasks. Additionally, students were encouraged to use their creativity to solve the programming tasks and to make the created artefacts unique to their liking. Once the artefacts were created, students could play with those alone or together with others.

[14] www.skillsdojo.nl

According to Huizinga [121] , the basic motivation to play is to experience the pleasure it grants. Hence, we assumed that a playful activity is a fun experience for the participants.

Figure 6.2 Children during the workshop. *Left*: interaction with the programming interface. *Right*: group play with the created stone-paper-scissors game.

6.3.2 Procedures

Each child was provided with a laptop or Chromebook, on which they could follow the video guide in their own pace. Additionally, each child was provided with a BBC micro:bit. Micro:bits are pocket-sized, powerful computers, with which one can easily learn to program and create electronics projects. It has a programmable LED display, buttons, sensors, and several input and output features allowing various ways for user interaction. Furthermore, students were also allowed to use paper, pencils, scissors etc. to dress up their microPets.

The workshop had the following structure:

1) Introduction of the topic and the structure of the workshop (~5 min)
2) Pre-activity data collection (~10 min)
3) Creative coding (~ 90 min)
4) Post-activity data collection (~10 min).

During the workshop, students were allowed to move around freely, to work alone or to interact with each other, which for no specific instructions or rules were given. When it was needed, the researcher provided students with further cues and helped when they asked for it. Thus, the role of the researcher was to help students troubleshoot eventual problems and was responsible for keeping the structure of the workshop, while the class teacher was present as an observer. During the workshop students used their creativity to solve the problems and the overall approach was playful throughout the introduction, exploration and outcome phases.

Although the data were collected from three workshops, the structure of the workshop and the collected data were similar (see the comparison in section 6.4.1).

6.3.3 Participants

Primary school teachers could voluntarily sign up their classes for the workshops, which were held in June 2019 in The Netherlands. The specific activity was not part of the curriculum, though the workshops were held in a classroom during school hours. Accordingly, students' participation in the workshop was compulsory as it took place during school hours, but their participation in the study (i.e., responding the questionnaires) was voluntary. Nevertheless, all students filled in the questionnaires. Given the students' age, informed consent was obtained from their parents across the schools, and the data was collected accordingly. The herein introduced results are based on the collected data from three creative coding workshop which were given in three Dutch primary schools. We collected data in total from 86 students (45 boys, 37 girls, 4 not indicated) between age 9 and 12 (M = 10.35, SD = 0.743). Before the workshop students reported on their previous experience with programming across a 5-step Likert-type question: 'Do you have any idea about programming?'. The frequency of the responses given to the question is displayed in Table 6.2. The mean of the responses is 2.48, which translates to knowing a bit. In other words, the participating students were mainly novices in the field of programming.

Table 6.2 The frequency of the responses for 'Do you have any idea about programming?'.

missing	(1) not at all	(2) I know a bit	(3) I know something	(4) I know much	(5) I am a pro
3 (3.5%)	15 (17.4%)	27 (31.4%)	24 (27.9%)	13 (15.1%)	4 (4.7%)

6.3.4 Measures

For the measurement of the elements in the model, we used a variety of tools (see summary in Table 6.3). When selecting the tools, we carefully considered the suitability of the selected measures for the responding students.

For assessing the experienced fun, we recorded the FunQ [299] with the students at the end of the activity. We present the items along with the related factors and descriptive statistics in Appendix G. We also used a single-item measure to cross-check the validity of the FunQ questionnaire 'During the activity I had fun'.

For addressing students' attitude towards the topic, students were asked to rate their agreement with the statement: 'I think that programming is my thing' ([1] not at all -- [5] absolutely). This single-item measure was adopted from earlier research [301] (see Figure 6.3) and was evaluated on a 5-point smiley face scale [108]. While some multi-dimensional scales exist for the measurement of primary school students' interest in programming (e.g., [151, 177]), they focus on various aspects (e.g., self-efficacy, utility, interest in programming). However, in this study, we aimed to measure attitude about programming as a global construct while keeping the length of the survey at a minimum for the sake of data quality taking into account children's attention span. Given that previous research [28] found that single-item measures are equally suitable for the measurement of concrete attributes as multiple-item measures, and using single-item measures for investigating attitudes about learning to program are not unusual [210], we favored the single-item measure. The same argument applies to the measurement of learning.

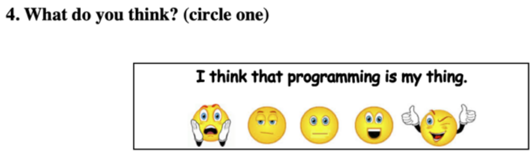


Figure 6.3 Questionnaire item assessing students' attitude about programming. Source of the representation of the 5-point scale is [108].

To measure learning we adopted the self-report measure from previous research [210], and asked students to answer on a 5-point Likert-type scale for the question 'Have you learned something new today about programming?'.

Table 6.3 The model components, their operational definition, and their respective measures.

Component	Operational definition	Measure
Fun	The degree to which students experienced fun during the activity.	FunQ and 'During the activity I had fun' ((1) Never - (5) All the time)
Attitude	The degree to which students indicate their attitude towards the subject.	'I think programming is my thing' ((1) Not at all - (5) Absolutely)
Learning	The degree to which students indicate their learning during the activity.	'Have you learned something new today about programming?' ((1) Not at all - (5) A whole lot)

6.3.5 Data Analysis

For the development of our model on the role of fun on learning, we used structural equation modelling (SEM) and mediation analysis [149]. Mediation analysis was performed to assess whether fun indirectly influences learning across its effect on students' attitude on the subject. For the data analysis the RStudio 1.1.453 [252] software, for modelling the lavaan [248] and psych [240] packages were used.

6.4 Results

6.4.1 Descriptive Results

To start with, we compared the data of the three workshops along the main elements of the hypothesized model: the experienced fun, the attitude and the reported learning. The descriptives are displayed in Table 6.4. One-way ANOVA analysis indicates no statistical difference in the data coming from the three workshops (p_{fun} = 0.985; $p_{attitude}$ = 0.737; $p_{learning}$ = 0.285) hence, analyzing them all together is justified.

Table 6.4 Comparison of the three workshops along the model components. No statistical difference is found.

	workshop1	workshop2	workshop3	sample
Fun	*M* = 70.06	*M* = 69.62	*M* = 69.68	*M* = 69.75
	SD = 10.18	*SD* = 6.587	*SD* = 9.280	*SD* = 8.283
Attitude	*M* = 3.96	*M* = 3.87	*M* = 4.07	M = 3.96
	SD = 1.107	*SD* = 0.957	*SD* = 0.900	*SD* = 0.974
Learning	*M* = 4.05	*M* = 3.84	*M* = 4.21	*M* = 4.02
	SD = 0.899	*SD* = 0.898	*SD* = 0.917	*SD* = 0.908

For calculating the FunQ scores, after reversing the scores of the *Stress* factor, we summed the values, resulting in an average score of 69.75 (*SD* = 8.283) from the possible range of 18-90. The internal consistency of the FunQ questionnaire appeared to be sufficient at both the first- and the second-order level ($\omega_{first\text{-}order}$= 0.721; $\omega_{second\text{-}order}$ = 0.778; ω > 0.6 is regarded as acceptable [107, 293]; for further statistics on the FunQ items, see Appendix G). Further, to safeguard the validity of the FunQ scores we calculated the correlation between those and the single-item measure 'During the activity I had fun.' (*M* = 4.56, *SD* = 0.859). The correlation coefficient indicates a positive and significant relationship (r = 0.422, p < 0.01), strengthening further our assumption about the reliability of the FunQ questionnaire and the measured values. Based on the aforementioned, we assume that students had fun during the workshops.

The reported average score for students' attitude towards programming measured after the activity is 3.96 (*SD* = 0.974) on a 5-point scale. The frequency of the responses is shown in Table 6.5 below.

Table 6.5 The frequency of the responses for 'Programming is my thing'.

missing	*(1)* *not at all*	*(2)* *a bit*	*(3)* *moderately*	*(4)* *much*	*(5)* *absolutely*
4 (4.7%)	2 (2.3%)	3 (3.5%)	19 (22.1%)	30 (34.9%)	28 (32.6%)

Students self-rated their learning on average 4.02 (SD = 0.908) on a 5-point scale, which translates to having learned 'much'. The frequency of the responses is shown in Table 6.6 below.

Table 6.6 The frequency of the responses for 'Have you learned something new today about programming?'

missing	(1) *nothing at all*	(2) *a bit*	(3) *something*	(4) *much*	(5) *a whole lot*
5 (5.8%)	0 (0.0%)	6 (7.0%)	14 (16.3%)	33 (38.4%)	28 (32.6%)

The descriptive statistics and the test of normality values for the model elements - fun, attitude, and learning - are displayed in Table 6.7.

Table 6.7 Descriptive statistics and test of normality of the elements of the model.

	Mean	Min.	Max.	Standard deviation	Skewness	Kurtosis	Shapiro-Wilk test of normality
Fun	69.75	44	84	8.283	-0.568	0.361	0.049
Attitude	3.96	1	5	0.974	-0.829	0.538	<0.001
Learning	4.02	2	5	0.908	-0.666	-0.309	<0.001

For assessing the strength of the relationship between the main elements of the model, we calculated the correlation coefficients. Table 6.8 displays the pairwise correlations among the experienced fun (measured by FunQ), the attitude and the reported learning. We conclude that all pairwise correlations are positive and significant.

Table 6.8 Correlation matrix of the model components.

	Fun	Attitude	Learning
Fun	-	0.388**	0.286*
Attitude	-	-	0.439**
Learning	-	-	-

*p-value = 0.022
**p-value < 0.01

6.4.2 *Path Analysis*

We applied path analysis to test the direct effects between the three main components of our model. The analysis revealed that the experienced fun, measured by FunQ, has no direct influence on the reported learning (p = 0.203, β_{std} = 0.136). Further, fun influences positively and significantly students' attitude about the subject (p < 0.001, β_{std} = 0.388). Regarding the effect of students' attitude about coding on the reported learning, our findings indicate a significant, positive relationship (p < 0.001, β_{std} = 0.386). The direct effects are displayed in Figure 6.4 below.

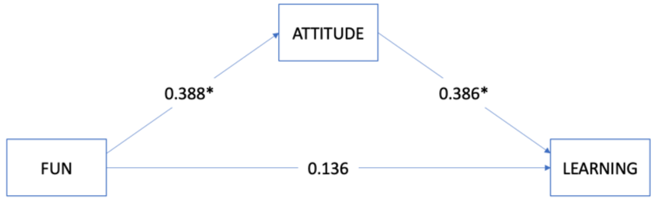


Figure 6.4 Path analysis: direct relationship between the components of the model. (*p < 0.05)

6.4.3 Mediation Analysis

To assess a possible indirect effect of fun on learning, we used mediation analysis. The analysis revealed that both the indirect effect of fun on learning across the attitude (β_{std} = 0.150, $p < 0.001$), and the total effect of fun on learning (i.e., indirect plus direct effect; β_{std} = 0.286, $p = 0.008$) are significant (see Figure 6.5). The model explains the 20.8% of the variance ($R^2 = 0.208$; $p < 0.001$).

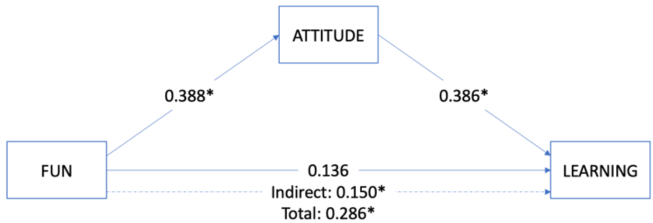


Figure 6.5 Mediation analysis: indirect relationship between fun and learning across attitude. (*p < 0.05)

6.5 Discussion

The relationship between motivation and learning has been researched extensively since the 1980s, but the identification of key emotions on learning is still at an early stage[112]. Studies on the effect of attitudes on learning appear to be controversial [162], and the role of the concept of fun on learning is still premature (e.g., [127, 210, 260, 275]).

Adopting FunQ [299] as a theoretical and measurement framework for fun during learning, our study has examined if and how fun influences students' learning to code. Our results showed that fun has a clear, positive, but indirect effect on the reported learning.

The herein introduced activity was playful by its nature as it was a non-formal learning activity using digital technology, following a learning-by-doing approach and building on students' voluntary participation [142]. During the workshops, students were invited to follow the video guide, which included playful elements regarding the design, the tone and the introduction of the tasks. Additionally, students were encouraged to use their creativity to solve the programming tasks, make the created artefacts unique and to play with those alone or together.

Evaluating the activities in general, we conclude that students experienced fun during the playful coding workshops, while the activities were meaningful in terms of learning, as students report on having learned much. Additionally, at the end of the workshop students expressed a positive attitude about programming as well.

Our study contributes to understanding the relationship between fun and learning, for which earlier research studies diverge. Some authors have found no significant correlation between fun and learning [127, 275], while others [210, 260] observe students having fun and learning without though directly investigating the relationship between the two.

In this study we hypothesized that i) fun has a positive direct effect on learning to code, ii) fun has a positive effect on students' attitude about coding, iii) students' attitude about coding has a positive effect on their learning, and iv) fun has a positive and indirect effect on learning to code as well.

Our hypotheses have been partially confirmed as we have not found a significant direct relationship between fun and learning to code. However, while previous work on the effect of attitude on learning is inconclusive [162], our results indicate a clear positive effect.

Regarding the final hypothesis, we found that fun has a significant and positive indirect effect on learning across attitude and that the total effect (indirect plus direct effect) of fun on learning to code is also significant. Our findings may explain why earlier studies [127, 275] have found no significant relationship between fun and learning, while they still report increased learning when students were observed having fun [210, 260]. The absence of a direct effect of fun on learning aligns with Sim, MacFarlane and Read [275] who did not find a significant correlation between neither the observed- nor the measured fun and learning, though they did find that children were more inclined to play again with educational software when they had more fun. We also concur with Iten and Petko [127] who found that the experienced fun while learning has a significant effect on motivation to learn and engage again with the learning game, but not on the reported- nor the measured learning. Our results extend those of the previous studies which only examined the direct relationship between fun and learning, despite having indications of a key intermediate element, namely the attitude. Further, our finding that fun does have a positive and indirect effect on learning is supported by previous work [210, 260], which observe students having fun and an increased learning. In those previous studies, however, the relationship between fun and learning was not investigated in depth.

In the herein introduced study we used self-reported measures with students at the edge of the formal operational Piagetian stage [218]. In child-computer interaction research - according to the theory of Piaget - age 11-12 is often considered as a strong border for using verbal questionnaires with children due to their state of cognitive development. However, in clinical psychology, the use of questionnaires with younger children is often difficult to avoid as those are used for assessment [187]. The study of Mellor and More [187] found that even young children are capable of giving reliable responses to Likert-type scales when certain criteria are met (e.g., scale labels reflect frequency of behavior/thoughts and not agreement; scale items refer to concrete situations/bodily feelings and not to abstract ones etc.). These findings are also supported by other studies on children as respondents in survey research [37, 38, 163]. While several design research studies in education include informal assessments of fun and learning, they do not allow quantifying the relationship between them (e.g., [103]). Therefore, in the herein introduced study we designed a survey with adhering to the above detailed criteria, and assessed students' attitude, knowledge and the fun they have experienced while learning by a self-report measure. We argue that the validity of the responses were not hindered by the format of the investigation (i.e., using self-reported measures), however, they allowed us to quantify the relationship between fun, attitude and learning, which would not have been possible with informal assessment methods (e.g., observation, qualitative measurements etc).

Our results support efforts of educational researchers and practitioners who try to make learning activities more fun for students. On the other hand, endorse the efforts of informal and non-formal learning environments - such as science museums, maker spaces and coding clubs - for making learning fun in order to engage students more and facilitate their learning process, as we see that when students are having fun while learning to code, then it has a significant effect not only on their attitude about coding but their learning as well. As an explanation, we propose that while fun has a positive effect on attitude, it serves as an enhancer for the willingness for learning about programming as well. In Glasser's words, fun *"is like a catalyst that makes anything we do better and worth doing again and again"* ([98], p. 28). This, we find especially important as programming is considered as the literacy skill of the 21st century [210], hence it should be in the interest of both educationalists and educational researchers to support the learning process of coding. Moreover, given that from previous research we know that students' success in computer science courses and their career choices can be affected by their attitudes towards programming [50], we find applying fun elements during learning to code crucial. Therefore, we strongly encourage both researchers and practitioners to utilize this property of fun especially when teaching or introducing to students such an important subject as programming.

6.6 Limitations and Future Work

In the herein introduced study we tested the role of fun on the reported learning, however, this bears with some limitations. To start with, we need to recognize that the assessment of learning is complicated. Accordingly, we cannot be sure what type of learning students considered when responding to the question (e.g., knowledge acquisition, learning new skills etc.), and the time constraint of the workshop could have played a role in the depth and extent of learning. The use of single-item measures, we believe, did not hinder the predictive validity given that both items measured concrete and simple constructs, nevertheless, in a future study more aspects of programming-related attitudes could be investigated with multi-item and multidimensional tools. Furthermore, we propose future research focusing on the eventual differences between the reported and the measured learning and the test of the herein introduced model's validity for the measured learning (a.k.a. knowledge acquisition) and other goals of learning (e.g., learning new skills) as well, including the investigation of learning during longer activities.

Another limitation of our research is that we did not investigate the effect of collaboration, which in future research could be an interesting angle to study.

Given that most of the participants were novices in the field of programming, novelty effect might have played a role in the magnitude of investigated aspects, but we believe it did not affect the relationships between those. Nevertheless, future research could investigate the effect of fun on learning in other scenarios where novelty effect is not present.

At last, investigating non-linear relationships between the components, including bi-directional effects in the future would contribute to a better understanding of the role that fun plays on learning to program.

6.7 Conclusion

In this study, we have examined the relationship between fun, attitude, and learning to code. Our results extend and explain earlier research suggesting that there is no direct relationship between the experienced fun and learning, as we found a positive indirect effect on learning across the attitude about programming and that the total effect (indirect plus direct effect) of fun on learning to code is also significant. This finding explains apparent contradictions in earlier studies where students were observed having fun and were found to learn more, while no significant association could be between fun and learning. Our results support efforts of educational researchers and practitioners who try to make learning activities more fun for students in the belief that making education - and especially learning to code - fun enhances students' learning. Therefore, this encourages us to work further in this direction to deepen our understanding on how fun influences not only the self-reported, but other levels of learning (e.g., the actual or measured learning), and to extend our research findings to different topics and age groups, with a special interest toward further personal and environmental influential factors.

Part IV.

PERSONAL AND ENVIRONMENTAL INFLUENTIAL FACTORS IN LEARNING

7 Self-Regulation and Fun in Learning – A Case Study[15]

In the previous part of this thesis we introduced the FiL model that quantifies the relationship between fun, attitude, and learning. In this part we aim to extend the FiL model by potential influential factors on learning. Accordingly, in this chapter we present a case study that investigates self-regulation (i.e., *"self-generated thoughts, feelings and actions that are planned and cyclically adapted to the attainment of personal [learning] goals"* ([346], p. 14)) as a possible personal influential factor on the previously established relationship between the FiL model elements. This chapter thus contributes with the examination of self-regulation as a potential extension to the FiL model.

Summary

Researchers have argued that self-regulation is an essential component for academic success. There is also substantial research effort invested towards enhancing fun in learning to improve student engagement and learning outcomes. However, little is known about the relations between self-regulated learning, emotions and learning outcomes. In this chapter we examine the relationship between self-regulation and fun in the context of a self-regulated cognitive training game. We hypothesized that a positive relationship exists between self-regulation, students' attitude about the topic, the experienced fun while interacting with the learning game and learning outcomes, including game performance and perceived learning. We collected data from 28 secondary school students before and after playing BrainHood, a self-regulated cognitive training game; and applied correlation- and path analysis to address the hypotheses. Our results suggest a positive and significant relationship between a) fun and learning, and b) fun and attitude towards the topic. However, c) no significant relationship was found between self-regulation and the other investigated dimensions. Our findings support efforts making learning more enjoyable; however, we challenge the long-standing general belief on the positive relationship between self-regulation and learning. We call on future research for investigating the relationship between self-regulation and fun in learning, with specific focus on the task-based (micro) level.

7.1 Introduction

In the last few decades, there has been substantial research effort invested worldwide to make learning enjoyable by using digital games for learning. Systemic reviews [54, 228, 284] on gamified learning, a.k.a. gamification, edutainment, digital game-based learning (DGBL), and serious games, conclude that gamification allows teaching systems to improve

---

[15] This chapter is based on the following publication: Tisza, G., Tsiakas, K., & Markopoulos, P. (2022). Exploring the relationship between self-regulation and fun in learning. Manuscript submitted for publication.

student engagement and motivation, and lead to increased performance. Typical areas for gamified learning are science education [141] and language learning [274], however, other fields such as business, marketing, tourism [59] and the health sector are also getting involved [143, 155, 341]. Accordingly, beyond entertainment and support in academic learning, gamification also acts as an enhancement to therapy and to promote health and well-being. Cognitive learning is defined by the Encyclopedia of the Sciences of Learning as a *"change in knowledge attributable to experience"* ([179], p. 594). One aspect of health and well-being is training cognitive skills for young [189] and old [170]. Related to learning, cognitive games, among others, can support the development of students' memory, spatial perception, and their executive functions as well. Executive functions, according to the Marriam-Webster dictionary, are sets of complex mental processes and cognitive abilities that control the skills required for goal-directed behavior[16]. These are also related to academic skills and competences [115] given that executive functions are the cognitive aspects of self-regulation. However, while it is suggested that students' self-regulation is an essential component of academic success [16], little is known about what role self-regulated learning skills and emotions play in the short term, during a specific educational activity or task, e.g., game-based or online learning. Plass et al. discussed properties of game-based learning and concluded that self-regulation is inherent to learning games as *"the player executes of strategies of goal setting, monitoring of goal achievement, and assessment of the effectiveness of the strategies used to achieve the intended goal"* ([224], p. 261). In connection to these properties of game-based learning, in this chapter we investigate two levels of learning according to Bloom's taxonomy [34]: perceived learning, which is linked to the *Evaluation* level (i.e., judgements about the value of the material for a specific purpose); and the game-based performance, which is linked to the *Application* level (i.e., use of abstractions in concrete situations). Azevendo et al. [18] summarized earlier research on self-regulation in various digital learning environments and concluded that i) the effectiveness of game-based learning is still contentious, and ii) it might be dependent on the self-regulatory skills required for processing information in the given digital learning environment. An important finding of theirs is that training self-regulation skills prior to using the digital learning environment led to increased learning effects, however, they do not provide exact values and ways of measurement. They also claim that *"despite the importance, there is a dearth of research studying the SRL [i.e., self-regulated learning] processes students employ or fail to employ during learning with these [i.e., digital] learning environments"* ([18], p. 590). Knowing more about the potential correlates of self-regulation and emotions related to learning, especially in the task-based or micro-level is essential to further contemporary research and our understanding of learning processes. Hence, in this study we set out to explore the relationship between self-regulation, learning

[16] https://www.merriam-webster.com/dictionary/executive%20function

outcomes, and fun in the context of game-based learning using a self-regulated cognitive training game.

In this chapter we propose an extension with the concept of self-regulation to the previously discussed FiL model (Chapter 6) capturing hypothesized relationships and testing them through the evaluation of Brainhood [310], a purpose made cognitive learning game for training executive function. Our model hypothesizes that students' self-regulation during learning will have a positive impact upon learning, and that itself will be positively affected by attitude and fun experienced in the learning activity (see Figure 7.1). The model was tested in an experiment, where we assessed the attitudes and performance of twenty-eight secondary school students. Our results confirm the expected relationship between fun and learning and between fun and the attitude towards the topic, but the hypothesized relations to self-regulation were not confirmed.

In the remainder of this chapter, we describe how related work leads to the hypothesized model, the design of the experiment that tested the model and the analysis of its results. We end the chapter with implications for further research on self-regulation, fun and game-based learning.

7.2 Background

7.2.1 Self-Regulation and Learning

Self-regulation is defined by Schunk and Ertmer as *"self-generated thoughts, feelings, and actions that are planned and systematically adapted as needed to affect one's learning and motivation"* ([267], p. 631). Pintrich [221] adds that self-regulated learning, in contrast with external regulation, is an active and constructive process where students define the learning goals for themselves, meanwhile they also regulate, monitor and control their cognitive and motivational processes in order to reach their goals.

It has been known from the field of educational psychology that students with higher motivation tend to reach higher academic success than their less motivated mates [16]. According to Flavell [86], a possible explanation therefore lies in the use of cognitive and metacognitive strategies. Students with higher motivation tend to use a higher variability of those strategies, helping them learn more efficiently. Pintrich [221] goes even further and concludes that highly motivated and self-regulated learners make the most academically successful students. Accordingly, we hypothesized that self-regulation positively influences learning outcomes, and that self-regulation is positively influenced by student's attitude about the topic.

Zhou et al. [345] investigated the relationship between online self-regulated learning and perceived learning gains. Their results indicate that online self-regulated learning (measured by the Online Self-Regulated Learning Questionnaire; OSLQ [21]) is significantly related and positively influences the perceived learning gains. However, they did not investigate the possible effect of emotions. Verstege et al. [316] investigated the relationship between self-regulation and learning outcomes in a gamified learning environment. They found a non-linear relationship between self-regulated learning and learning outcomes, as the medium self-regulated learners learned the least as opposed to the low and high self-regulated learners. These results were obtained by measuring game-specific learning, reflecting the performance in the game rather than the class grade by which the teacher assesses the learning outcomes at the end of the semester. Nevertheless, these results further support our hypothesis regarding the positive relationship between self-regulation and learning outcomes.

Regarding the relationship between emotions and self-regulated learning, Pekrun, Goetz, Titz, and Perry [215] found that academic emotions (e.g., enjoyment, hope, pride, anger, anxiety, boredom) in general, and enjoyment in specific, are significantly correlated with self-regulated learning. However, they did not investigate the direction of the relationship, nor did they investigate learning outcomes. In line with them, Artino and Jones [15] tested the relationship between self-regulated learning behavior and emotions (among others enjoyment) in the online learning setting. Their study results are consistent with those of Pekrun et al. [215] as they found that enjoyment is significantly associated with self-regulated learning behaviors, however, they also did not investigate learning outcomes. According to these findings, we hypothesized that having fun while learning will positively influence self-regulation.

Recently, An et al. [8] investigated the relationship between self-regulated English learning, enjoyment and learning outcomes. Their research results indicate that enjoyment significantly influences students' self-regulation, and that self-regulation significantly influences students' learning outcomes (measured by a national English language test score). They also found a significant indirect relationship between enjoyment and learning outcomes. These findings strengthen our hypothesis that the experienced fun while learning positively influences students' self-regulation, and led to our next hypothesis, namely, that fun also indirectly affects learning across self-regulation.

Contrary to these results, Villavicencio and Bernardo [318] did not find a relationship between self-regulation and the final grade (as a measure for learning), but, they found an interaction effect of self-regulation and enjoyment: for students who reported a higher level of enjoyment, self-regulation was positively associated with grades. But for students who experienced lower level of enjoyment, self-regulation was negatively related to grades. The authors speculate that high levels of perceived positive emotions reinforce students' sense of control over the learning situation, and therefore boost the benefits of self-regulation. Low level of positive emotions, on the other hand, might be an indicator for negative task value and eventually negative outcome appraisal, which leads to the impairment of the often-described positive effect of self-regulation on learning outcomes. Therefore, Villavicencio and Bernardo [318] challenge the widely accepted positive effect of self-regulation on learning, and call for its further qualification as this effect, they argued, can be undermined by experiencing low levels of positive emotions while learning.

They conclude that *"the experience of positive emotions may be a necessary condition for the positive relationship between self-regulation and achievement to be obtained"* ([318], p. 338.). Their study is among the very few who examined the possible interaction between emotions, self-regulation and learning outcomes. *"We should underscore that there has been very little evidence about the moderating effect of academic emotions on the relationship between cognitive–motivational variables on the one hand and learning and achievement on the other"* ([318], p. 337). Nevertheless, their study pertains to a few academic emotions (enjoyment and pride), and hence, it does not extend to the fun experienced while learning, leaving space for further examinations.

What appears to be typical for the studies investigating the relationship between self-regulation and learning is that most studies used the course grades as a performance measure for learning [16]. Accordingly, we know little about the relationship between self-regulation and short-term learning (i.e., the influence of self-regulation on the performance of a given learning activity or task). Some studies in the field of robotics [23, 136] investigate short-term interactions between students and Intelligent Tutoring Systems (ITS), aiming to design personalized support for self-regulated learning (SRL) skills during an educational activity. In such comparative studies, the authors aim to investigate the effects of SRL support by assigning each student to either SRL support or no support. Results of such studies [23, 136] suggest that providing *personalized* support of self-regulated learning skills to students can improve students' learning and affective outcomes, compared to no such support. In our study, we investigate self-regulated learning and experienced fun in the context of a self-directed educational activity (brain game) which provides SRL support through game features; however, where the student has full control regarding to level of SRL support they want to get, by using specific game features.

7.2.2 Game-Based Learning

Game-based learning, also known as digital game based learning (DGBL), gamification, gamified learning, serious games and edutainment, are notions that all refer to a recently emerged phenomena, namely, teaching through and acquiring knowledge and skills by playing engaging educative computer games [229]. A recent literature review [40] found that game-based learning is most frequently applied in relation to science, technology, engineering, mathematics (STEM) and health-related topics. They also found that the main aim of such games was knowledge acquisition, but other goals also appeared such as aiming for behaviour change or perceptual, affective, cognitive and physiological outcomes. Game-based learning builds on the idea that learning games resemble free-time activities such as playing video games, hence it is fun and intrinsically motivating [205]. However, while intrinsic motivation is a well-defined concept, the same is not true for fun.

7.2.3 Fun, Attitude, and Learning

As discussed in earlier chapters, in academic literature, fun is often handled as a *commonsense* term, and accordingly, there is a *"lack of conceptual clarity in the literature concerning the nature of fun"* ([184], p 160). Without a commonly accepted definition, measurement is complicated and as a consequence, research results might be divergent. To overcome this lacuna, we developed a theoretically grounded definition of fun experienced during learning (see Chapter 2) and introduced a questionnaire for its measurement (see Chapter 3). According to our previously established definition, fun is an affective state that is dominated by positive emotions while negative emotions are limited, during which one feels in control of the activity and is intrinsically motivated for participation, is immersed in the experience by losing sense of time and space, letting go of social inhibitions, while an optimal level of challenge is present - meeting the level of one's skills.

From educational psychology we know that students are more keen to invest time and effort into a learning activity that is interesting and enjoyable compared with activities that are dull and stressful [90]. In accordance with this, previous studies indicate a possible positive relationship between fun and learning [51, 171, 232, 296, 317, 327], however, this relationship has only occasionally been directly investigated, more frequently it is reported as an observation.

The FiL model introduced in Chapter 6 proposed a positive relationship between fun, attitude, and learning, which relationship got supported by our quantitative research results. Based on this previous research in the current study we hypothesized that the FiL model is valid for cognitive learning as well, and accordingly, fun has a positive direct effect on students' attitude about the topic, students' attitude has a positive direct effect on learning, fun has a positive direct effect on students' learning, and fun has a positive, indirect effect on students' learning across students' attitude about the topic.

7.2.4 Brain Games

Brain games (a.k.a. cognitive games, cognitive educational games, or cognitive training) are problem solving activities that aim to improve one's cognitive skills and capacities by requiring players to look for patterns and pay attention to details. Education psychologists argue [16] that successful learners have better mental capabilities (i.e., cognitive and executive functions) than their less successful mates. Accordingly, brain games *"focus on enhancing cognitive functioning in children with different profiles of executive functions and cognitive development"* ([311], p. 521). Brain games aim to assess and train cognitive and executive functioning, and they have been successfully used with an educational purpose previously with typically and atypically developing children [11, 325, 328]. Brain games are used as educational tools to monitor and enhance cognitive skills related to learning and not to help the student to acquire a specific item of knowledge.

7.2.5 Hypotheses

The purpose of the study was to explore the relationship between self-regulation and fun in the context of a cognitive training game. Hereby we collect the above formulated research hypotheses in a row. The hypothesized relationships are depicted on Figure 7.1:

- **H1**: The FiL model is valid for cognitive learning as well.
  - **H1a**: Fun has a positive direct effect on students' attitude about the topic
  - **H1b**: Students' attitude has a positive direct effect on learning
  - **H1c**: Fun has a positive direct effect on students' learning
  - **H1d**: Fun has a positive, indirect effect on students' learning across students' attitude about the topic.
- **H2**: Self-regulation positively influences learning measured by task-based performance and perceived learning.
- **H3**: Fun positively influences self-regulation.
- **H4**: Student's attitude about the topic positively influences self-regulation.
- **H5**: Fun indirectly effects learning across self-regulation

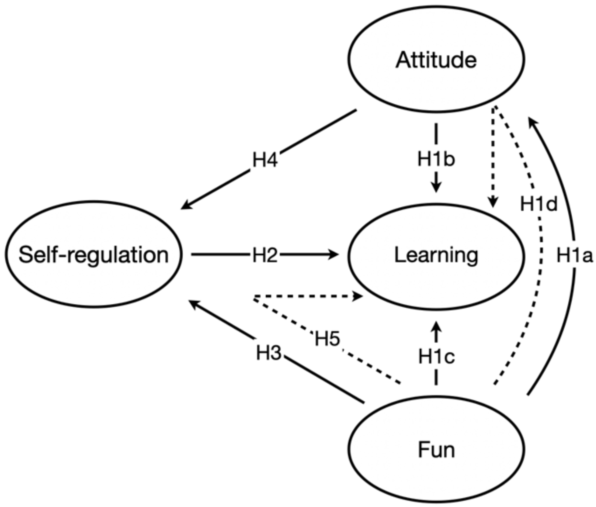


Figure 7.1 Hypothesized relationships between the model elements. Straight line indicates a direct relationship. Dashed line indicates an indirect relationship.

7.3 Method

To test our hypotheses, we carried out an experimental study with a pre-test post-test design at a school environment, where two classes of students used a cognitive training game, and we assessed the variables of the model through questionnaires and logging their performance in the game. The methodology is further detailed below.

7.3.1 *Participants*

We carried out the study in June 2021, a period in which schools were open and physical presence at the school was allowed during the Covid-19 pandemic. A schoolteacher from a secondary school scheduled the study as an extra-curricular activity for two of her classes during school hours. Participation in the study was voluntary for all students; students not wishing to participate would be otherwise occupied but stay in the classroom. The study was approved on 28 May 2021 by the Ethics Review Board of Eindhoven University of Technology, Department of Industrial Design. The participants were above age 16, hence participants provided informed consent and no parental consent was needed according to the applicable Dutch regulations. Two students decided not to take part in the study. The results below are based on the collected data from two secondary school classes, including a total of 28 students (7 boys, 18 girls, 3 not given; M_{age} = 16.4, SD = 0.50).

7.3.2 *Procedure*

For the study, each student was equipped with a computer with internet access. At the beginning of the study, the researchers briefly introduced the game and its rules, and students were allowed to ask their questions. Before playing the game, students were asked to fill in the online pre-game questionnaire (anticipated response time is 5 minutes) and watch an introduction video for playing the game. Thereafter, they were directed to the online platform to play the game (https://brainhood-test.netlify.app/#). Figure 7.2 depicts the study setup and a student interacting with the game. Students could spend approximately 30 minutes on the game, after which they were asked to fill in the post-game questionnaire (anticipated response time is 5 minutes).

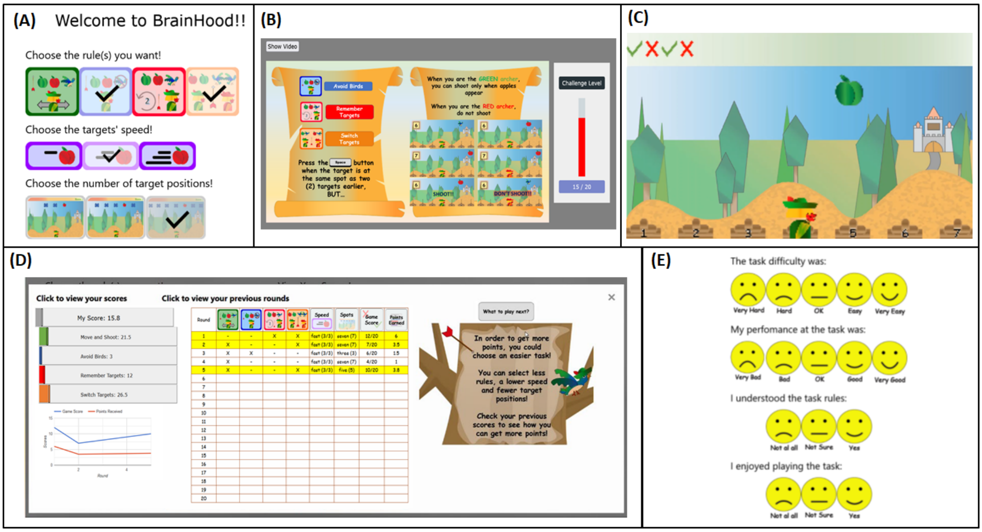
Figure 7.2 Study setup. A student interacts with BrainHood.

Before playing the game, we asked students to report on their knowledge and experience with brain games. In the pre-game questionnaire, we also recorded students' self-regulated learning, and in the post-game questionnaire we recorded students' attitude towards the topic, the fun they have experienced and their perceived learning. During the game students reported after each round on the level of fun and the perceived performance (in-game data), and we also recorded their game scores (log data).

7.3.3 BrainHood

To test the hypotheses, we used BrainHood, a self-regulated cognitive training game [310] designed to enhance the effects of cognitive training by supporting the player's self-regulation skills. Despite the usability of the game has not yet been previously reported, the design of this game emphasizes on *system transparency* and *user autonomy*, as game design features for self-regulated cognitive training (see Figure 7.3). More specifically, game features support players in monitoring their performance and managing their training regimen. BrainHood is a simple shooting game, where self-regulation plays an important role to achieve a good game performance. Namely, the game is designed to elicit gameplay behaviors related to three self-regulation components: *task selection* strategies (select the appropriate task to reach their goal), *self-efficacy* (making accurate self-assessment of the game skills), and *goal setting* (set a target score for a session). Players can select from four types of shooting rules (further: rules). These rules can be used either individually or combined (15 possible combinations), which can provide a wide range of task complexity and difficulty. Players can also adjust the speed (3 levels) and the number of targets (3 levels), allowing for a total combination of 15 x 3 x 3 = 195 different setups. Players can select appropriate task combinations to maximize the points they can collect. It is expected that players with higher self-regulation can optimize the game setup and adjust it better to their improving skills in comparison with players having lower levels of self-regulation.

Figure 7.3 (A) BrainHood rules and difficulty selection menu. (B) Description of task rules and challenge level. (C) Screenshot of BrainHood game. (D) Open Learner Model and Task Recommendation. (E) In-game survey.



7.3.4 Measures

To assess knowledge and experience with brain games, students responded to two 5-step Likert-style questions: 'Do you have any idea what a brain game is?' and 'How many brain games have you played before?'.

We used several measurement tools to assess the fun students experienced while playing the game, their attitude about the topic, their self-regulation, the perceived learning, and game performance (see summary in Table 7.1).

For assessing self-regulation, we used two validated questionnaires: Children's Perceived use of Self-regulated Learning Inventory (CP-SRLI; [314]) and the Motivated Strategies for Learning Questionnaire (MSQL; [223]). Both questionnaires are evaluated on a 5-point Likert-type scale (1 – Never, 2 – Rarely, 3 – Sometimes, 4 - Often, 5 – All the time). From CP-SRLI we used the Planning (PL), the Self-efficacy regulation (SER) and the Monitoring (MT) subscales. Cronbach's alpha values indicate an acceptable internal validity on our data (α_{PL} = 0.725; α_{SER} = 0.825; α_{MT} = 0.690). Along with the more general SER subscale, we selected the other two subscales as they reflect on the required skills to successfully play BrainHood: in the game, players need to follow a goal-setting strategy (addressed by the PL subscale) and need to monitor their progress to achieve the highest score possible (addressed by the MT subscale). From MSQL we used the Metacognitive Self-regulation (MSR) subscale. The internal validity testing resulted in an acceptable value for MSR as well (α_{MSR} = 0.731).

For assessing fun, we used FunQ. FunQ is evaluated on a 5-point Likert-type scale (1 – Never; 5 – All the time). To cross validate FunQ, we also recorded the validated Enjoyment in Math scale (EiM; [49]), which was adjusted to fit the study purpose (e.g., 'I enjoy doing math' was rephrased onto ' I enjoy doing brain games'). Both questionnaires have a reliable internal consistency on our data set (α_{FunQ} = 0.875; α_{EiM} = 0.845) and the correlation between the two measures is significant (r = 0.773, p < 0.001). Additionally, during the game, after each round students were asked by the interface to report on how much fun the previous task was, ('I enjoyed playing this task'), for which we used a 3-point smiley face scale, (a.k.a. Smileyometer [233]), a widely used instrument for assessing experiences with interactive products.

For the assessment of learning, we used both a self-reported measure (i.e., perceived learning) and the performance in the game to assess different levels of learning. According to Bloom's taxonomy of learning [34], we link perceived learning to the *Evaluation* level; while the game-based performance is linked to the *Application* level. For the perceived learning, students indicated their agreement on a 5-point Likert-type scale across three items (1 – Totally disagree; 5 – Totally agree). Perceived learning is considered as a good complement to other types of learning measures as it has the potential to capture learning aspects that other, more exact measures fail to capture. The internal consistency of the perceived learning items is adequate ($\alpha_{perceived\ learning}$ = 0.767) on our data set. For measuring achievement in the game, we used two measures. The *game score* (i.e., the number of correct targets in a given round) and an adjusted score based on the total correct targets and the difficulty level in a given round (i.e., *game performance*). The game performance was calculated as follows:

Game performance = correct_targets x number_of_rules x speed_level x target_positions

Where *correct_targets* is the number of the correct targets for each round [0-20], the *number_of_rules* is a factor related to the number of selected rules [1 rule: 0.25, 2 rules: 0.5, 3 rules: 0.75, 4 rules: 1.0], and the *speed_level* and *target_positions* are factors related to the speed and target positions selected [level 1: 0.8, level 2: 0.9, level 3: 1.0]. The scoring approach aimed to reward appropriate selections which match the player's abilities. Players could get good scores if they selected tasks they can perform well. Maximum *game performance* could be achieved if the student hit all correct targets and avoided hitting wrong ones (20 points) at the most difficult level (4 rules, maximum speed, and maximum number of targets). Minimum score was achieved if the student missed all targets at any level. Additionally, after each round in the game, students were asked to report on their perceived performance on the task (i.e., perceived performance; 'My performance at this task was…' (1) Very bad – (5) Very good on a 5-point smiley face scale (a.k.a. Smileyometer [233]).

For measuring students' attitude about the topic, we used a single item measure ('I think that brain games are my thing'; AMT) over which previous studies indicate a suitable and reliable use for the measurement of students' general attitude about subjects [300, 301, 306]. Additionally, we recorded students' attitude on six specific dimensions about brain games ('Do you think that brain games are fun / easy to do / easy to understand / pleasant /

exciting / something I want to do again?; AS). All items were evaluated on a 5-point Likert-type scale (1 – Not at all; 5 – Absolutely). On our data set Cronbach's alpha indicated an adequate internal consistency for the questions on the specific attitude dimensions (α_{AS} = 0.752) and a significant correlation with the single general attitude item AMT (r = 0.750, p < 0.001). Additionally, to cross validate the aforementioned, self-generated items, we recorded the previously validated Brief Scale on Attitude Towards Learning of Scientific Subjects (ATLoSS; [49]), for which we rephrased the items to fit the study purpose (e.g., 'There should be more hours of scientific subjects at school' was rephrased onto 'There should be more hours of brain games at school'). The items are evaluated on a 5-point Likert-type scale (1 – Totally disagree; 5 – Totally agree). The internal validity value of ATLoSS is acceptable on our data (α_{ATLoSS} = 0.875).

Table 7.1 Study dimensions, their operational definition, and their respective measures. All items were evaluated on a 5-point scale (1 – Totally disagree /Never; 5 – Totally agree / All the time).

Component	Operational definition	Measure
Self-regulation	The degree to which students can self-regulate.	CP-SRLI [314] and MSQL [223]
Fun	The degree to which students experienced fun during the activity.	FunQ [299] and EiM [49], and in-game self-report 'I enjoyed playing this task'
Perceived learning	The degree to which students indicate their learning during the activity.	'I got better by playing the game.' 'I felt that I was training my brains.' 'I learnt new skills today.'
Perceived performance	The degree to which students indicate the level of their perceived performance after each round in the game.	In-game self-report 'My performance on this task was… (1) Very bad - (5) Very good'
Game score	The final score collected in the game.	The number of the total correct targets.
Game performance	The adjusted game score, taking into account the total correct shots and the difficulty of the task selection (i.e., rules, speed, and position).	Game score adjusted by the difficulty of the selected rule(s)
Attitude	The degree to which students indicate their attitude towards the subject.	ATLoSS [49] and 'I think that brain games are my thing.' (AMT) 'Do you think that brain games are fun / easy to do / easy to understand / pleasant / exciting / something I want to do again?' (AS)

7.3.5 Data Analysis

The descriptive data analysis and the correlation analysis was done with SPSS Statistics software version 27.0.0. For the path analysis RStudio 1.1.453 [252] software, and the lavaan [248] and psych [240] packages were used.

7.4 Results

7.4.1 Descriptive Results

For the data analysis we used data from two sources: the pre- and post-game surveys and the in-game data.

The responses to the questions regarding knowledge and experience with brain games indicate that students had some idea about the topic ($M = 2.08$, $SD = 0.812$) which translates to knowing a bit. 13 students reported not to have played a brain game before and 12 to have played five games or less prior to the experiment (3 responses were missing; $M = 1.92$, $SD = 0.997$). We concluded that students had some anticipation about brain games but could safely be considered as novice players.

To address participants' self-regulation, we calculated the mean scores on the recorded dimensions of CP-SRLI and MSQL. Students' mean scores are as follows on the CP-SRLI dimensions: Planning (PL; $M = 3.49$, $SD = 0.83$), Self-efficacy regulation (SER; $M = 3.16$, $SD = 0.64$) and the Monitoring (MT; $M = 3.28$, $SD = 0.59$). Students' mean score on the Metacognitive Self-regulation (MSR) dimension of MSQL is 2.92 ($SD = 0.47$). All four subscales are significantly correlated with each other ($p < 0.001$). These average values translate roughly to students 'sometimes' having the investigated self-regulatory behaviors and thoughts.

To assess the experienced fun while playing the game we calculated the sum score on FunQ, correcting for the reversed items. From the possible range of 18 to 90, the average score students reported on FunQ was 55.81 ($SD = 10.70$). Independent sample t-test indicates no significant gender difference in the FunQ values ($p = 0.458$). Based on this result we conclude that boys and girls evaluated the game evenly and had a moderate level of fun while interacting with it. Additionally, students also reported on a 3-point scale on the fun they have experienced after each round in the game. The mean of the in-game fun score is 2.46 ($SD = 0.29$), which also indicates a moderate level of fun.

The mean score for students' perceived level of learning is 3.04 ($SD = 0.94$) on a 5-point scale, which translates into having learnt something during the game. Regarding performance, based on the in-game performance data, students self-rated their performance on average as 3.30 ($SD = 0.38$) on a 5-point scale. The game score ranges from 34 to 259, with a mean of 161.89 ($SD = 51.88$). The adjusted game score ranges from 5.8 to 123.3, with a mean of 42.22 ($SD = 20.38$).

Regarding students' attitude about the topic, we utilized three measures. To start with, we calculated the mean score on ATLoSS (2.75, $SD = 0.97$), followed by the mean on the simple item measure 'Brain games are my thing' (2.78, $SD = 1.19$) and the mean score on the specific attitude dimensions (3.19, $SD = 0.63$). These results suggest that students had a moderately positive attitude about brain games. All three measures are significantly correlated with each other ($p < 0.001$).

7.4.2 Survey Data

7.4.2.1 Fun, Learning and Attitude (H1a-H1d)

To start with, we calculated the correlation between the FunQ score, students' perceived learning (mean of the 3 items) and their attitude toward the subject (ATLoSS, single item measure, and average on the 6 items spec. dimensions) to address the fitness of the FiL model in the cognitive learning domain. All pairwise correlations are positive and significant ($p < 0.05$; see Table 7.2).

Table 7.2 Survey data correlation coefficients between fun, learning, and attitude. All correlation coefficients are significant (p < 0.05).

	Fun/Enjoyment		Learning	Attitude		
	FunQ	EiM	Perceived learning	ATLossS	AMT	AS
FunQ	1					
EiM	0.773	1				
Perceived learning	0.564	0.665	1			
ATLoSS	0.730	0.839	0.667	1		
AMT	0.593	0.798	0.444	0.763	1	
AS	0.802	0.924	0.640	0.802	0.750	1

7.4.2.2 Self-Regulation and Fun, Learning and Attitude (H2-H5)

To explore the relationship between self-regulation, fun, learning, and attitude, we calculated the correlation between the CP-SRLI and MSQL dimensions as indication for self-regulation, the FunQ score, the enjoyment of brain games (EiM; mean of 3 items), students' perceived learning (mean of the 3 items) and their attitude toward the subject (ATLoSS, single item measure (AMT) and average on the 6 items specific dimensions (AS)). All pairwise correlations were non-significant ($p > 0.05$) indicating no link between self-regulation and fun or perceived learning or attitude.

7.4.2.3 Path Analysis

To address the research hypotheses, we conducted a path analysis. The analysis results support the findings of the correlations analyses, as we found no significant relationship between self-regulation and either attitude, or fun or learning. However, we found a significant and positive association between fun and learning ($p = 0.047$) and fun and attitude ($p < 0.001$). The total effect of fun on learning is also significant ($\beta_{std} = 0.660$, $p < 0.001$; for all standardized factor loadings see Figure 7.4).

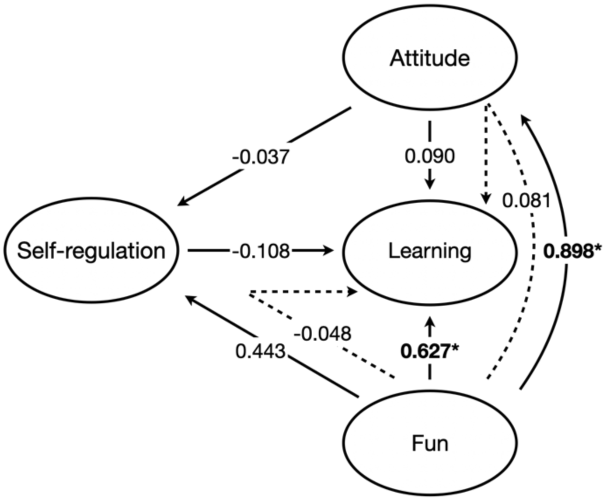


Figure 7.4 Path analysis findings: standardized coefficients. * indicates a significant relationship (p < 0.05). Straight line indicates a direct relationship. Dashed line indicates an indirect relationship.

7.4.3 In-Game Data

To further investigate the possible relationship between self-regulation, fun and learning, we analyzed the in-game data with respect to the in-game self-reported fun, the in-game self-reported performance (as a subjective indicator for learning), the game scores (as an objective indicator for learning, obtained from game logs) and self-regulation (i.e., CP-SRLI and MSQL dimensions). Pearson's correlation analysis resulted in non-significant values ($p > 0.05$) for all pairwise correlation between the self-regulation subscales and perceived fun or perceived performance or game score or game performance. Hence, no path analysis was conducted.

7.5 Discussion

The aim of the study was to investigate the relationship between self-regulation and fun in relation to learning in the specific field of cognitive games, and to possibly extend the previously presented FiL model with self-regulation as a personal influential factor on learning. To explore these relationships, we selected a cognitive game called BrainHood [311], which provides players with a full autonomy to explore and utilize the game features to achieve high scores, this way promoting self-regulation. Accordingly, it was expected that students with higher levels of self-regulation will perform better than students with lower levels of self-regulation. Additionally, based on earlier research we expected fun to have a positive impact on students' attitude about the topic and on students' learning (H1) and to have a positive indirect effect on learning across self-regulation as well (H5). Also, we hypothesized a positive influence of self-regulation on learning (H2), and a positive effect of both fun (H3) and attitude (H4) on self-regulation. Our research results partially support our hypotheses.

Regarding the relationship between fun, attitude, and learning (H1) we found that the experienced fun while interacting with the game had a significant positive effect on both students' attitude about the topic (H1a) and on their perceived learning (H1c). These results are aligned with our study presenting the FiL model (see Chapter 6), where we reported on similar findings in relation to a playful programming activity. Additionally, these findings are also in line with our earlier findings in relation to digital game-based learning (see Chapter 5) and with the findings of Connolly et al. [58] who found a positive association between fun and learning in the digital game-based learning environment. However, our findings contradict earlier research [300] in the sense that the current study did not find a significant relationship between students' attitude and their perceived learning (H1b), and we also did not find an indirect effect of fun on learning throughout students' attitude (H1d). As for the effect of attitude on self-regulation (H4), we expected a positive association based on the work of Pintrich [221], in which he claimed a positive association between high motivation and self-regulation. However, we did not find such an association. Despite attitude and motivation are two notions that are incrementally linked to each other [167], the difference between the two could have caused the contradictory finding. Other, possible explanations lie in the different domain of the studies, the study design, the number of study participants or the game itself, which may have prevented the expected relationship between the constructs to be manifested. Therefore, these findings should be further examined in next studies for a thorough understanding of the relationship between attitude, self-regulation and learning in the specific field of cognitive games.

For the assessment of the relationship between self-regulation and students' learning (H2), we used questionnaire data (i.e., perceived learning) and in-game data (game performance and perceived performance as indicator for task-based learning) as well. Nevertheless, neither the questionnaire data, nor the in-game data indicated any significant relationship between students' learning and their self-regulation. These research results are in line with those of Villavicencio and Bernardo [318] who did not find a relationship between self-regulation and learning. The results are partially in line with those of Verstege et al. [316] who did not find a linear relationship between self-regulation and a learning, but they found a nonlinear relationship between self-regulation and learning - which has not been investigated in the current study. Yet, our findings further earlier research as we distinguished two levels of learning, and accordingly, investigated both of those; and we conducted our study with secondary school students (in contrast with earlier studies, which focused on undergraduate students) this way providing more details onto the understanding of the learning processes. However, our findings contradict earlier findings of An et al. [8] who found a positive relationship between self-regulation and learning, and the findings of Zhou et al. [345] who found a positive relationship between online self-regulation and perceived learning. Among this previous research, however, there were substantial differences that could explain the discrepancy. While all studies were conducted with undergraduate students, Villavicencio et al. [318] and Verstege et al. [316] examined science-related subjects (trigonometry and enzymology), An et al. [8] investigated language learning, and the sample of Zhou et al. [345] consisted of students from various study disciplines. Therefore, it is possible, that the hypothesized positive relationship between self-regulation and learning is dependent on the subject. This, we argue, might be related to the earlier discussed cognitive processes as processing new information from various subjects could require different cognitive strategies. Another difference between the referred studies is that the study of Villavicencio et al. [318] and An et al. [8] took place in the traditional learning environment, while the study of Zhou et al. [345] investigated specifically students' online self-regulated learning during the COVID19 pandemic. The study of Verstege et al. [316] was the closest to our study setting, as they investigated the learning by utilizing a virtual experiment environment (thus a digital learning platform), and in their study, students were also required to apply the freshly obtained knowledge immediately withing the digital game. Obtaining similar outcomes as them (but with a different age group and discipline) further strengthens our research findings. Nonetheless, as it is reflected in the aforementioned, the positive association between self-regulation and learning should not be taken granted, as it might vary across different fields and different levels of learning, and we call on future research for a deeper understanding of this topic.

With respect to the relationship between fun and self-regulation (H3) our study findings indicate no significant association. This finding is in contrast with the earlier findings of An et al. [8], Artino et al. [15] and Pekrun et al. [215], who found a positive association between self-regulation and enjoyment. A possible explanation to the contradicting fining might lie in the difference in the nature of the evaluation of enjoyment. An et al. [8] investigated general enjoyment of a subject (i.e., students' liking for learning English as a foreign language), Artino et al. [15] investigated students' course-related enjoyment (in the online learning environment), and Pekrun et al. [215] investigated learning-related and class-related emotions. Accordingly, neither of these studies investigated enjoyment in the setting of a single learning activity, but at a more general (course-) level. Therefore, the granularity of the earlier research and our study differs in merit. We argue that the effect of self-regulation on a single-learning task is more difficult to ascertain than the effect of self-regulation on a longer learning task (i.e., course), as in the latter case there is more time given for the self-regulatory skills to unfold. We also suggest that task-related (and not general) emotions can affect self-regulation and learning at a different level or in different ways. Hence, there can be a difference between e.g., the enjoyment of English learning in general, and the enjoyment of a specific English learning task. Accordingly, to further our understanding of learning processes, future research should examine the relationship between emotions and self-regulation in learning at the task-based level.

Regarding the expected indirect effect of fun on learning across self-regulation (H5) we also contradict the previous findings of An et al. [8] as we did not find a significant relationship. The above-named arguments, highlighting the differences in the granularity of the measurements, are also valid as reasoning for this. On top of that, we find it important to note that all these previous findings investigated the academic emotion *enjoyment*, which is conceptually close to the concept of *fun*, however, is not completely overlapping, as fun is rather an emotional experience or a mental state than a single distinguished emotion. Hence, for a better understanding on the relationship between self-regulation, fun and learning, and the difference between fun and enjoyment, further research is required.

7.6 Limitations and Future Work

As indicated above, the study has certain limitations. The experiment, by default, was not controlled. Students were asked to interact with the cognitive game and its features, which means that their interaction was of exploratory nature. Based on the comments we received after the study, we conclude that the understanding of the game rules and its features took time for many players. Accordingly, future research could replicate the study in a controlled environment.

Furthermore, the validity and generalizability of the results from the path analysis are limited by the sample size. Therefore, testing the herein introduced effects on a bigger sample would contribute to the generalizability of the results.

Additionally, to further contemporary understanding of learning processes, research could focus on the development and deployment of similar, flexible recommendation systems for a better understanding of the relationship between self-regulated learning and emotions.

Another possible direction for future work would be to assess whether and how students used the game features to optimize the settings to reach higher scores. While students had to select the game setting for each round, in the herein introduced study we did not examine whether they have *actually* showed self-regulatory behavior (reflected in their setting selection and its relation to the achieved points). This is planned to be investigated in a follow-up study.

7.7 Conclusion

While it is generally held that self-regulation plays an important role in learning, our contradicting findings call on further research. As Villavicencio et al. [318] stated, we know little about the effect of academic emotions on the relationship between self-regulation and learning. Our study, to our best knowledge, is the first of its kind to look outside of the scope of academic emotions, and investigate the relationship between self-regulation, fun and learning. Moreover, we did not only go outside the scope of the academic emotions but zoomed into the setting of a single learning activity, compared with the generally studied course-level effects. Despite our results are partially contradicting those of earlier scholars, we must state that the comparison is difficult as we investigated the special domain of cognitive learning, using perceived learning and task-based performance as an indication for learning as opposed to the generally used course grade [16], and investigated the experienced fun while being busy with a single learning task, instead of the general enjoyment of a course or topic. Especially given these properties, our study results complement those of earlier findings by bringing in a new research angle, namely, the investigation of fun at the learning activity level. Additionally, we contribute to the literature by testing the fit of the FiL model in a new domain. This way, providing supportive evidence towards the generalizability of the model into different learning setups.

8 SES and Fun in Learning[17]

In the previous chapter we investigated whether the earlier introduced FiL model (Chapter 6) could be nuanced by the personal factor of self-regulation. In the current chapter we aim to investigate another potential extension to the FiL model that could weigh on the role and effect of fun in learning. Namely, we examine students' socioeconomic background, as an environmental influential factor. This chapter, accordingly, contributes with the extension of the FiL model with students' socioeconomic background.

Summary

Programming and creative thinking are important skills for the 21st century. A large body of evidence suggests that a playful approach to learning helps students engage deeply with programming, improves their creative thinking skills, and shapes a positive attitude towards programming. However, such research has rarely considered how differences in socioeconomic background impact the way students experience such programming activities. The theoretical perspective of science capital suggests that students from high income families will hold more positive attitudes towards science and technology and will perform better in programming than students from lower income families based on their generally higher exposure to experiences involving computing technology. To examine this assumption, we designed and implemented single-occasion programming workshops lasting two hours that followed the Lifelong Kindergarten Approach and investigated differences in students' attitudes, their learning outcomes (measured by a pre-post-test, perceived learning, and task-based performance), and the fun they experienced during the workshops. We collected data from three primary schools in three distinct socioeconomic neighborhoods (i.e., high-, middle-, and low-income areas), involving, in total, 138 students. Findings indicate that the workshops had a positive effect on the students' attitude towards programming in the middle- and low-income schools only. The self-reported learning was similar in the three schools, but students from the low-income school significantly outperformed students from the high-income school in their task-based performance. Children from the middle-income school had the most fun, however, the experience of fun only significantly affected the low-income school students' perceived learning scores. We conclude that students from the middle- and low-income schools profited the most from the playful programming workshop and call on future research to investigate further underlying factors of perception, acceptance, and enjoyment of out-of-school programming activities in relation to participants' socioeconomic background when studying students' participation of programming in school.

[17] This chapter is based on the following publication: Tisza, G., Markopoulos, P., & King, H. (2022). Socioeconomic background influences children's attitudes and learning in creative programming workshop. Education and Information Technologies. https://doi.org/10.1007/s10639-022-11467-w

8.1 Introduction

In recent decades, out-of-school STEM learning, aimed at teaching Science, Technology, Engineering and Mathematics (STEM) subjects in a playful and engaging way has gained ground through learners' participation in maker spaces[18], Fab Labs[19], programming clubs and science museums [220, 255]. These venues typically provide children with a collaborative (work)space that enables exploring, learning, creating, and sharing. In case of maker spaces and Fab Labs, the emphasis is on making. Such settings offer a wide range of readily available tools from high-tech to no-tech. In programming clubs, the focus is on programming and robotics, whilst in science museums, a variety of scientific topics may be addressed, including making and programming. The overarching approach for out-of-school learning, is to develop learning environments that support learners' intrinsic motivation and trigger their curiosity.

This trend relates to a worldwide pursuit to increase children's interest in scientific topics, and especially in computer science, as computational thinking and programming are frequently seen as some of the main literacy skills of the 21st century [210]. According to Saez-Lopez et al. *"the ability to be a creator rather than just a consumer of technology is increasingly seen as an essential skill in order to participate fully in a digital society"* ([260], p. 131). This observation reflects the need to cultivate creativity from early age on. Non-curricular and out-of-school programming clubs can play a significant role in teaching children to program as programming is not yet an integral part of the primary school curricula. The UK, Estonia, Spain, and Finland are examples from Europe where programming is already a compulsory subject in primary education. In other countries, such as the Netherlands, primary schools can decide whether to teach programming to their students or not.

Despite this worldwide pursuit, we know little about what influences children's and adolescents' interest and willingness to participate in such activities. This study aimed to broaden our knowledge on possible underlying factors for children's and adolescents' participation in programming-related activities, and hence provide cues for a more successful design and implementation of such activities. We focused this examination around the possible effect of children's socioeconomic background, as earlier research shows that it can have an effect on children's academic achievement in general (e.g., [276, 323]), and in their STEM interest specifically [35, 201, 334]. However, we know very little about the relationship between children's socioeconomic background and their participation and learning to program. To this end, we designed and implemented a 2-hour-long, playful, programming workshop to introduce programming with micro:bits to primary school students. We investigated how the participating students' socioeconomic

[18] http://www.makerspaceforeducation.com/makerspace.html
[19] https://fabfoundation.org/getting-started/#fablabs-full

background and their attitude toward programming influenced the fun they experienced while learning to program, and ultimately, their learning outcomes.

8.2 Theoretical Background

8.2.1 Lifelong Kindergarten

The Lifelong Kindergarten is an often-used approach for teaching programming to children. It is described as being *"ideally suited to the needs of the 21st century, helping learners to develop the creative thinking skills that are critical to success and satisfaction in today's [digital] society"* ([237], p. 1). This approach emulates a traditional kindergarten environment where, during play, children design, create, experiment, and explore continuously. In this approach, learning takes place through a spiraling process that starts with imagining, and followed by creating, playing, sharing, reflecting, before returning to imagining, and so on. The widely used visual programming environment Scratch [174, 239], provides a framework which applies the key elements of the Lifelong Kindergarten for those learning to program. Since this approach has been argued to be universally suitable [237], and given that for the workshop we used a visual-programming interface, we hypothesized that all the students would find the workshop equally fun regardless of their gender or socioeconomic background (H1a and H1b).

8.2.2 Attitude, Fun and Learning

Besides an appropriate teaching method, having a positive attitude towards the subject can arguably play a key role in obtaining high (academic) achievements. Moreover, previous research with university students found that their attitude towards programming not only influenced their academic achievement, but also affected their career choices [50].

To examine the question of attitude and learning, Bakar et al. [19] investigated university students' attitude and academic performance and found a significant positive correlation between the two. Narmadha and Chamundeswari [199] investigating secondary school students' science-related attitudes and their academic achievement in science class and found a positive correlation between attitude towards learning science and students' academic achievement in science class. With respect to technology-related learning, Gunbatar and Karalar [104] found that programming with a visual programming environment - called mBlock - had a positive influence on middle school students' attitudes towards programming. Saez-Lopez et al. [260] found the same association with primary school children: after students learned to program with Scratch, their motivation and commitment about programming increased significantly. In Chapter 6 we introduced the FiL model after investigated primary school students' attitude towards programming and the learning outcomes of a programming workshop. According to the model, a more positive attitude towards programming is associated with higher levels of learning. We also reported that having fun while learning to code significantly and positively influenced students' attitude towards programming and their learning outcomes. Based on these latter findings, we in the current study hypothesized that the experienced fun while learning will have a positive effect on students' learning outcomes (H2).

8.2.3 Socioeconomic Status and Learning

Despite the vast evidence that has accumulated regarding the importance of students' positive attitude towards programming, we know little about what influences attitudes beyond learning to code with a visual programming interface. Arguably, socioeconomic background may play a role as, we suggest, students from lower socio-economic backgrounds will have limited access to programming opportunities. This affect may be more pronounced in countries where computer science or programming is not yet a compulsory subject in primary education.

The concept of *science capital* has been coined by Archer et al. in 2015 [13]. The concept provides an explanation as to how children's socioeconomic background could influence their science-related attitudes and interests. S*cience capital* encapsulates *"all science related knowledge, attitudes, experiences and social contacts that an individual may have"* ([99], p. 5). Grounded in Bourdieusian notions of capital and accrued privilege, the concept science capital acknowledges that particular advantages, such as socioeconomic status, will positively affect the science-related resources, contacts and experiences that a learner holds [13]. Other factors, such as ethnicity and gender, have been shown to shape one's science capital [69], and research mapping the intersectional affects – gender, ethnicity and social class – of learners' participation with science, technology and engineering is ongoing [193]. Since 2015, the concept of science capital has gained considerable traction in STEM education research, practice and policy [202] as findings have indicated that the higher one's science capital the more likely one is to engage with science and STEM related activities and to have a 'science identity', the latter indicating an increased likeliness to continue with science related studies after age 16 [14].

With respect to the relationship between socioeconomic status (SES) and academic achievement in general, the meta-analysis of Sirin [276] on primary and secondary school students concluded that there is an overall positive correlation. In another meta-analysis of early research on this topic with students from pre-school to high-school White [326] noted that the strength of the relationship between SES and academic achievement depended on how SES is defined and what is considered as the unit of analysis, with weak correlations found when the individual student is the unit of analysis and stronger correlations when the unit of analysis is the school. In this thesis we adopt the definition of Sirin, who considers socioeconomic status to be *"an individual's or a family's ranking on a hierarchy according to access to or control over some combination of valued commodities such as wealth, power, and social status"* ([276], p. 418).

In the specific domain of ICT literacy, a few studies have examined the effect of socioeconomic differences. Hatlevik and Christophersen [113] identified SES as a significant influencer on secondary school students' digital competence, with students from higher socioeconomic backgrounds having higher levels of competence than students from lower socioeconomic backgrounds. Senkbeil et al. [268] found in their study with lower secondary school students that students' ICT literacy was dependent on their family's social background and school achievement (mathematics and German grade). Another qualitative study involving secondary school students [323] observed that ICT-related knowledge and skills are dependent on socioeconomic status, with young, well-educated people of a higher SES having the highest knowledge and the most skills. In their recent meta-analysis, Scherer and Siddiq [264] concluded that ICT literacy is dependent on students' socioeconomic status, however, they emphasize that the relationship between SES and ICT literacy was weaker than those reported in other educational subjects such as mathematics or reading.

Regarding STEM education, the study of Niu [201] with college students found that low-SES students were disadvantaged in pursuing a STEM major, as they may not possess the skills and/or information (or indeed, science capital) required to make a well-informed decision on STEM enrollment. Niu also found that gender and racial gaps in STEM enrollment narrow for high SES students. The study of Yerdelen, Kahraman, and Tas [334] investigated low SES middle school students' STEM career interests and found that they had positive attitudes towards pursuing a STEM career, however, they did not compare these results with students from different socioeconomic background, hence it is difficult to assess how students' SES influenced their attitudes. The study of Blums et al. [35] aimed to examine early SES and later STEM achievement on a large, longitudinal data set. Their study results indicated that maternal education (as an often-used factor to determine SES) had a strong positive influence on children's cognitive abilities which are, on the long term, related to children's STEM achievement.

Based on the research reviewed, and building on the theory of science capital, we hypothesized that students from high income schools would perform better based on their higher exposure to STEM in general, and to computing and programming experiences specifically (H3) and would thus hold more positive attitudes towards programming (H4).

8.2.4 Study Aim and Hypotheses

In this multiple-case study we set out to investigate students' attitudes towards programming and their learning outcomes in relation to their socioeconomic status. More specifically, we aimed to examine whether students with different socioeconomic backgrounds profit evenly from a non-curricular creative programming workshop. Based on the above detailed earlier research we hypothesized that:

- **H1**: All the students will find the workshop equally fun regardless of their gender (**H1a**) of socioeconomic background (**H1b**).

- **H2**: The experienced fun while learning has a positive effect on students' learning outcomes.
- **H3**: Students from high income schools will perform better on the programming tasks, in other words, will have higher learning outcomes in comparison with students from lower income schools.
- **H4**: Students from high income schools hold more positive attitudes towards programming in comparison with students from lower income schools.

8.3 Method

8.3.1 Participants

The study was conducted in February 2020 in the Netherlands. Figures from 2019 indicate that the average yearly income per person in the Netherlands was 26 140 euro [20]. Accordingly, for the workshop and hence for participation in this study, we selected three socioeconomically distinct neighbourhoods with a low, an average, and a high yearly income[21], and recruited primary school classes from the selected neighbourhoods. In the rest of the chapter, we will refer to the schools as low-income, middle-income, and high-income schools. Detailed descriptive information about the schools can be found in Table 8.1 below. In total, three schools participated with six school classes and 138 students. The average age of the participants was 9.89 years (*SD* = 1.124). The gender distribution was relatively balanced, with 73 boys (52.9%), 64 girls (46.4%) and one who did not specify their gender (0.7%).

Table 8.1 Descriptive statistics of the three schools.

	High-income school	**Middle-income school**	**Low-income school**
Nr. of students	60	16	76
Age (*M*)	8.88 (*SD* = 0.640)	10.53 (*SD* = 0.516)	10.81 (*SD* = 0.696)
Gender distribution	31 boys (51.7%) 29 girls (48.3%)	10 boys (62.5%) 6 girls (37.5%)	39 boys (51.3%) 36 girls (47.4%) 1 not given (1.3%)
Average yearly income in the neighborhood of the school	€ 31.800	€ 26.600	€ 20.200

From the pre-workshop data collected examining prior experiences in programming, we found that most of the students participating in the activity were novices. A total of 22.5% of the students reported having no idea about programming, and 36.2% of the students reported knowing a bit. This is also reflected in the sample mean for the 5-step scale (*M* =

[20] https://allecijfers.nl/ranglijst/gemiddeld-inkomen-per-provincie-in-nederland/

[21] Source: https://allecijfers.nl; Average gross yearly income per habitant in the neighbourhood of the school (2019), used as an indication for socioeconomic background.

2.39, which translates to 'a bit'; SD = 1.114). In other words, almost 60% of the students were new to programming. When comparing the three schools we found that there was a significant difference between students' prior knowledge or understanding of programming (p = 0.005, F = 5.451, η^2 = 0.069). Namely, students from the middle-income school reported the highest values (M = 3.31, SD = 1.352), followed by the high-income school students (M = 2.47, SD = 1.033) and, lastly, the low-income school students (M = 2.31, SD = 1.097).

A total of 39.1% of the students reported never having participated in a programming activity, and 23.9% reported having participated in one programming activity only. These numbers reflect the current situation in the Netherlands in that programming is not a compulsory subject in primary education and schools can decide whether to teach it or not. Importantly, programming was not taught in any of the three schools. When comparing students' previous experience with programming activities between the schools we found a significant difference (p < 0.001, F = 11.598, η^2 = 0.137). Namely, students from the middle-income school had the highest average reported (M = 3.06, SD = 1.526), followed by the high-income school students (M = 2.46, SD = 1.222) and the low-income school students (M = 1.78, SD = 0.896).

To summarise, there is a clear difference in students' experience and self-reported initial knowledge between the three schools. Namely, students from the lowest socioeconomic neighborhood had on average the least previous experience with programming while students from the middle-income school had the most. However, most of the students across the total sample were novices to programming.

8.3.2 Ethical Considerations

Participation in the learning activity was compulsory as it took place during school hours in the classroom setting, but participation in the study (i.e., responding to the questionnaires) was voluntary. The data was collected anonymously, nevertheless, informed consent was obtained across the schools from both the students and their parents / caregivers. Neither the school nor the students received any incentives for participating in the study. The study was approved on 10 January 2020 by the Ethics Review Board of the Eindhoven University of Technology.

8.3.3 Procedure

In collaboration with SkillsDojo (a foundation that produces open-source STEM learning materials for children between 6 and 14) we designed a single-occasion, two-hour-long creative programming workshop for primary school students. The workshop aimed to introduce programming with BBC micro:bits (www.microbit.org) for students. We prepared three tasks of increasing complexity and difficulty levels. In the first, introductory task students wrote a program to display their names. In the second, they created a stone-paper-scissors game. In the third task, they either created a micropet that reacted to kinetic stimuli, or they could decide to choose themselves what to code. Examples from the workshop for the programming tasks are shown on Figure 8.1.

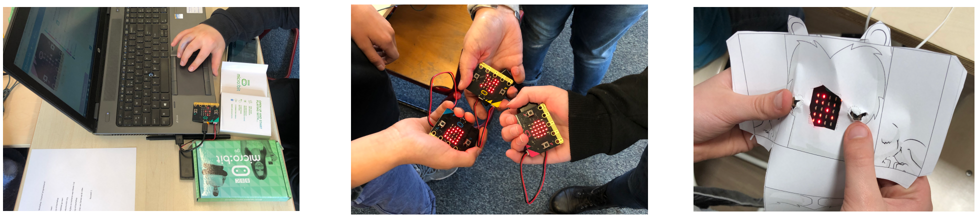


Figure 8.1 The three programming tasks (from left to right): Program your name; Program stone-paper-scissors game; Program a micropet.

The workshops were held in a classroom but were not part of the school's formal curriculum. The activity was designed with and based on interactive video guides that follow the Lifelong Kindergarten approach to introduce programming to students. Accordingly, students' imagination and curiosity were triggered through the use of micro:bits, and during the whole workshop students were encouraged to play, share, and reflect on their codes, games and artefacts, for example, by helping each other with debugging of the code. Once code had been developed, students could play together with the game they had made, and if they wished, they could refine the code further. In line with the Lifelong Kindergarten approach, the workshop aim was not only to learn to program with micro:bits, but to do so in a creative and deeply engaging way.

During the workshop, students were equipped with their own laptops/Chromebooks which further supported the personal authorship of the activity. Nevertheless, students were allowed and encouraged to work with each other, thereby fostering communication and collaboration, prompting, sharing, and reflecting [320]. In addition, students were also permitted to move freely around the room, ask questions as they liked - of each other and/or the facilitators - and interact with each other. This aimed to further foster a sense of agency and to disrupt traditional classroom structures pivoted on getting answers right.

Three researchers and the teacher were present during the workshop. At the beginning of the workshop and after the introduction the researchers handed out the pre-workshop questionnaire to the students. While students were busy filling the questionnaire, the researchers prepared the Chromebooks/laptops and distributed the micro:bits and a printed step-by-step guide. Once the questionnaires were collected, students were asked to explore the micro:bits, and then assemble and plug them in the Chromebooks/laptops. Thereafter, students were asked to open the website of the videos (www.skillsdojo.nl/workshop (Dutch) or www.kidzcourse.com/workshop (English)) and the website from which they could programme the micro:bit (i.e., programming interface; www.makecode.microbit.org). The researchers helped students with these steps and encouraged them to start watching the videos and follow the instructions. When the time was over, students were asked to tidy up their tables and the post-workshop questionnaire was handed to them.

8.3.4 Materials

As noted earlier, students followed a how-to video guide to complete the programming tasks described as a set of SkillsDojo missions. This video guide was created by the SkillsDojo Foundation implementing the Lifelong Kindergarten approach where participants learn how technology works through a digital or physical project, building on 21st century skills e.g., working together, problem solving, critical thinking. All SkillsDojo missions have a 'low floor' making it easy for everyone to begin and to complete the mission, a 'high ceiling' so that in each mission there is plenty of room to grow and students are constantly being challenged, and 'wide walls' so that anyone can make any mission relevant to themselves.

The videos build on the dual programming principle, namely, the videos use two channels (audio and picture) and this supports double-barreled learning and, in line with the cognitive load theory prevent overloading working memory by following the segmentation principle (i.e., they are built of 'chunks') and signaling (highlighting the important parts). Students can set their own pace and follow a declining guidance strategy (phasing out guidance). Finally, the videos use the redundancy principle i.e., combination of audio and picture instead of audio and word and the worked-example effect, the learning effect observed when working examples are used as part of the instruction.

8.3.5 Measures

For the assessment of students’ socioeconomic background, and to stratify our sample, we used the average gross yearly income per habitant in the neighbourhood of the school. We decided to use this measure for multiple reasons. First of all, in educational contexts this method has been applied successfully before (e.g., [323]). Second, obtaining precise data from the parents about their SES would have introduced unnecessary ethical concerns, raising questions of anonymity and issues with willingness for participation, ultimately resulting in the introduction of sampling bias. Third, previous findings consistently indicate a positive association between students’ educational outcomes and their schools’ neighbourhood SES (for a systemic review, see Nieuwenhuis and Hooimeijer [200]; example studies in the Dutch context are Kuyvenhoven and Boterman [157] and Sykes and Musterd [288]). Fourth, previous research [326] found that the relationship between SES and academic achievement is stronger when the unit of the analysis is the school in comparison with the individual. Therefore, we concluded that using the average yearly income in the neighbourhood of the school is a reliable proxy for students' socioeconomic background and it is a suitable method for the assessment of differences in learning outcomes.

In the pre-workshop questionnaire, we measured students' self-reported knowledge on programming by two questions: 'Do you have any idea about programming?' ((1) not at all --- (5) I'm a pro)) and 'How many programming activities have you participated before? ((1) none --- (5)-six or more). Additionally, we measured students' attitude towards programming across six bi-polar items [206, 210] both at the beginning and the end of the workshop (see Figure 8.2). By collecting responses on these items both before and after the workshop we aimed to understand whether the workshop had a positive effect on students' attitude about programming. For the attitude items we used the smiley-face scale designed and validated by Hall et al. [108]. In addition to these six specific attitude items, we used a more general item ('Programming is my thing'), which we adopted from earlier research [300] where it has been shown to be a reliable measure for students' general programming-related attitude, and which was evaluated on a 5-point scale. The internal consistency of the seven attitude dimensions appeared to be adequate both before and after the workshop ($\alpha_{pre\text{-}workshop}$ = 0.781, $\alpha_{post\text{-}workshop}$ = 0.833).

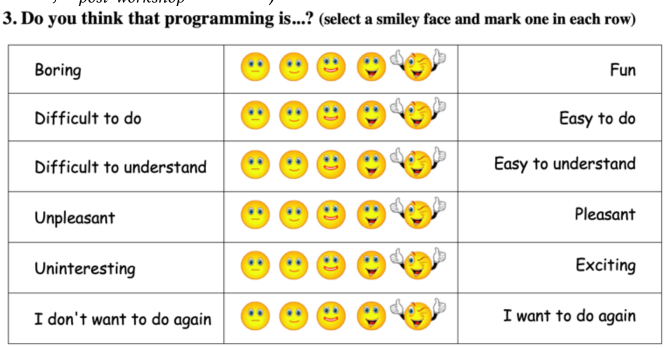


Figure 8.2 Attitude questions of the pre- and post-workshop questionnaires.

Since earlier research mostly either pertained to reported or measured learning, - and those who used the combination of these found that the two measures do not necessarily align [127, 306] - in order to gain a comprehensive picture, we decided to use for the assessment of learning three measures that reflect three levels of learning according to Bloom's taxonomy [34]. Accordingly, we recorded a knowledge assessment test both before and after the workshop and calculated the *measured learning* scores by subtracting the pre-workshop scores from the post-workshop scores (*knowledge* level of Bloom's taxonomy; possible range was -5 to +5). Additionally, at the end of the workshop students self-reported on their *perceived* level of *learning* ('Have you learned something new today about programming?', (1) not at all --- (5) a whole lot; *evaluation* level of Bloom's taxonomy). As a third measure, we calculated students' *task-based performance* on the second task (*application* level of Bloom's taxonomy). We have chosen the second task as it was expected to be the most reliable part of the workshop for task-based performance, given that the first task had an introductory nature and that many students decided to develop their own code after the second task. For rating the task-based performance, due to resource limitations, eleven randomly selected students' screens could be captured in each of the six classes, from which 54 could be used to rate students' task-based performance. Twelve screen captures were damaged or lost during the data recording or saving process due to overheating of laptops and/or freezing of the system and/or freezing of the screen-capture program. Since the second task involved five distinct steps to complete, students were rated on a scale of 0-5 by two raters on their performance. For each correctly conducted step, 1 point could be earned. The inter-rater agreement was 100%.

For assessing students' perceived level of fun during the workshop, we used the FunQ, which was evaluated on a 5-point Likert-type scale. The internal consistency of the FunQ appears to be adequate on our sample (α = 0.833).

8.3.6 Data analysis

For the analysis of students' pre- and post-workshop questionnaire data we applied quantitative data analysis techniques, including one-way ANOVA, multivariate general linear models, and repeated measures general linear models. For the data analysis we used the SPSS Statistics version 27 software.

8.4 Results

8.4.1 Fun (H1)

To assess the level of fun students experienced, we recorded FunQ after the workshop. For testing H1, we applied one-way ANOVA to compare the FunQ scores across schools and genders. The overall minimum FunQ score was 39 and the maximum was 90 in our sample (M = 70.48, SD = 10.205) from the possible range of 18-90. According to Leven's test, equal variances across the three schools were assumed (*Leven's test* (2, 116) = 2.706, p = 0.071). Our results indicate that there is a significant difference in the level of fun experienced between schools (p = 0.001, F(2) = 7.493, η^2 = 0.114). The average FunQ score for the high-income school students was 70.33 (SD = 8.832), for the middle-income school students was 79.20 (SD = 6.05), and for the low-income school students was 68.34 (SD = 10.971), meaning that students from the middle-income school - where we have seen the most positive attitudes about programming - experienced the most fun during the workshop. We found no gender difference in the level of fun experienced (p = 0.436, F(116) = 1.165, *Cohen's d* = 0.144, M_{boys} = 71.19 (SD = 10.619), M_{girls} = 69.71 (SD = 9.839)).

In sum, students experienced the workshops as fun rating them in the upper third of the range, however, the level of fun they experienced varied across the schools significantly (students from the middle-income school experienced the workshops as most fun; H1b), while the experienced fun was not gender dependent (H1a). Therefore, H1 is only partially supported (i.e., H1a is supported and H1b is refuted) as we expected that all the students will find the workshop equally fun regardless of their socioeconomic background or gender.

8.4.2 Learning (H2 and H3)

As discussed above, we used three measures to address students' learning that indicate different level of learning according to Bloom's taxonomy [34]. For testing H2 we examined the effect of fun on the three levels of learning, and for testing H3 we addressed how the school, as a proxy for students' SES, influenced students' learning outcomes. The average measured learning, perceived learning, and task-based performance scores are displayed in Figure 8.3.

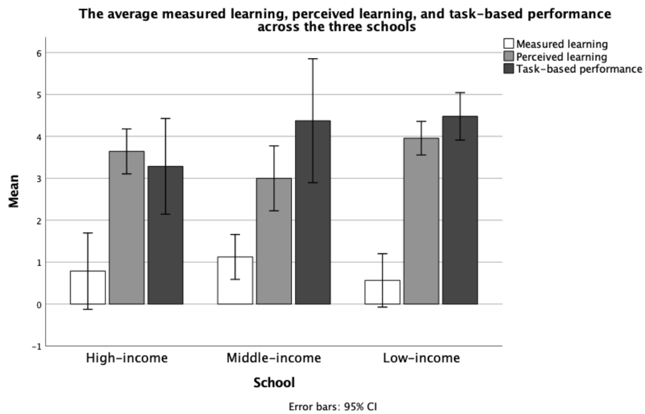


Figure 8.3 Average measured learning, perceived learning, and task-based performance across the three schools. Possible range for the measures are -5 to +5; 0 to 5; and 0 to 5 accordingly.

8.4.2.1 Measured learning

The sample mean for the measured learning is 0.7328 (*SD* = 1.41) and we did not find a significant gender difference (p = 0.175, F (128) = 0.762, *Cohen's d* = -0.240). The average measured learning score in the high-income school is 0.833 (*SD* = 1.68), it is 1.067 (*SD* = 0.88) in the middle-income school and 0.565 (*SD* = 1.25) in the low-income school. We did not encounter a ceiling effect. One-way ANOVA test indicates that these differences are not statistically significant (p = 0.373, F(2) = 0.995, $\eta^2 = 0.015$; *Leven's test* (2, 128) = 3.160, p = 0.056), however, we see that students from the lowest socioeconomic neighborhood performed the worst.

To assess whether having fun while learning affected students' measured learning, we conducted linear regression analyses. We found that fun is not a significant predictor of students' measured learning (high-income school: p = 0.855, β_{std} = 0.029; middle-income school: p = 0.289, β_{std} = 0.305; low-income school: p = 0.165, β_{std} = 0.198). When investigating the differences by applying univariate general linear model with fixed factors 'school' and 'fun', we find that neither fun (p = 0.574, F (37) = 0.940, partial η^2 = 0.471) nor the school (p = 0.399, F (2) = 0.941, partial η^2 = 0.046) or their interaction effect (p = 0.349, F (27) = 0.773, partial η^2 = 0.349) is significant.

In sum, we found that students from the middle-income school outperformed students from the other two schools in the learning assessment test (see Figure 8.3), however, this difference was not significant. Additionally, we did not find a significant link between having fun while learning and the learning outcomes in any of the schools.

8.4.2.2 Perceived learning

We recorded students' perceived learning at the end of the workshop. To test the differences between the schools, we applied one-way ANOVA According to the Leven's test, equal variances across the three schools were assumed (*Leven's test* (2, 135) = 1.823, p = 0.166). The average perceived learning in the high-income school was 3.95 (SD = 0.934), it was 3.50 (SD = 1.155) in the middle-income school, and 3.84 (SD = 1.153) in the low-income school. In other words, students from the middle-income school report on having learnt the least. Nevertheless, these differences are not statistically significant (p = 0.335, F (2) = 1.101, η^2 = 0.016). We add that we did not encounter a ceiling effect.

To assess how having fun while learning affected students' perceived learning, we firstly conducted a regression analysis for each school. In the high-income school, the perceived fun while learning is not a significant predictor for perceived learning (p = 0.163, β_{std} = 0.216). It also accounts for less than 5% of the variance in the learning scores (R^2 = 0.047). In the middle-income school we see a similar tendency. Fun is not a significant predictor for students' perceived level of learning (p = 0.099, β_{std} = -0.442). In the low-income school, however, having fun while learning accounts for approx. 50% of the learning scores (R^2 = 0.519), and hence, fun is a significant predictor of the perceived learning (p < 0.001, β_{std} = 0.721).

When investigating the differences by applying univariate general linear model with fixed factors 'school' and 'fun', we see that fun has a significant effect on students' perceived level of learning (p = 0.017, F (38) = 1.968, *partial* η^2 = 0.640), however, neither the school (p = 0.110, F (2) = 2.328, *partial* η^2 = 0.100), nor the interaction effect between fun and the school (p = 0.265, F (27) = 1.230, *partial* η^2 = 0.459) is significant.

In sum, we found no significant difference among the schools in students' perceived level of learning. However, we found that fun affected differently students' perceived learning depending on their socioeconomic background as indicated by the schools they attend. Accordingly, for low-income school students, having fun while learning had a strong influence on their perceived learning, while this is not true for students from the two other socioeconomically better situated schools.

8.4.2.3 Task-based performance

The mean task-based performance of them is 4.02 (SD = 1.754) and the scores vary between 0 and 5. The average task-based performance in the high-income school was 3.07 (SD = 2.086), it was 4.38 (SD = 1.768) in the middle-income school, and 4.37 (SD = 1.450) in the low-income school. One-way ANOVA indicates a significant difference in the task-based performance between schools ($p = 0.05$, $F(2) = 3.169$, $\eta^2 = 0.112$; *Leven's test* (2, 50) = 3.114, $p = 0.053$). Children from the high-income school performed significantly worse than students from the low-income school ($p = 0.019$, $F(43) = 6.670$, *Cohen's d* = -0.772).

To assess how fun influenced students' task-based performance, we conducted regression analysis. We found that fun is not a significant predictor for students' task-based performance in either of the schools (high-income school: $p = 0.185$, $\beta_{std} = 0.185$; middle-income school: $p = 0.902$, $\beta_{std} = 0.058$; low-income school: $p = 0.060$, $\beta_{std} = 0.417$).

When investigating the differences by applying univariate general linear model with fixed factors 'school' and 'fun', we see that neither fun ($p = 0.511$, $F(26) = 1.074$, partial η^2 = 0.823), nor the school ($p = 0.245$, $F(2) = 1.795$, partial $\eta^2 = 0.374$), or their interaction effect ($p = 0.542$, $F(5) = 0.889$, partial $\eta^2 = 0.426$) has a significant influence on students' task-based performance.

To summarize, we found that students from the high-income school performed significantly worse on the task-based performance than the other two schools. However, we did not find a significant relationship between students' perceived fun while learning and their task-based performance in any of the schools.

To conclude on learning, we found that students from the middle-income school thought that they have learnt the least (i.e., perceived learning) compared with students from the low- and high-income schools. Regarding the measured learning scores, students from the low-income school gained less knowledge than students from the high-income school. However, students from the low-income school significantly outperformed students from the high-income school on the task-based performance. Interestingly, students from the middle-income school thought that they have learnt the least (i.e., perceived learning), yet they outperformed students from the other schools on both the measured learning and the tasked-based performance scores. Therefore, H3, in which we expected that students from high-income schools would perform better based on their higher exposure to computing is supported in case of the perceived- and the measured learning but is rejected in case of the task-based performance. The average measured learning, perceived learning, and task-based performance scores are displayed in Figure 8.3. Regarding H2, in which we expected that the experienced fun while learning would have a positive effect on students' learning outcomes, our results are partially supported as we found that fun had a positive effect only in case of the low-income school and students' perceived learning scores.

8.4.3 Attitude toward programming (H4)

To test H4 and to investigate the development of science-related attitudes, and specifically, students' attitude toward programming, we asked them before and after the workshop across six bi-polar scales and a 5-step Likert scale, and compared the results along the schools, which we used as a proxy for students' SES. A summary of the scores on the seven attitude dimensions, before and after the workshop, according to the schools is displayed in Table 8.2. All statistical results regarding the effect of school and gender are displayed in Table H1-H4 in Appendix H.

Table 8.2 Average scores on the attitude dimensions before and after the workshop across the three schools. * indicates significant change ($p < 0.05$).

Attitude dimension	High income school		Middle-income school		Low-income school	
	Pre	Post	Pre	Post	Pre	Post
Boring – fun	M = 4.56 SD = 0.676	M = 4.58 SD = 0.889	M = 4.88 SD = 0.500	M = 4.94 SD = 0.250	M = 4.33 SD = 0.822	M = 4.33 SD = 1.048
Difficult to do - Easy to do	M = 3.05 SD = 0.860	M = 3.13 SD = 1.241	M = 3.56 SD = 1.094	M = 4.19* SD = 1.047	M = 3.09 SD = 1.018	M = 4.00* SD = 0.986
Difficult to understand - Easy to understand	M = 3.08 SD = 0.952	M = 3.40 SD = 1.224	M = 4.06 SD = 0.854	M = 4.75* SD = 0.577	M = 3.23 SD = 1.085	M = 3.90* SD = 1.024
Unpleasant – Pleasant	M = 4.25 SD = 0.863	M = 4.30 SD = 1.046	M = 4.81 SD = 0.544	M = 5.00 SD = 0.000	M = 4.14 SD = 0.857	M = 4.19 SD = 0.928
Uninteresting – Exciting	M = 4.05 SD = 0.782	M = 4.02 SD = 1.084	M = 4.88 SD = 0.342	M = 4.88 SD = 0.342	M = 4.13 SD = 0.984	M = 4.01 SD = 1.110
I don't want to do - I want to do	M = 4.19 SD = 1.051	M = 4.41 SD = 1.131	M = 4.81 SD = 0.403	M = 5.00 SD = 0.000	M = 4.27 SD = 0.994	M = 4.39 SD = 1.040
I think that programming is my thing	M = 4.02 SD = 0.881	M = 4.10 SD = 1.115	M = 4.56 SD = 0.512	M = 4.80 SD = 0.414	M = 3.65 SD = 0.905	M = 3.84 SD = 1.163

8.4.3.1 Pre-workshop attitudes

We investigated whether there is a difference across the three schools controlling for students' gender in the pre-workshop attitude scores by applying multivariate general linear model with fixed factors 'school' and 'gender'. We see that students from the middle-income school scored on average higher on all items than students from the other schools. The effect of school is thus accordingly significant in all but the 'difficult to do/easy to do' (p = 0.267, F (2) = 1.336, *partial* η^2 = 0.020) and 'I don't want to do/I want to do' (p = 0.111, F (2) = 2.232, *partial* η^2 = 0.033) attitude scores. The effect of gender on the attitude questions was not significant (p = 0.944, F (7) = 0.283, *partial* η^2 = 0.013).

8.4.3.2 Post-workshop attitudes

Here again we applied the multivariate general linear model with fixed factors 'school' and 'gender', to test the difference between the schools, controlling for students' gender. We see that, in general, students from the middle-income school scored on average higher in all attitude items than students from the other two schools. The effect of 'school' is thus significant in all but one ('I don't want to do again/I want to do again', $p = 0.094$; $F(2) = 2.411$, *partial* $\eta^2 = 0.036$) attitude score. The effect of gender was not significant ($p = 0.604$, $F(7) = 0.782$, *partial* $\eta^2 = 0.043$).

8.4.3.3 Attitude change

We applied the repeated measures general linear model to test whether students' attitude had changed differently across the schools, and to see whether there is a gender effect.

For the bi-polar scale *Do you think that programming is boring/fun* we found no significant change in the pre- and post-workshop scores (within subject effect; $p = 0.693$, $F(1) = 0.156$, partial $\eta^2 = 0.001$). The effect of school ($p = 0.003$, $F(2) = 6.033$; *partial* $\eta^2 = 0.081$) was however significant, but the effect of gender ($p = 0.867$; $F(1) = 0.135$; *partial* $\eta^2 < 0.001$) was not. In other words, students' attitude regarding whether programming is boring or fun was not significantly affected by the workshop, however, students' attitude differed between the three schools.

For the bi-polar scale *Do you think that programming is difficult to do/easy to do* we found a significant change in the pre- and post-workshop scores (within subject effect; $p = 0.000$, $F(1) = 20.241$, *partial* $\eta^2 = 0.131$). The effect of school ($p = 0.001$, $F(2) = 7.125$, *partial* $\eta^2 = 0.096$) is also significant, but the effect of gender ($p = 0.298$, $F(1) = 1.094$, *partial* $\eta^2 = 0.008$) is not. In other words, students' attitudes on whether programming is difficult or easy to do was significantly and positively affected by the workshop and was different among the three schools.

For the bi-polar scale *Do you think that programming is difficult to understand/easy to understand* we found a significant change in the pre- and post-workshop scores (within subject effect; $p < 0.001$, $F(1) = 20.679$, *partial* $\eta^2 = 0.134$). The effect of school ($p < 0.001$, $F(2) = 10.489$, *partial* $\eta^2 = 0.135$) is also significant, but the effect of gender ($p = 0.186$, $F(1) = 1.767$, *partial* $\eta^2 = 0.013$) is not. In other words, students' attitude whether programming is difficult or easy to understand was significantly and positively affected by the workshop. Moreover, students' attitude was different among the three schools.

For the bi-polar scale *Do you think that programming is unpleasant/pleasant* we found no significant change in the pre- and post-workshop scores (within subject effect; $p = 0.282$, $F(1) = 1.168$, *partial* $\eta^2 = 0.009$). The effect of school ($p = 0.001$, $F(2) = 7.033$, *partial* $\eta^2 = 0.097$) is however significant, but the effect of gender ($p = 0.879$, $F(1) = 0.023$, *partial* $\eta^2 < 0.001$) is not. In other words, students' attitude as to whether programming is unpleasant or pleasant was not significantly affected by the workshop, however, students' attitude was different among the three schools.

For the bi-polar scale *Do you think that programming is uninteresting/interesting* we found no significant change in the pre- and post-workshop scores (within subject effect; $p = 0.068$, $F(1) = 3.375$, *partial* $\eta^2 = 0.025$). The effect of school ($p = 0.003$, $F(2) = 5.935$, *partial* $\eta^2 = 0.083$) is however significant, but the effect of gender ($p = 0.506$, $F(1) = 0.445$, *partial*

$\eta^2 = 0.003$) is not. In other words, students' attitude whether programming is uninteresting or interesting was not significantly affected by the workshop, however, students' attitude was different among the three schools.

For the bi-polar scale *Programming is something I don't want to do/I want to do* we found no significant change in the pre- and post-workshop scores (within subject effect; $p = 0.135$, $F(1) = 2.259$, *partial* $\eta^2 = 0.017$). The effect of school ($p = 0.039$, $F(2) = 3.329$, *partial* $\eta^2 = 0.049$) is however significant, but the effect of gender ($p = 0.735$, $F(1) = 0.115$, *partial* $\eta^2 = 0.001$) is not. In other words, students' attitude whether programming is something they want to do or not was not significantly affected by the workshop, however, students' attitude was different among the three schools.

For the 5-step Likert scale *Programming is my thing* we found no significant change in the pre- and post-workshop scores (within subject effect; $p = 0.144$, $F(1) = 2.158$, *partial* $\eta^2 = 0.017$). The effect of school ($p = 0.001$, $F(2) = 7.459$, *partial* $\eta^2 = 0.105$) is however significant, but the effect of gender ($p = 0.452$, $F(1) = 0.569$, *partial* $\eta^2 = 0.004$) is not. In other words, students' attitudes regarding whether programming is their 'thing', or not, was not significantly affected by the workshop, however, students' attitudes differed between the three schools. A summary of the scores on the seven attitude dimensions, before and after the workshop, according to the schools is displayed in Table 8.2.

For a general impression on the attitude change across the three schools, we calculated the average aggregate score (i.e., compound score) on all seven attitude dimensions (see Figure 8.4). We found that students' general attitude about programming has increased significantly in case of the middle- and low-income school, but not in the high-income school.

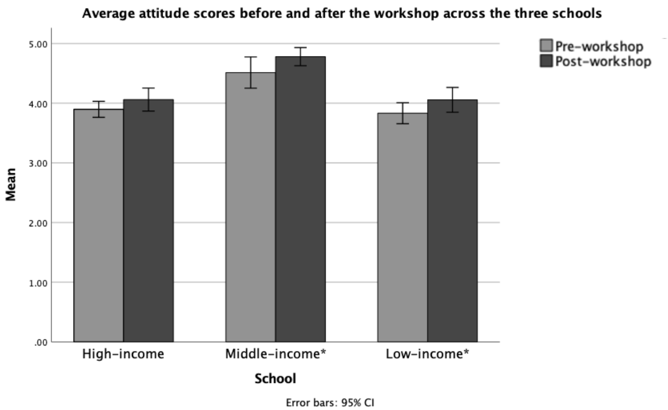


Figure 8.4 Average attitude score before and after the workshop across the three schools. * indicates a significant change ($p < 0.05$).

In sum, we see a tendency that students' attitude scores were positively influenced by the participation in the workshop, and this positive influence was significant in case of the 'difficult to do/easy to do' and 'difficult to understand/easy to understand' items. Considering the aggregated average attitude scores, we conclude that the workshop had a significant positive effect on students' attitude about programming at the middle- (p = 0.008, t = 3.068, *Cohen's d* = 0.792) and low-income (p = 0.021, t = 2.384, *Cohen's d* = 0.327) school, but not at the high-income school (p = 0.138, t = 1.506, *Cohen's d* = 0.201). Further, the effect of school was overall significant, in other words, students' attitude score was dependent on the school they attended. In case of the middle-income school the effect size indicates a strong relationship with students' attitude about programming, while in case of the low- and high-income schools the effect is considered to be small. However, we did not find a gender difference in any of the attitude scores and score changes.

8.5 Discussion

In this study we aimed to investigate whether students with different socioeconomic backgrounds profit evenly from a non-curricular creative programming workshop that follows the Lifelong Kindergarten approach. We evaluated students' attitude about programming and their learning outcomes, while controlling for gender differences. Our research results indicate that both students' attitudes about programming and their learning outcomes were affected by their socioeconomic status, as this is indicated by the average yearly income of the neighborhood of the schools. This influence though was not as expected based on previous literature.

Programming and creative thinking are frequently seen as the most important skills of the 21st century [210, 237, 260]. As computer science or programming is still not a mandatory subject in primary education around the world, non-curricular learning activities play a crucial role in introducing programming to children in a playful and engaging way. Previous research indicated that adopting the Lifelong Kindergarten approach [237] and using a visual programming interface [104, 260] as the way of introducing programming provides learners with the possibility to deeply engage with the topic, to improve creative thinking skills, and to develop a positive attitude about programming. Our research results partially support these previous findings as we found that participation in the visual-programming interface- based creative programming workshop significantly increased the low- and middle-income school students' general attitude about the topic (i.e., compound score), but it did not have a significant effect on the high-income school students. Since earlier research did not provide a quantitative comparison of attitudes of children from different socioeconomic backgrounds, we call for a replication of this study in different setups for a wider understanding of factors affecting children's and adolescents' attitudes about programming. Nevertheless, our findings suggest that changing students' perception about the difficulty of programming is a key element to attract them to similar activities in the future. This is especially true for students from a low socioeconomic background, as they had on average the least previous experience with programming before the workshop.

We also saw that students from the middle-income school reported on average higher attitude scores (both before and after the workshop) than students from the other two schools. This finding complements that of Yerdelen, Kahraman, and Tas [334], whose investigation of low SES students found that they had a generally positive attitude toward pursuing a STEM career. However, our findings question assumptions based on theory of science capital [13] which would suggest that students from high income families with greater access to science and technology related resources and contacts will hold more positive attitudes towards science (and by extension science and technology-related disciplines) in general compared with students from lower income families. As a possible explanation for these findings, we speculate that the middle-income school students have well-educated parents that grant values and interest in new technologies alongside a reasonable income to enable access to a variety of related activities, but we propose that such experiences are still sufficiently unusual to prompt novelty and more positive attitudes. On the other hand, students from the high-income school are more likely to have had high exposure to new technologies at home, and hence the workshop could have had less novelty and offered less challenge prompting minimal attitudinal affects. This speculation is supported by our finding that their attitude about programming did not change significantly in any of the seven investigated dimensions. Another possible explanation for these results is related to our finding that students from the middle-income school had the most prior knowledge in comparison with the students from the high- and low-income school, which could have had an influence on the attitude of these students, however, it does not provide an explanation for the post-workshop attitude scores.

Children from the low-income school reported the lowest attitude scores before the workshop. However, the workshop was as effective as for children in the middle-income school in terms of attitude change as in both cases students' attitude about programming has changed significantly in two out of the seven investigated dimensions. Based on these findings we conclude that students' attitude about programing, and the effect of the playful programming workshop is dependent on students' socioeconomic background, with middle- and low-income school students profiting the most, regardless of their gender. These results align with that of earlier research on the positive association between SES and STEM interest [35, 201, 334], however, our study goes a step further by focusing on the specific topic of learning to program, and provides new insights relating to primary aged students' participation in programming.

To address the effect of the workshop on students' learning, we investigated three levels of learning according to Bloom's taxonomy [34]. Regarding students' performance on the knowledge assessment test (i.e., measured learning) we found no statistically significant difference between the students from the three schools. Nevertheless, we see that students from the lowest socioeconomic neighborhood had the lowest learning gain scores. These findings align well with previous studies with secondary and college students that indicated a positive association between socioeconomic status and academic achievement

[264, 276, 323]. Sirin [276] report on an overall positive relationship between socioeconomic status and academic achievement in their meta-analytic review. Warschauer et al. [323] investigated access and use of new technologies in a group of low- and high-SES high schools and based on interviews with teachers and students they concluded that students from low-SES schools were more often assessed as being below grade-level in English and mathematics than students from the high-SES schools. The meta-analysis of Scherer and Siddiq [264] suggested a difference in ICT literacy between children from different socioeconomic background. Our study extends these findings, by investigating programming-related learning outcomes in a programming workshop in relation to the participating students' socioeconomic background and with primary school students.

Regarding the perceived learning scores, we found no statistically significant difference in the scores between the schools. However, we see that students from the middle-income school reported the lowest scores for their perceived level of learning (meanwhile they scored the highest on the knowledge assessment test). This result is only based on comparing three schools and would need further replication before we can generalize this conclusion. Regarding the task-based performance, we found that students from the high-income school performed significantly worse than students from the low-income school. This finding, we argue, might be related to students' engagement with the activity, and accordingly, we suggest that students from the high-income school (with some prior programming experience) found the workshop less engaging than students from the middle- and low-income schools, which is reflected in their task-based performance. Another possible confounding element is the academic level or general intelligence of the students, which we did not investigate in this study. Accordingly, future studies addressing this question could shed light on further factors that influence students' programming-related learning outcomes.

In sum, we found that the playful programming workshop was partially successful in terms of learning, as we found no statistically significant difference in students' measured and perceived learning between the three schools, but we found that the task-based performance of students from the high-income school was significantly lower than that of the low-income school students. This is a novel finding as previous studies did not directly investigate the relationship between socioeconomic background and programming-related learning outcomes.

We also aimed to understand better whether the enjoyment of the workshop had an influence on students' learning. We found that students from the middle-income school experienced the workshops as most fun, while the experienced fun was not gender dependent. This is a novel perspective on playful learning, as previous literature has not examined the fun experienced in non-curricular programming activities in relation to participants' socioeconomic background. While this result is only based on comparing three schools and would need further replication before we can generalize this conclusion, we argue that this perspective is key to better understanding what sort of activities are appreciated in different socioeconomic contexts.

Regarding the effect of fun on learning, our study introduced in Chapter 6 with primary school students suggested a positive association between having fun while learning to code and students' perceived learning. Furthermore, the study discussed in Chapter 5 with secondary school students in the field of digital game-based learning found the same positive association between students' perceived learning and the fun they experienced while learning, however, in that study we did not find the same effect in case of measured learning. In this study we found no significant association in either of the schools between fun and the measured learning or the task-based performance, which, in general, aligns with the research with our earlier findings discussed in Chapter 5, but it extends those by providing a more nuanced picture by investigating students from different socioeconomic background. Similarly, our finding that having fun while learning to code had a significant and positive effect on students' perceived learning in the low-income school, but not in the other two extends the FiL model with a more nuanced picture that takes SES into account. Considering the role played by SES is important as it demonstrates that playful programming workshops can contribute to a more positive perception of programming among low-income school students. Since we know very little about the aforementioned relationship, we call on future research to explore how exactly fun affects students with different socioeconomic background to learn to program.

As a final discussion point, we address the assessment of students' socioeconomic status. We must state that in case of young children, the assessment of their socioeconomic status is difficult, as children are unlikely to know their own relative status or understand differences between individuals. knowledge about it. In addition, involving the parents to clarify the situation is not always possible, and thus the response rate could be low, and further it may undermine the anonymity of the data collection. A possible way to overcome this issues is using the average yearly income of the neighborhood of the school [323] – the protocol we have followed in this study. However, this approach assumes that most children go to the school in their neighborhood, and that people in the same neighborhood have an approximately equal yearly income, and hence, approximately equal socioeconomic status. While the former is in general true in Holland (i.e., the majority of children attend the closest school in their neighborhood), the latter is only an assumption, which is nonetheless frequently used in the field of sociology. To further strengthen the findings of this study, future research could adopt different ways for the assessment of students' socioeconomic background, for example, surveying the parents, or using other proxies, such as the Family Affluence Scale [308].

8.6 Limitations and Future Work

While our study results are partially supported by previous research, our findings are still limited to the study location. Accordingly, future research should investigate students from a broader spectrum regarding their socioeconomic status, and eventually, in other, less wealthy countries than the Netherlands.

Additionally, our study involved a 2-hours long intervention, due to which we could only expect a limited effect on students' attitude, and we did not investigate the permanence of this effect. Therefore, we call on further studies to examine students' STEM and programming-related attitudes over time.

Regarding the assessment of learning, despite we have applied three different measures to address three levels of learning according to Bloom's taxonomy [34] (i.e., perceived learning, measured learning, and task-based performance), we need to acknowledge that the measurement of learning is complicated and capturing *actual* learning is challenging. A possible way to improve learning assessment in future research could be to use more sophisticated measures for capturing the measured learning than the difference of the post-test and pre-test scores, as this latter has the limitation of not taking into account students' relative learning in reflection of their initial knowledge, and it can also be prone to floor effect.

Furthermore, we selected three schools based on the socioeconomic neighborhood they are located in. However, this choice has some limitations as there could have been other factors that could have differentiated the schools, for example the school pedagogy that we were not aware of, the academic level of the participating students or their general intelligence. Accordingly, the structure of the workshop and the applied Lifelong Kindergarten pedagogy could have been variably suitable for the different schools, perhaps because of the school itself, and not because of the socioeconomic status. Whilst none of the schools expressed following a specific pedagogy, we acknowledge that the freedom provided in the Netherlands for schools to organize their curriculum and way of teaching may have created pedagogical differences. Therefore, to completely exclude these limitations, a future study should compare schools not only based on their socioeconomic neighborhood, but their applied pedagogy as well, possibly investigating schools with a specific pedagogy like Montessori or Dalton Plan schools.

Another possible factor that could have affected students' performance is the time of the day in which students were asked to code. One could expect that performance and learning may fluctuate at different times of the day, especially as students may get tired after several hours of schooling. This fluctuation, however, is equally affected the high-income and low-income schools as in both schools the workshop was given during both the morning and the afternoon hours. In the middle-income school, we only gave the workshop during the morning hours.

8.7 Conclusion

We designed and implemented a series of single-occasion playful programming workshops that followed the Playful Kindergarten Approach [237] to introduce programming in a playful and engaging way to primary school students. In this setup, we aimed to investigate whether students from different socioeconomic neighborhoods profit differently from such learning activities, taking into account gender differences. Our findings indicate that students' socioeconomic background is related to their pre-workshop attitude about programming, and it has an influence on how students' attitude changed during the workshop. Accordingly, the workshop did not cause a significant change in students' attitude about programming in the high-income school, but it did have a positive effect on students' attitude in the middle- and low-income school. Regarding students' learning outcomes we also found that the workshop was the least effective with students from the high-income school, while students from the low-income school outperformed students from the high-income school in terms of their task-based performance. Our findings, thus, shed light on the previously understudied effect of the socioeconomic background and students' attitude about programming and their learning outcomes during the course of a non-curricular playful programming workshop. Based on our findings we suggest that targeting with similar activities students with a middle and low socioeconomic background is more beneficial in terms of attitude change and learning outcomes than targeting students with a high socioeconomic background.

Part V.

FUN AND LEARNING IN REFLECTION OF PHYSIOLOGICAL DATA

9 Fun and Learning in Reflection of Physiological Data[22]

In the previous chapters we introduced the FiL model, which quantifies the relationship between fun, attitude, and learning, and investigated how adolescents' self-regulatory skills and their socioeconomic background could possibly influence the FiL model. In this chapter the FiL model is put into test by a multimodal data analysis study, in which we examine how fun impacts learning during a programming activity by combining physiological data with self-reports. Therefore, the chapter contributes with a deeper understanding on how fun occurs during learning to program, and which physio-affective states can be used as a predictor of fun.

Summary

The role of fun in learning, and specifically in learning to code, is critical but not yet fully understood. Fun is typically measured by post session questionnaires, which are coarse-grained, evaluating activities that sometimes last an hour, a day or longer. Here we examine how fun impacts learning during a coding activity, combining continuous physiological response data from wristbands and facial expressions from facial camera videos, along with self-reported measures (i.e., knowledge test and reported fun). Data were collected from 53 primary school students in a single-occasion, two-hours long coding workshop, with the BBC micro:bits. We found that a) sadness, anger and stress are negatively, and arousal is positively related to students' relative learning gain (RLG), b) experienced fun is positively related to students' RLG and c) RLG and fun are related to certain physiological markers derived from the physiological response data.

9.1 Introduction

Coding skills are gaining increased attention especially as they are often considered as a core literacy skill of the 21st century [210]. Accordingly, more and more countries are introducing computer science (CS) and coding competence to their curricula[23]. Despite this ongoing momentum and development of new CS courses (e.g., [146]), currently, children's participation in out of formal education activities is the main way children obtain competence in CS and coding. Designing fun and engaging learning activities is essential to attract children as fun provides the affective coloring for all our day-to-day events and interactions.

---

[22] This chapter is based on the following publication: Tisza, G., Sharma, K., Papavlasopoulou, S., Markopoulos, P., & Giannakos, M. (2022). Understanding fun in learning to code: a multi-modal data approach. In Interaction Design and Children (IDC '22), June 27–30, 2022, Braga, Portugal. ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3501712.3529716

[23] https://www.euractiv.com/section/digital/infographic/infographic-coding-at-school-how-do-eu-countries-compare/

In the field of interaction design and children, evaluation of fun has been largely focused on self-reported data from children, asking them to assess specific activities in single-item scales or to compare the experienced fun in relation to different elements of the design [234]. This pragmatic and widely used approach addresses the difficulties children have in responding to surveys but does not provide a theoretically grounded definition of fun and a corresponding psychometrically validated measurement. Such an undertaking is reported in Chapter 2 and 3, where a theoretical account of the nature of fun as an affective state is proposed together with FunQ, a validated questionnaire for measuring the fun children and adolescents experience during learning. This approach has arguably a better theoretical foundation and is more reliable than single item measures, but still, suffers from being coarse grained, providing a single retrospective measure for a whole activity rather than a measure that considers fun as a changing state that varies over time and in relation to the momentary activity of the child.

Despite the great promise on designing coding activities that can be both instructional and perceived as fun, there are several challenges with this endeavor. First, there are various available methods used to measure children's affect in design research, with limited agreement among researchers about the definition and an acceptable measurement of fun. Moods and emotions, as well as human's affective preferences (i.e., what someone likes or dislikes) are complex constructs, and different methods have been developed to understand and measure them. Three broad categories are the following:

1) Methods that rely on automatic affect recognition (e.g., objective signals portraying specific physiological and behavioral response patterns that represent emotions) and inspired by theories (embodiment of affect) [74, 77].
2) Methods that rely on self-report (e.g., questionnaires, rankings), such items can be of verbal or pictorial scales (e.g., Smileyometer [234]).
3) Methods that rely on text or discourse analysis (can be automated via natural language processing methods or via thematic analysis [55]).

The three categories have different strengths and weaknesses but can also co-exist allowing us to capture different aspects of children's mood and emotions and understand their affective preferences comprehensively (although the third category is not relevant for this study, since there was no discourse and text). Despite the great interest in designing fun learning activities, as yet there is little known regarding the impact of fun on learning.

To further contemporary approaches and understand children's affective preferences comprehensively, we adopt the use of a multimodal approach. In particular, our approach involves the use of objective automated measures coming from children's physiological response data (collected by wristbands and facial video recordings, the latter allowing us the extraction of facial Action Units (AUs, [76])), self-reported fun and their learning gain (via a standard test). This approach has been proven to be effective, among others, for predicting cognitive performance [271], and hence indicated that using physiological response data allows us a new level of examination. However, we know little about the nature of fun in learning, and fun in coding activities has never been examined previously from the physiological perspective. This study aims to fill the gap in the literature by investigating the relationship between the experienced fun while learning how to code (as self-reported), the learning outcomes (based on standard tests) and students' affective states derived from unobstructive subjective measurements. In particular, this study focuses on the following research questions (RQ):

RQ1: What is the relationship between students' learning and their affective states (i.e., affect from the Action Units (AUs), physiological stress and arousal) and processes during a coding activity?

RQ2: What is the relationship between students' perceived fun (as measured by FunQ) and their affective states and processes during a coding activity?

To tackle the aforementioned RQs, we designed a non-curricular 2-hour-long playful coding workshop (introducing coding with BBC micro:bits) and implemented it in six primary school classes. Our findings indicate that both students' learning (i.e., relative learning gain - RLG) and the level of fun they have experienced while coding are associated with specific set of physiological predictors. On top of that, we also found a positive and significant association between fun and students' RLG. To summarize, we present the following contributions:

1) We offer insights from a study where students, aged 8-12 years, participated in a coding workshop and their experience and learning were captured by standardized tests and physiological devices.
2) We identify the relationship between students' learning, perceived fun and affective processes (captured using the transitions among the affective states) during the coding activity.
3) We discuss how our approach and findings can be used to design future coding workshops.

9.2 Background

9.2.1 Affective Processes and Learning

Pekrun [214] introduced the Control-Value Theory (CVT) of achievement emotions by integrating assumptions from expectancy-value approaches to emotions, theories of perceived control, attributional theories of achievement emotions, and models that involve effects of emotions on learning and performance. More specifically, Control-Value Theory builds on the idea that experiencing emotions during learning is dependent on whether learners consider the learning activity important, and the extent to which learners have control over the achievement activities and outcomes [101]. Accordingly, emotions can be mapped on a two-dimensional plot based on their valence and activation, and thus we can distinguish positive activating (e.g., enjoyment, curiosity), negative activating (e.g., frustration, confusion), positive deactivating (e.g., relief, relaxation), and negative deactivating (e.g., boredom) emotions. In relation to the students learning experience during a coding activity, this research focuses on the four Control-Value Theory emotions - happiness, sadness, anger, and surprise together with physiological stress and arousal to capture the affective states and processes; while for capturing fun the already discussed definition was adopted [299]. These four CVT emotions were selected as other emotions (e.g., disgust, contempt, relief) accounted for less than 3% of the total interaction time.

9.2.2 Fun and Learning

Related research into the relationship between fun and learning has been already discussed in detail in the previous chapters. In sum, while earlier research appeared to be inconclusive on the role that fun plays on learning, recent empirical research results are supportive that fun contributes positively to the learning outcomes. This shift is proposed to be due to a better understanding of the notion of fun and accordingly, improved ways for the assessment of it. The previous chapters of this thesis contribute to this body of knowledge essentially. However, from the literature review it is also clear that the relationship between fun and learning has never been examined before in reflection of physiological data.

9.2.3 Multimodal Data and Learning

Learning is a complex process and involves cooperation and coordination of several cognitive processes (e.g., information processing, creating, maintaining and updating mental schemas) and affective mechanisms (e.g., frustration, boredom, confusion, stress, arousal;[277]). These processes and mechanisms could incur an affective disequilibrium that might be detrimental for learning, when students struggle to maintain and update their existing mental models with new information [102]. Given the range of processes involved, it would make sense that a single data stream would not be able to capture all these processes. Depending on the process of interest, combining different data streams may be more appropriate. Some of these data streams currently used within education include video, system logs, and physiological response data such as, electrodermal activity, heart rate variability, blood volume pulse, and skin temperature. Individually, these data streams have been used to explain and predict aspects of the cognitive processes and affective mechanisms [272]. By extending these findings into interventions, researchers have used the data streams to scaffold the learning process to provide better learning support to students.

Given that a single data stream cannot capture all processes happening during learning activities as each data stream can only provide a partial view when used on its own, an upcoming field of research, multimodal learning analytics (MMLA), combines several of these data streams to serve as a virtual observer and analyst of learning activities [33, 272].

MMLA provide an unprecedented opportunity to understand students' behavior and performance during and after the learning sessions by understanding their relations with cognitive processes and affective mechanisms [62]. MMLA can provide insights into a multitude of behaviors including reasoning patterns [331], short-term memory usage [154], artefact quality [280], help-seeking and help-giving behavior [63], tentative and casual problem-solving behavior [9], and problem-solving phases [10, 279]. MMLA can be used to differentiate and identify different learning processes and behaviors [279, 331], as well as to explain the relationship between two behaviors, such as a student's physical actions and their reasoning during learning [9]. MMLA can shed light to learning processes that may be invisible to the human eye and that students cannot self-report on [63, 161, 210, 269]. Therefore, MMLA can complement our understanding on how children learn, providing more information on children's affective aspects during the coding activities.

9.3 Method

9.3.1 Participants

The herein introduced study was conducted in mid-February 2020 in the Netherlands. Primary school teachers across the country were approached to participate in the study. We recruited 53 students (M_{age} = 10.13 yrs, SD = 1.103, 27 boys, 26 girls) from three schools and six school classes. Participation in the activity was compulsory for students as the workshop took place during school hours, however, participation in the study (i.e., responding to the questionnaires and allowing us to capture their screens and cameras) was voluntary. Given students' age, informed consent was obtained across the schools from both the students and their parents/guardians before the study started. The study was approved on 10 January 2020 by the Ethics Review Board of Eindhoven University of Technology, Department of Industrial Design.

9.3.2 Procedure

A single-occasion, two-hours long workshop was designed to introduce coding with BBC micro:bits in a playful way. The workshop consisted of five main sections. First, the pre-workshop data collection section, then three distinct coding tasks and the workshop ended with the post-workshop data collection. Both the pre- and post-workshop data collection took approximately 10 minutes. Children had approximately 90 minutes to spend on the coding tasks. The coding tasks were guided by the videos provided by the SkillsDojo Foundation, just as in the studies introduced in earlier chapters. The first coding task had an introductory nature, during which students learned the basic properties of the micro:bits and thereafter they learned to program their names. In the second task, students programmed a stone-paper-scissors game. In the third task, students could create a micropet that reacted to kinetic stimuli (guided by the instructional video) or they could choose to create their own code. By their nature, the coding tasks required individual work, however, collaboration was also allowed and facilitated by the researchers.

9.3.3 Data Collection

To address the research questions, multimodal data were collected. Alongside with students' demographics, we collected their self-reported fun via questionnaires, their pre- and post- coding competence via a test, facial expressions from facial videos and physiological arousal and stress from wristband sensors. The setup is displayed on Figure 9.1 below.

Figure 9.1 Setup during the workshop. Left top: questionnaire, left middle: micro:bit, left bottom: wristband on student's wrist, middle: laptop with facial camera, right middle: device displaying and storing wristband data.

In particular, the pre-workshop questionnaire, captured students' demographics and background information that included their perceived experience and knowledge on coding, using a 5-point Likert scale ('Do you have any idea about programming?' (1) not at all - (5) I'm a pro; 'How many programming activities have you participated before?' (1) none - (5) six or more; see Figure 9.2).

2. Do you have any idea about programming? (mark one)

Not at all	I know a bit	I know something	I know much	I am a pro

3. How many programming workshops (e.g. on Scratch, Alice, Lego Robots, Python) have you participated before? (circle one) (mark one)

None	1	2-3	4-5	6 or more

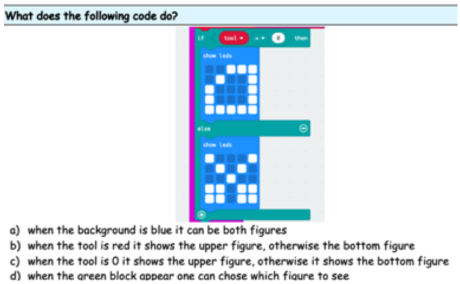


Figure 9.2 *Left*: Example of the questionnaire. *Right*: Example of the knowledge test.

To assess their coding competence, we employed a pre- and post-workshop knowledge test. This allowed us to assess students' competence levels before and after their participation in the workshop, and calculate their RLG (*knowledge* level of Bloom's taxonomy [34]). The test was developed specially for the purpose of the study to cover the material of the how-to videos. It consists of seven multiple-choice questions with four response options, from which four ask about terms that are explained in the videos (e.g., What/who is a variable?) and three questions address the working of a piece of code – which are necessary to complete the programming tasks (see example in Figure 9.2; this piece of code is part of the stone–paper–scissors game (task 2)).

For the assessment of fun during the workshop, we employed the FunQ [299] instrument as part of the post-workshop questionnaire. FunQ is a validated instrument in several languages (including Dutch) and consists of 18 easy to understand (considering students' age) questions.

Besides using questionnaires, we collected students' physiological response data. We collected arousal data via wristbands and facial expression via facial cameras data of 11 randomly selected students in each workshop, thus from 66 students in total. However, data from 13 students was damaged or lost during recording, hence our data set used for the analysis comprises of data from 53 students. Those multimodal data were collected while the students were engaged with coding tasks. Regarding the data collection of the different data modalities, we used the Empatica E4 wristband to capture students' physiological response data consisting of 4 different variables: Heart rate variability (HRV, 1Hz), Electrodermal Activation (EDA, 64Hz), skin temperature (4Hz), and Blood Volume Pulse (BVP, 4Hz) and for the facial video we used the web camera of each laptop the students were working on. The frame rate was set to 24 frames per second.

9.3.4 Measurements

Relative Learning Gain (RLG): To address the previously discussed limitation (in Chapter 8) of the frequently used learning gain (i.e., difference on post-test and pre-test scores), in this study we have decided to use a more sophisticated measure. From the pre and post knowledge acquisition test, we calculated students' RLG that has been used previously in similar studies [208]. This measure is more accurate than typical learning gain (i.e., difference between the post-test and pre-test scores), since it considers students' initial knowledge when assessing learning gain and avoids potential floor effects. RLG captures how much students learn beyond what they knew prior to the intervention.

$$RLG = \begin{cases} \dfrac{Posttest - Pretest}{Max.\,in\,pretest - Pretest}, if\ Posttest \geq Pretest \\ \dfrac{Posttest - Pretest}{Pretest}, if\ Posttest < Pretest \end{cases}$$

Fun Dimensions: FunQ [299] was employed to measure the experienced fun along its six dimensions, FunQ has eighteen questions (items), and it uses a 5-point Likert scale. The six dimensions are Autonomy (perceived control over participation and the activity itself), Challenge (experienced challenge), Delight (perceived positive emotions and related desires), Immersion (perceived loss of time and space), Loss of Social Barriers (perceived social connectivity), and Stress (perceived negative emotions).

Affect: We used the face images coming from the videos to extract the facial Action Units (AUs, [76]) using the OpenFace framework [7]. Facial Action Coding System (FACS) is a taxonomy for human facial movements as they appear on the face. Movements of individual facial muscles are encoded by FACS from slight instant changes in facial appearance. Using FACS it is possible to code nearly any anatomically possible emotion, deconstructing it into the specific AU that produced the facial expression. Accordingly, a combination of AUs define an emotion, and each AU is a set of facial muscles moving. Therefore, there can be one AU that is contributing to more emotions and many emotions can have one AU in their set of difining AUs, because one emotion can take many muscles to move, and it can overlap with other emotions. FACS is an established scheme for coding facial expressions, which is supported by multiple studies that have evaluated FACS with positive results with adults [57, 262]. Additionally, studies used the scheme in the previous years with children with positive results as well [209, 210]. Furthermore, it is a common standard to objectively describe emotions from facial expressions using such techniques [309]. Figure 9.3 shows the AUs detected for this work, and Table 9.1 shows how to define emotions from the AUs.

In this study, we are using the proportion of each emotion during the coding activity. We define happiness, sadness, anger, and surprise from the action units (shown in table 9.1). Therefore, despite we detected more AUs, we used the thirteen AUs that define the four affective states we focused our analysis on. These affective states are a subset of achievement emotions included in Control-Value Theory [214]. These four CVT emotions are use because other emotions (e.g., disgust, contempt, relief) make less than 3% of the total interaction time. Therefore, we discarded the emotions that are not detected with a significant proportion of the interaction time. The interpretation of facial expressions can change from one situation to other however, the coding is well-evaluated and the qualitative interpretation in the context we studied will be done in our future analysis. This study focused more (being the first of its kind, to the best of our knowledge) on finding the relationships between CVT-based emotions and sensor and facial data.

Table 9.1 Defining emotions from Action Units.

Affective state	Action units	Affective state	Action units
Happiness	AU6, AU10	Anger	AU1, AU2, AU5, AU26
Sadness	AU1, AU4, AU15	Surprise	AU4, AU5, AU7, AU23

Affective states transition: the second set of measurements were the transition probabilities between two affective states. These transitions capture the affective process during the coding activity. We did not consider the self-loops in this work, because we are already using the proportion of the duration of each individual emotion as the first set of measurements.

Figure 9.3 Action Unites detected for this study. The facial images are taken from https://www.cs.cmu.edu/~face/facs.htm The action unit numbers are mentioned in parentheses next to the action unit names to have a mapping with Table 9.1.

Physiological Stress: This is computed as the heart rate's increasing slope. The more positive the slope of the heart rate is in a given time window, the higher the stress is [291]. Heart rate has been used to measure stress in educational [272] and problem-solving [191] contexts. In the rest of the chapter, physiological stress is referred to as stress among the physio-affective states and processes.

Physiological arousal: EDA signal is comprised of two parts: the tonic and phasic components. The tonic component of the EDA signal is the one with slow evolving patterns. The phasic component of the EDA signal is the one with rapid changes and is found to be related to physiological arousal [159]. In this work, we consider only the mean phasic EDA component as a measure of physiological arousal. In the rest of the chapter, physiological arousal is referred to as arousal among the physio-affective states and processes.

9.3.5 Data Pre-processing[24]

To remove noise, and potential conditional biases from the sensor data, the following pre-processing was conducted.

[24] The multimodal data analysis described in this chapter, including the pre-processing of the physiological data was done by Kshitij Sharma.

Wristband data: A simple smoothing function was used to remove any unwanted spikes in the time series in the 4 data streams originating from the E4 wristband (HRV, EDA, Skin Temperature, and BVP). This was a simple running average with a moving window of 100 samples, and an overlap of 50 samples between two consecutive windows. Physiological response data, such as HRV, BVP, and skin temperature, are susceptible to many subjective and contextual biases. These biases include the time of the day, physical health condition, gender, age, overnight sleep, and others. All 4 data streams were normalized using the first 30 seconds of the data (which decision was based on earlier studies [95, 96, 230]) to remove the subjective and contextual biases from the data. Normalizing the data allowed us removing personal biases, especially, as we were not interested in absolute values, but we were interested in the variations in those.

Facial data: For most of the frames in the video recordings, only one face was visible. However, sometimes the researcher overseeing the activity appeared in the field of view of the camera. For some other frames there were a few other students in the frame as well (visualized in the Figure 9.4). First, we used the OpenFace [7] library in the videos, in order to detect the faces for every frame. Thus, each face is given a label starting from left to right (1 to N, where N is the number of faces in each frame). There are three cases where the left-to-right labeling of faces fails as shown in Figure 9.4. First, when students are with the teacher and/or the researcher. Second, when classmates join the student for a short time. We need to keep the face to which the recording belongs. To achieve this, we used a pre-trained deep neural network, INCEPTION-v4 [289], to extract features from the individual face images and used a k-nearest neighbor prediction algorithm to recognize the original student in every recording. Figure 9.4 shows the example for all the three cases. The first few minutes are used to create the feature vectors for the original student in each recording.

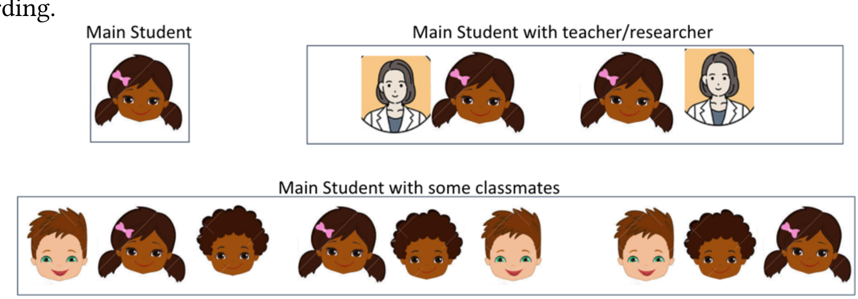


Figure 9.4 Facial data pre-processing. Detecting the main student.

9.3.6 *Data Analysis*

First, to get a better idea of the results of our study, a descriptive and correlational analysis has been conducted on the main variables. Then, to address our RQs, appropriate data analyses have been conducted.

To examine possible predictors of students' learning from their physio-affective states (i.e., affect from AUs and physiological stress and arousal; RQ1), multiple regression equations were calculated. We use the RLG as the dependent variable and all the measurements from the facial and wristband data as the regressors in a regression model. The adjusted R-square value of the models shows the variance of the RLG explained by the physio-affective variables. Second, we use a t-test to find which individual coefficients from the regression model contribute significantly to the dependent variables and explain the relationship between the RLG and the physio-affective states. For multiple t-tests, the p-values are corrected using a Bonferroni correction.

Similarly, to examine possible predictors of students' perceived fun (the dimensions measured by FunQ) from their physio-affective states (RQ2), a series of multiple regression equations were calculated (Table 9.2-9.4).

To check the gender bias in the data, we use one-way ANOVA with the physio-affective measurements, FunQ dimensions and RLG as dependent variable and gender as the independent variable. Regarding the age bias in the data, we use Pearson Correlation between age and the other measurements (i.e., physio-affective measurements, FunQ dimensions and RLG).

9.4 Results

From the descriptive analysis, we found that just above half of the students in our study were novices to coding. For the question 'Do you have any idea about programming?', 22.6% of the students reported on having no idea about coding, and 35.8% of the students reported knowing a bit. The mean for this question is 2.38 (with 5 being the highest) (SD = 1.105) which also translates into 'knowing a bit'. As for the responses for the question 'How many coding workshops have you participated in before?', 43.4% of the students reported never having participated in a coding activity, and an additional 13.2% reported having participated in 1 coding activity only. The mean of the responses for this question is 2.25 (with 5 being the highest) (SD = 1.356). Therefore, some of the students who participated in our study had some previous experience with coding, and most of them had none or very limited. When it comes to students' relative learning gain it was on an average 0.61 (SD = 0.22, min = 0, max = 1).

Children's average FunQ score is 71.55 (SD = 9.756; *Cronbach's alpha* = 0.819, min = 50, max = 89), which is quite high if we consider that the possible scores range from 18 (lowest fun) to 90 (highest fun).

We also checked for any age- or gender-related biases for the RLG and the different FunQ dimensions (i.e., autonomy, challenge, delight, immersion, social barrier, stress). There was no correlation between the age of the students and their RLG or any of the FunQ dimensions. However, there was one exception. The social barrier was higher for boys than that for girls ($F(1,37) = 4.63$, $p = 0.03$, nine students had missing values). As we show in the main analysis that we did not find any significant relationship between the social barrier and physio-affective states, this bias will not be discussed in the light of the results reported in this chapter.

9.4.1 Results from Modeling the Relative Learning Gain (RQ1)

We modeled the relative learning gain (RLG) using the proportions of emotions, the transition among them, stress, and arousal. The overall model was significant ($F(10, 37) = 10.41$, $p < 0.001$, $R^2 = 0.72$), accounting for 72% of explained variance in students' RLG. We found *arousal* and the *transitions between happiness and surprise* to be positive predictor for RGL, while *sadness*, *anger*, *stress*, and *transition between sadness and anger* contributed negatively to students' RLG. The coefficients of the significantly contributing predictors are in Table 9.2, the complete model is to see in Appendix I.

Table 9.2 The model for RLG with Control-Value Theory affective states, transitions among them, arousal, and stress. This table shows only the significant terms.

	β	**Error**	***t*-value**	***p*-value**
Intercept	0.27	0.69	0.52	> 0.05
Sadness (sad)	-1.37	0.009	-2.38	0.01
Anger (ang)	-1.56	0.003	-3.12	0.001
Trans:hap <-> sup	1.13	0.04	2.03	0.02
Trans:sad <-> ang	-1.88	0.004	-3.42	0.0006
Stress	-1.03	0.01	-1.78	0.04
Arousal	1.19	0.02	1.71	0.04

9.4.2 Results from Modeling the FunQ Dimensions (RQ2)

We modeled the FunQ Total Score and all the dimensions using the proportions of emotions, the transition among them, stress, and arousal. Below are the details for each of the dependent variables.

FunQ Total Score: The overall model for the total score of FunQ was not significant ($F(10, 37) = 1.44$, $p = 0.20$, $R^2 = 0.26$). We have provided the model details in Appendix I. In other words, from the used physio-affective states we could not predict the total score of the FunQ.

FunQ Autonomy: The overall model for the Autonomy dimension of FunQ was not significant ($F(10, 37) = 1.47$, $p = 0.19$, $R^2 = .32$). We have provided the model details in Appendix I. In other words, from the investigated physio-affective states we could not predict the FunQ Autonomy scores.

FunQ Challenge: The overall model was significant ($F(10, 37) = 10.19$, $p < 0.001$, $R^2 = 0.71$), accounting for the 71% of explained variance in students' FunQ Challenge. H*appiness*, *anger*, *arousal*, and the *transitions between happiness and sadness* predict FunQ Challenge positively. On the other hand, *sadness*, *surprise*, and *transition between sadness and surprise* contribute negatively to FunQ Challenge. The coefficients of the significantly contributing predictors are shown in Table 9.3.

FunQ Delight: We found the overall model to be significant ($F\,(10, 37) = 9.93$, $p < 0.001$, $R^2 = 0.65$). The predictor model accounts for 65% of the explained variance in students' FunQ Delight. In details, *happiness* and *surprise* are positive predictors for FunQ Delight whereas, *stress*, and *transition between happiness and anger* contribute negatively to FunQ Delight. The coefficients of the significant predictors are shown in Table 9.3.

Table 9.3 The models for FunQ Challenge and Delight with control-value theoretic affective states, transitions among them, arousal and stress. This table shows only the significant terms (p < 0.05).

Model Challenge	β	Error	t	***Model Delight***	β	Error	t
intercept	0.31	0.12	0.43	intercept	0.44	0.24	0.54
Happiness	1.41	0.005	2.92	Happiness	1.34	0.051	2.13
Anger	1.02	0.004	3.13	Surprise	1.93	0.009	3.19
Sadness	-1.32	0.012	-2.10	Trans. Hap-Ang	-0.93	0.001	-4.18
Surprise	-1.35	0.003	-2.44	Stress	-1.34	0.003	-3.53
Trans. Hap-Sad	0.99	0.001	3.38				
Trans. Sad-Sup	-1.45	0.001	-3.04				
Arousal	1.23	0.014	2.23				

FunQ Immersion: The overall model was significant ($F\,(10, 37) = 8.16$, $p < 0.0001$, $R^2 = 0.63$), accounting for the 63% of explained variance in students' FunQ Immersion. H*appiness* and *arousal* are positive predictors for FunQ Immersion while the *transition between sadness and anger* contribute negatively to FunQ Immersion. The coefficients of the significant predictors are shown in Table 9.4.

FunQ Social Barrier: The overall model for the social barrier dimension of FunQ was not significant ($F\,(10, 37) = 1.09$, $p = 0.39$, $R^2 = 0.17$). We have provided the model details in Appendix I. In other words, from the investigated physio-affective states we could not predict the Social Barrier dimension of FunQ. As we mentioned earlier, there was a gender bias for this sub-construct. Boys ($M = 9.00$, $SD = 3.22$) reported a higher social barrier than girls ($M = 7.11$, $SD = 1.99$). However, because there is no relationship between this construct and the RLG or any other physio-affective measurements, we will not explore this bias in this contribution.

FunQ Stress: We found the overall model to be significant (F (10, 37) = 10.02, $p < 0.001$, $R^2 = 0.70$), accounting for the 70% of explained variance in students' FunQ Stress. We found that *sadness*, *anger*, *stress*, and *transitions between sadness and anger* predict FunQ Stress positively, while the *transition between happiness and surprise* contribute negatively to FunQ Stress. The coefficients of the significant predictors are shown in Table 9.4.

Table 9.4 The model for FunQ Immersion and Stress with control-value theoretic affective states, transitions among them, arousal, and stress. This table shows only the significant terms (p < 0.05).

Model Immersion	β	Error	t	***Model Stress***	β	Error	t
intercept	0.19	0.89	0.52	intercept	0.16	0.21	0.89
Happiness	1.46	0.001	4.34	Sadness	1.28	0.005	2.48
Trans. Sad-Ang	-1.73	0.003	-3.28	Anger	0.94	0.001	3.32
Arousal	2.01	0.001	4.26	Trans. sad-ang	1.27	0.017	2.70

9.5 Discussion

In this study we set out to investigate the relationship between students' coding learning, the experienced fun, and their physio-affective states during a coding activity. We collected data from a questionnaire, and physiological response data collected by wristbands and facial video recordings. Using data from different modalities and analyzing them we provide a novel approach as earlier research has been limited to either the investigation of affective states (e.g., by interviews or surveys (e.g., [300]) or physiological measures (e.g., [272]). By combining these we extended our current body of knowledge by adding a new, physiological level of understanding of learning procedures. One can argue that the two measurements are not exactly the same, which is evident by the results reported in the chapter. We have shown that there is a significant overlap between the retrospective measurement of fun (through questionnaire) and the spontaneous measurement of affect (through sensor data). Both measurements have been evaluated separately [57, 262, 299]. This study is an attempt to find a relationship between the two measurements to have more real-time information about the semantic beliefs and memories using the sensor data. Accordingly, we found that RLG and most of the FunQ dimensions can be explained by the CVT affective states (i.e., happiness, sadness, anger, surprise, and the transition between these). Therefore, the introduced results indicate that there is a link between learners' affective states, their learning outcomes, and the fun they have experienced while learning.

More specifically, regarding students' learning and their affective states during coding activity (RQ1), we found that sadness, anger, and stress contribute negatively on students' learning, while arousal positively on it. This finding is in line with previous research that has investigated this relationship with traditional methods (i.e., questionnaires and observations) [168]. However, it also goes beyond those by applying MMLA and physiological measures.

Our research results indicate that from physiological data we could not predict the level of fun - measured as the total score on FunQ – that students experienced while learning to code. Nevertheless, we found that the total FunQ score significantly correlates with the RLG (Pearson correlation = 0.33, $p < 0.05$). This finding is in line with earlier research, which suggests that having fun while learning contributes to the learning outcomes [169, 171, 260, 300, 306]. Although some dimensions of fun could be predicted from the physio-affective states of the child, we found that the physio-affective states do not predict fun comprehensibly. This aligns with previous works in physiological response measures that indicate challenges with achieving perfect one-to-one relationship between physiological response measures and psychological constructs [46].

Concerning our findings about the dimensions of FunQ, we conclude that as just mentioned, not all its dimensions could be predicted from physiological response data. Accordingly, neither the Autonomy nor the Loss of Social Barriers dimensions could be modeled by the affective states. We believe that these results rather reflect the characteristics of the activity rather than general tendencies. Namely, despite students were provided with some freedom and attributes that are atypical in a formal learning environment (e.g., they could decide whether they wanted to follow the activity, they were allowed to move around freely and ask the instructors whenever they wanted), the workshop was still scripted as it followed a fixed sequence of tasks. Hence, students might have not felt a sufficient level of autonomy (or not frequently enough) to be able to relate it to physiological response data. Regarding the Loss of Social Barrier dimension, on top of the aforementioned possible explanations, the scripted structure of the workshop might not have provided enough space for social interactions that would have led to an increase in social connectedness. Hence, it appeared not to be possible to link this dimension to physiological response data. Regarding both dimensions, further research is required to establish general tendencies as our findings might be a consequence of the activity design and be activity specific.

The Challenge dimension of FunQ could be predicted from the CVT affective states happiness, anger and arousal positively, and the transitions between happiness and sadness. These transitions' positive contribution to FunQ challenge can perhaps be because students were in a constant loop of succeeding and failing, as not everything worked at their first attempt. On the other hand, sadness, surprise, and transition between sadness and surprise, contribute negatively to FunQ Challenge. Connected to the previous finding, when students were failing the task, it could have been a sign that the task was (momentarily) too difficult for them and it can explain of why the transition between sadness and surprise contributed negatively to FunQ Challenge. During a coding activity, the students need to deal with different aspects of the tasks, like debugging, problem solving and reflecting iteratively on the needed actions, and this process can be difficult and challenging [207].

As for the Delight dimension, happiness and surprise appeared to be a positive contributor, whereas stress, and the transition between happiness and anger contributed negatively. We propose that Delight can be seen as an emotion related to solving or understanding a problem or even having a desired outcome in a given possible task [65] and this can be triggered from positive emotions or an unexpected outcome. Regarding Immersion, happiness and arousal are found to contribute positively to it, while transition between sadness and anger turned out to be a negative contributor to Immersion.

Concerning the Stress dimension, which is a contra-indicative dimension of FunQ with reversed items, we found that the physio-affective states sadness, anger, stress, and the transition between sadness and anger contributed positively, while the transition between happiness and surprise contributed negatively. In other words, the physio-affective states sadness, anger and stress are inducing stress, while changing from happy to surprised, and vice-versa is a contra-indicative signal, indicating low levels of stress.

For all the constructs that we have used in this contribution (i.e., RLG and the FunQ dimensions), stress and/or arousal have been a significant predictor. We found that arousal is positively associated with RLG and Challenge, Delight, and Immersion dimensions of FunQ; stress is negatively associated with RLG and positively associated with the Stress dimension of FunQ. The positive association of arousal and the negative association of stress with the RLG (or in other words learning or cognitive performance) is consistent with various other studies. For example, in game-based learning settings with children Lee-Cultura et al. [160] and Sharma et al. [270] found physiological stress to be negatively associated with learning performance and experiences. Similarly, Joëls et al. [134] showed that the memory-based learning performances decrease under stress. These studies are also in line with the finding that higher levels of stress are negatively associated with the RLG extends the consensus from these studies. Furthermore, the physiological response measurement of stress being positively associated with the self-reported stress is indicative of the measurements' validity in the context of children coding.

On the other hand, physiological arousal provides us with a reliable proxy of engaged behavior [36, 159, 161]. The high levels of engagement have been shown to be positively associated with learning [39, 52, 105]. In our case, higher levels of physiological arousal indicate high levels of engagement which in turn increases the probability of students with high physiological arousal also having a high RLG. Moreover, with high levels of engagement, students might also feel immersed and challenged at appropriate levels, which in turn might increase their ratings for the delight dimension of FunQ.

9.5.1 Implications

Our findings support endeavors of educators, designers, and researchers to make learning to code a fun experience, as we found a positive relationship between those. Further research studies could aim to improve the applicability of physiological measure devices (e.g., wristbands) for children. Beyond research purposes, such improvements in the physiological measure devices could pave the way for everyday (classroom) use. If the devices became more comfortable, easier to use/calibrate and non-disturbing, they could support the personalized learning experience at a new level (e.g., based on the wristband data the level of stress could be monitored and the content of the learning material could be adjusted accordingly). Our research also opens ways for at-the-moment measurement of fun that will allow us a precise insight into the activity, in contrast with the post-hoc tests and get a more holistic understanding. This way, micro-level investigations and interventions are enabled for supporting fun, leading to increased learning outcomes – a finding introduced by this study and supported by previous research indicating a clear relationship between children's perceived fun while learning and their learning outcomes [300]. Additionally, it can be particularly relevant for the development of different systems for educational purposes, to use multimodal data to support both teachers and students in their everyday learning activities. Earlier studies [85] provide a foothold for questions one should consider when designing such learning experiences (e.g., What is being personalized? or How is the personalization carried out and who are the beneficiaries?). Accordingly, supporting students and teachers in their everyday learning activities can happen for example by providing systems with affordances for reflective purposes, indicating students' disengagement to support better classroom management. We propose, this feature can be especially beneficial for junior teachers, or for teachers of bigger groups, as this can help teachers with spotting disengaged students early on and help them to get back on track. From the students' perspective it can also be helpful because they will be able to signal when they are in need for more support from the teacher/instructor. Moreover, integrating affordances for reflective purposes will provide students with insights to their bodily reactions (making students more conscious about them) and the translations of those into practical matters will help them to deal with, for example, stressful situations. An example for such system message is as follows: 'Your heart rate jumped from 70 to 80. It seems that you are stressed. Why don't you take a break?' or 'It seems that you are stressed and that you have a bug in your code. Why don't you look at this example to solve the bug in your code?'. Future systems with different functionalities can also exploit multimodal data, to, for example, automatically adjust the difficulty level of a learning task, providing personalized learning to students on a given task. Personalization should then also take into account the learning setup, whether the task requires individual work or collaboration, an ultimately, at which level the personalization should happen (standardization, personalization, customization, or individualization

[156]). Knowing the affective state of the students can be powerful information helping them to overcome affective states that may hinder their learning or fun during coding.

9.6 Limitations and Future Work

Besides the applicable findings and the new approach introduced, the limitations of this study should be mentioned. First, we highlight the practical difficulties involved in collecting of physiological response data from children as these technologies are designed for adults. One example is the difficulty we faced in attaching the wristbands to some of the students' wrists. This led to some uncontrolled data loss. To resolve this issue, further research could assist the design of the wristbands to be more suitable for young users. In the same lines, the use of sensing devices increases students' curiosity, therefore researchers need to spend time to explain in simple words how each device works, what data we collect and why, letting the children interact with them. One example from our study is that we observed that some children wanted to see on the mobile device connected with the wristband how their heart rate is shown or how it changes. While satisfying students' curiosity has the potential to increase their involvement, interacting with the sensing device, if uncontrolled, can lead to some data removal.

Second, the coding activity was designed as a non-curricular activity, but in a classroom setting, aiming to provide participating students with autonomy over their participation and the activity itself. Since we did not find physiological response correlates for the Autonomy and Loss of Social Barriers dimension of FunQ, we speculate that given the activity was scripted (i.e., three tasks were given to be followed), students might not have felt the desired level of autonomy, and in relation to this, they also might not felt enough freedom to connect to each other more than usual. Future studies, hence, should examine the physiological response correlates of the FunQ dimensions in relation to a broad range of learning activities, including possibly informal learning setups as well.

Another limitation of this study is the use of only quantitative data. We expect that triangulating the quantitative data of our study with qualitative data from interviews or observations could provide us insights at multiple levels [194]: they would support the herein introduced findings (convergence) or would shed light on eventual discrepancies (divergence), and we could gain a deeper understanding of the herein introduced dynamics (complement), ultimately leading to a better understanding of students' learning-related behavior. The study of Lee-Cultura, Sharma, and Giannakos [160] is a good example on how to triangulate results from a qualitative analysis, a mixed-method study, and a predictive analysis. Lastly, more studies are needed to better understand the cognitive and affective states of students during coding and to monitor how they may shift naturally or not with the ultimate goal to offer more effective and efficient learning experiences.

Although facial expressions have been used in many studies to extract emotions, this method comes with some limitations [22, 114], such as annotation and label subjectivity, cultural differences in emotion expression, dependency on benchmark data sets, occlusions, computational efficiency and computer vision. In our case, occlusions, and limitations in relation to computer vision were the most prominent. Namely, facial data collection is dependent on computer vision, as for example, the light in the room must be sufficient, the camera should be in a proper angle to capture the whole face etc. Accordingly, we paid special attention during our study to prevent these issues. But beyond these, the person captured should sit in front of the camera, without covering (part of) their face (e.g., resting on elbow). In the herein introduced study we removed those occlusions when occurred from the data, but there were not many instances, and accordingly, we believe that they did not affect the herein introduced results.

While this study is the first to connect FunQ and sensor data focusing on a quantitative exploration, we call on further research, including both qualitative, quantitative, and especially mixed-method approaches to provide more insights into this relationship, and to triangulate the results of the herein introduced study. Hence, in future mixed-method studies participants could be closely monitored on the individual level in order to pinpoint moments that are detrimental or helpful for learning and/or fun, and those moments could be studied in depth to extend our understanding on the topic. Additionally, future mixed-method studies could also help removing contextual bias from the data, for example, by video coding, to clarify whether the observed variation at the physiological level occurred from a task-related activity, or from something else (e.g., smiling about the task or smiling because the neighbor was joking).

9.7 Conclusion

We contribute to the literature regarding the role of fun in how children learn to code in a number of ways. First, we investigated fun, a construct, which is frequently in the focus of evaluation in design and educational research, however, our knowledge is still limited about its nature. By using multimodal data, we went a step further than earlier research as it either pertained to surveys or to physiological response data only. Using the combination of the two allowed us a deeper understanding on how fun occurs during learning to program, and which physio-affective states can be used as a predictor of fun. Being able to predict fun from physiological signals can help assess different learning activities, but potentially can be developed further to support timely interventions to get disengaged students on track again, ultimately leading to better learning outcomes. In contrast to surveying students about their level of fun, using physiological response data by its unobstructive nature can provide us with immediate feedback, without disrupting the learning experience (surveying several times during an activity) and without inducing recency bias (surveying once at the end). Developing new tools or further improving existing ones that address the potential of unobstructive physiological response data could support both teachers and students in their everyday life, by providing systems with affordances for reflective purposes, indicating students' disengagement or other features to support better classroom management.

Part VI.

AN EXTENDED MODEL OF FUN IN LEARNING

10 Fun, Self-Regulation, and SES in Learning – An Extended Model[25]

In the previous four parts of this thesis (Part II, Part III, Part IV, and Part V) we introduced several studies that examined the relationship between fun and learning, taking into account attitude as a moderating element, self-regulation as a personal, and socioeconomic background as an environmental influential factor. We have examined fun in learning to code by traditional (i.e., survey; Part III-IV) and novel (MMDA, Part V) methods. We have proposed the FiL model in Chapter 6, which we tested in various settings. In this part of the thesis, we introduce a final study that aimed to investigate simultaneously the previously considered factors that could affect the relationship between fun and learning. Therefore, this chapter contributes with an extended model on the role of fun in learning.

Summary

Due to the technological advancements of the 21st century, increasing students' interest in STEM (Science, Technology, Engineering and Mathematics) subjects has become a worldwide pursuit. However, our understanding is limited on what and how influences a) students' willingness to participate in such activities, and b) their learning outcomes. This study aims to broaden our knowledge in this topic, with specific focus on programming, by investigating the relationship between students' attitude about programming, the level of fun they experience while learning, their learning outcomes, self-regulation (as a personal factors) and socioeconomic background (as an environmental factor). To investigate these questions, we designed and implemented a 90-minute AI programming workshop with 122 secondary school students (M_{age} = 12.47, SD = 0.501). Our results indicate that fun has a significant and positive direct effect on students' attitude about the topic, that students' attitude about the topic influences positively and significantly their learning, that the indirect effect of fun on learning across attitude is also significant, and that students' socioeconomic status has a negative and significant effect on students' attitude about the topic. We did not find a significant relationship between self-regulation and any of the study dimension. In this chapter we provide explanation for our findings and highlight possible directions for future work.

10.1 Introduction

In the recent decades the importance of STEM (Science, Technology, Engineering and Mathematics) teaching and learning has become apparent due to the technological advances of the 21st century. Accordingly, related educational activities are gaining momentum both within the formal and the informal context. Increasing children's interest

[25] This chapter is based on the following publication: Tisza, G., & Markopoulos, P. (2022). Path analysis to understand better how self-regulation, socioeconomic background, and having fun while learning influence students' attitude towards programming and their learning outcomes. *Manuscript submitted for publication.*

in scientific topics from early ages on thus has become a worldwide pursuit, with specific focus on computer science, as computational thinking and programming is often regarded as one of the main literacy skills of the 21st century [210]. Despite this worldwide pursuit, our knowledge is limited on what factors influence children's interest and learning outcomes in STEM subjects in general, and in coding or programming in specific.

In the previous chapters of this thesis, we have introduced a number of potential influential factors on children's and adolescents' interest and learning outcomes in STEM related activities. However, we also have shown that these factors are interrelated with each other, and hence, to get a better understanding on these relationships, the analysis of a complex model is required. Therefore, we decided to investigate factors previously studied in this thesis simultaneously by applying path analysis. Path analysis allows the examination of complex models, and *"in particular, it can examine situations in which there are several final dependent variables and those in which there are "chains" of influence, in that variable A influences variable B, which in turn affects variable C"* ([282], p. 115). Accordingly, the herein introduced study aimed to broaden our knowledge on this subject by investigating the relationship between students 'attitude about learning to program, the level of fun they experience while learning, their learning outcomes, their socioeconomic background (as an environmental factor) and their self-regulatory skills (as a personal factor). We designed a 90-minutes long workshop for secondary school students to introduce programming with AI in a playful way, which we implemented with five school classes and 122 students in total. In the remained of this chapter, first, we introduce our background theories, which provide ground for our hypotheses, then we present the study methods and our research findings, followed by the discussion and interpretation of the study findings in reflection of earlier research. We close the chapter with the study limitations and directions toward future research.

10.2 Background

10.2.1 Fun, Attitude and Learning

Most research into the relationship between fun and learning is to be found in the context of educational technology, such as gamification and child-computer interaction research. While earlier research appeared to be inconclusive whether is a relationship between these two or not [127, 275], more recent research tend to point toward one direction. Namely, a number of recent research has found supportive evidence that a relationship exists between fun and learning, and that this relationship is positive [169, 171, 260, 306]. In Chapter 6 we proposed the FiL-model, in which we described not only the positive relationship between fun and learning but considered attitude as an important moderating element. The FiL model was further supported by the studies introduced in Chapter 7 and Chapter 8. Based on these earlier findings in this current study we hypothesized that the FiL model is supported in the context of programming AI (H1), and in specific, we expected the previously described relationships between the FiL model elements (H1a – H1d; see Figure 10.1).

10.2.2 Self-Regulation, Motivation, Learning, and Fun

We have discussed extensively previous literature in relation to self-regulation, motivation, learning and fun in Chapter 7. Here, we briefly summarize those earlier findings, which provide ground for our hypotheses in the current study.

Zimmerman defined self-regulation as *"self-generated thoughts, feelings and actions that are planned and cyclically adapted to the attainment of personal [learning] goals"* ([346], p. 14). According to Pintrich [221], self-regulated learning is an active and constructive process, during which students define the learning goals for themselves, and parallel, they also regulate, monitor, and control their cognitive and motivational processes to attain their self-set goals.

Regarding the relationship between self-regulation and learning (or students' academic achievement), previous research found an important correlate, namely, attitude, to play a crucial role. According to Pintrich [221], highly motivated and strongly self-regulated students make the academically most successful ones. Flavell [86] provided an explanation for this by investigating the use of cognitive and metacognitive strategies. He proposed that more motivated students tend to utilize a wider range of strategies, leading to more efficient learning. However, when we tested the relationship between self-regulation and attitude by means of path modeling (Chapter 7), no supportive evidence was found. While we have proposed possible explanations for those findings, in this study, based on the work of Pintrich and Flavell we hypothesized that self-regulation has a positive effect on students' learning outcomes (H2) and that students' attitude has a positive influence on their self-regulation as well (H4).

As for the relationship between self-regulation and fun, previous research is scarce and inconclusive – but it is important to mention that previous studies mostly investigated enjoyment rather than fun. Artino and Jones [15] found in the online learning setting that the enjoyment of the learning activity was significantly associated with students' self-regulated learning behaviors, however, their study pertains to the examination of self-regulated learning behaviors and emotions (among others, enjoyment), and did not examine learning outcomes. Pekrun, Goets, Titz, and Perry [215] found similar results when investigating academic emotions and self-regulated learning: enjoyment was found to be significantly correlated with self-regulated learning. However, their study also lacks the investigation of students' learning outcomes. This issue is addressed in the study of An et al., [8], where both self-regulated learning, enjoyment and learning outcomes were examined. Their study results provided further supportive evidence for the relationship between self-regulation and learning and indicated that the enjoyment of the learning activity has a positive influence on students' self-regulation. Based on these findings, thus, our earlier hypothesis is further strengthened regarding the positive relationship between self-regulation and learning (H2), and we also hypothesized that having fun while learning has a positive influence on students' self-regulation (H3).

10.2.3 Socioeconomic Status and Learning

In relation to socioeconomic status and learning, we have also discussed previous literature extensively in Chapter 9. Therefore, here we shortly summarize those earlier findings, which provide ground for our hypotheses.

Investigating the relationship between students' socioeconomic background and their academic achievement has a long tradition. The meta-analysis of White [326] already back in the 80s concluded that the way SES is defined and measured (i.e., unit of the analysis) influenced the strength of the relationship between SES and academic achievement. A more recent meta-analysis in the US context [276] investigating the relationship between SES and academic achievement (in general) found further supportive evidence that there is an overall positive correlation between those two. However, Sirin [276] noted that he observed a change in the strength of the correlation between SES and academic achievement, namely, the correlations has become weaker over time. He also found that the strength of the correlation between SES and academic achievement increased significantly from primary school to middle school, suggesting that the gap between low- and high-SES students is mostly likely to remain the same over time, if not widen. More recently, Hernandez, Cascallar and Kyndt [244] conducted a meta-analysis on the relationship between SES and academic achievement, but in the higher-educational context without specifically focusing on one country's or region's research outputs. They found a weak correlation between the two, indicating also that academic performance is more strongly related to, among others, prior academic achievement or working status than to SES. A very recent review [166] on the same relationship but in the context of primary and secondary education found a moderate correlation between SES and academic achievement across the world, concluding that despite worldwide efforts to increase educational opportunities, the applied measures do not seem to reduce inequalities in students' academic outcomes between low- and high-SES students. In the Chinese context, Liu, Peng, and Luo [165] found a moderate relationship between SES and academic achievement, and that - in their context - the strength of this relationship has gradually decreased over the past decades. Moreover, they also found that in their context SES was more strongly related to language achievement than to science/math achievement.

However, while in general, as just shown, the relationship between SES and academic performance is well studied, we know much less about the relationship between SES and STEM education in specific. Blums et al. [35] in their longitudinal study examined the effect of early years SES and children's later STEM achievement. They found that the maternal education (as the main indicator for SES) had a strong influence on how children's cognitive abilities developed, which abilities on the long term were found to be related to their STEM achievement. In other worlds, the higher the mother's education level, the stronger the relationship would be with their young children's cognitive skills (on the short term), and those children's STEM achievement at older ages. In our earlier study (see Chapter 8) with primary school students we found that students from low- and middle-income schools benefitted the most from a playful programming workshop in terms of knowledge acquisition and attitude change, however, the intervention was not effective with high-income students. Investigating middle school students' STEM interest and their socioeconomic background, Yerdelen, Kahraman and Tas [334] found positive attitudes towards pursuing a STEM career, however, their study pertained to low-SES students, hence, the interpretation and generalization of their results is limited. Niu [201] found with college students that, on one hand, low-SES students were disadvantaged in pursuing a STEM major due to lack of information from their environment, on the other hand, the gender and racial gaps in STEM enrollment narrowed for high-SES students.

According to the aforementioned research, in this study we hypothesized that students' SES influences their attitude about the topic (H5), that there is a positive relationship between SES and student's learning outcomes (H6), and that SES has an indirect effect on learning through the fun students experience while learning (H7).

10.2.4 Research Question and Hypotheses

The driving motivation of this research was to broaden our understanding on students' interest and learning in STEM subjects by link earlier research findings and confirming the following hypothesized relationships between self-regulation, socioeconomic status, and fun in learning. Thereby, this study aims to extend the FiL model (introduced in Chapter 6) with two additional factors simultaneously, which previously have been shown to have importance on students' learning outcomes. Namely, with students' self-regulatory skills (as a personal factor) and their socioeconomic background (as an environmental factor). Accordingly, we hypothesized that (see Figure 10.1):

- **H1**: The FiL model is further supported in the context of programming AI, and with secondary school students.
    - o **H1a**: fun has a positive direct effect on students' attitude about the topic.
    - o **H1b**: Students' attitude has a positive direct effect on learning.
    - o **H1c**: fun has a positive direct effect on students' learning.
    - o **H1d**: fun has a positive, indirect effect on students' learning across students' attitude about the topic.
- **H2**: Self-regulation has a positive effect on learning.
- **H3**: fun has a positive effect on self-regulation.
- **H4**: Student's attitude about the topic has a positive effect on self-regulation.
- **H5**: SES influences students' attitude about the topic.
- **H6**: SES influences the learning outcomes.
- **H7**: SES has an indirect effect on learning by the mediating the effect of fun.

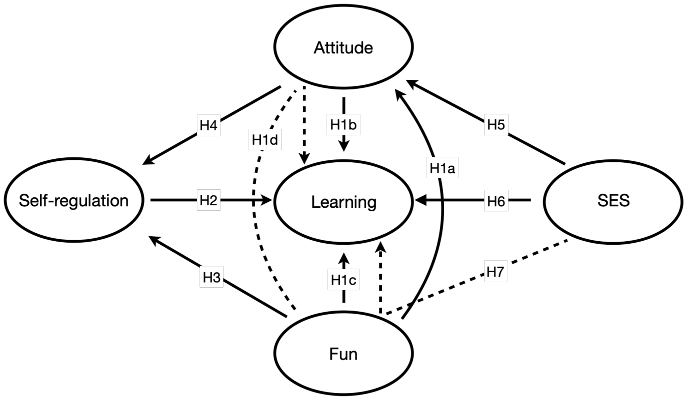


Figure 10.1 Hypothesized relationships between the model elements. Straight line indicates a direct relationship. Dashed line indicates an indirect relationship.

10.3 Method

10.3.1 The Activity

For designing the activity we followed similar steps as described in earlier chapters (Chapter 4, 6, 8, and 9), however, in the current study we aimed to collect data from slightly older children than before (12 to 14 years in comparison with 10 to 12 years), and hence, we have selected another series of videos that were better fitting to this purpose. Accordingly, to test our hypotheses, we designed a 90-minute AI programming workshop targeting 12 to 14-year-old students. The main aim of the workshops was to introduce the working of AI to secondary school students in a creative and playful way, and hence, to get them acquainted with the basic terms, and evoke their interest towards the topic. Accordingly, the workshop followed a learning-by-doing approach, and it utilized a set of instructional videos[26], thereby supporting students' individual needs for processing the workshop material in their own pace. The instructional videos were created by the SkillsDojo foundation, and they follow the MIT's Lifelong Kindergarten approach [237]. For the workshops an introduction video and three tasks were prepared. The three tasks reflected the three input modalities for AI: text (i.e., creating a chatbot), picture-based input (i.e., creating a program that recognizes colors from the computer camera), and voice-based input (i.e., creating a program that recognizes the tone of emotions (i.e., positive, neutral, negative)). Given the varied (but mostly basic) previous experience of the students with programming in general, most of them could only finish within the timeframe of the workshop only the first task. A few of them completed the second task. Figure 10.2 depicts how students wrote the script for the chatbot on paper first, then programmed it on the visual-programming interface.

---

[26] www.skillsdojo.nl/programmeren-met-kunstmatige-intelligentie/

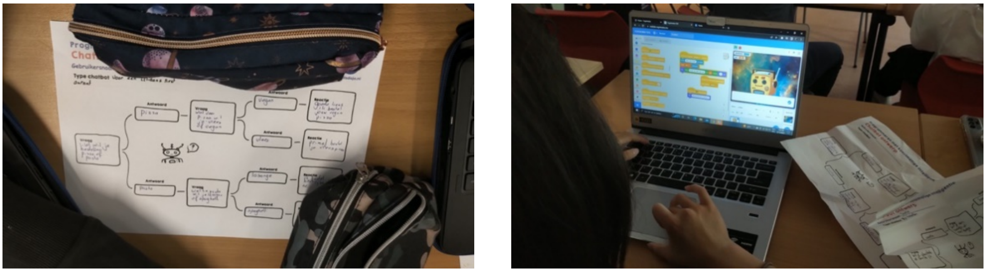


Figure 10.2 ***Left***: **Creating the script for the chatbot.** ***Right***: **Programming the chatbot.**

10.3.2 Procedure

We collected data from the participants at the beginning and at the end of the workshops. For the workshops, each student was equipped with their own laptop, and they could use their earphones as well. The workshops had the following structure:

1) Pre-workshop data collection (~10 minutes)
2) Displaying the introduction video to the whole class on the whiteboard (~5 minutes)
3) Students following the AI programming task(s) on their own pace (~65 minutes)
4) Post-workshop data collection (~10 minutes)

During the workshop, students were allowed to move around, to interact with each other, and to help each other. If they got stuck, the researcher provided them with further cues and helped when the students asked for it.

10.3.3 Participants

We applied convenient sampling to invite secondary school teachers across the Netherlands to participate with their classes in the workshop during the spring semester of 2022. In total, we approached three teachers who all responded positively. Due to personal circumstances, we could collect data from the classes of two out the three teachers. The specific activity was not part of the curriculum, though the workshops took place in the classroom during regular school hours. We note here that in the Netherlands teachers are in general free to design learning activities that fulfill the curricular goal, and our study was in line with their current teaching approach. Therefore, students' participation in the workshop could be a compulsory learning activity, but their participation in the research study (i.e., responding the survey) was voluntary. Given the students' age, informed consent was obtained from their parents across the schools, and the data was collected accordingly. In the study, in total, five school classes participated from two secondary schools, both within the agglomeration of Eindhoven, with an approximately equal average yearly income in the neighborhood[27]. Therefore, the herein introduced results are based on

[27] https://www.allecijfers.nl/; Average yearly income in the neighborhood of school A is €29.900, and in school B is €28.400.

data from 122 students (65 boys, 52 girls, 5 prefer not to say or not given), between age 12 and 13 (M_{age} = 12.47, SD = 0.501).

At the beginning of the workshop, students reported on their previous knowledge about AI across two 5-steps Likert-type scale. The questions and the frequency of the responses for the questions are displayed in Table 10.1. The sample mean for the question 'Do you have any idea about how artificial intelligence (AI) works?' is 1.81 (SD = 1.059) on a 5-step scale where (1) was 'not at all'; and the sample mean for the question 'How many times have you created your own AI program before?' is 1.17 (SD = 0.576) on a 4-step scale where (1) was 'none'). Therefore, we conclude that students were novices in the field of programming for artificial intelligence.

Table 10.1 The frequency of the responses for 'Do you have any idea about how artificial intelligence (AI) works?' and 'How many times have you created your own AI program before?'.

Do you have any idea about how artificial intelligence (AI) works?		How many times have you created your own AI program before?	
Missing	7 (5.7%)	missing	7 (5.7%)
(1) not at all	63 (51.6%)	(1) none	105 (86.1%)
(2) I know a bit	23 (18.9%)	(2) 1	3 (2.5%)
(3) I know something	19 (15.6%)	(3) 2-5	5 (4.1%)
(4) I know much	8 (6.6%)	(3) 6 or more	2 (1.6%)
(5) I know a lot	2 (1.6%)		

10.3.4 Measures

To compare the school classes along the model dimensions, we applied One-way ANOVA analysis. The analysis indicated no statistical difference in the data coming from the five workshops (p_{fun} = 0.069, $p_{attitude}$ = 0.367, $p_{learning}$ = 0.129, $p_{self\text{-}regulation}$ = 0.444, p_{SES} = 0.060), hence, analyzing them all together is justified.

Table 10.2 Comparison of the five workshops along the model components. No statistical difference is found.

	workshop1	workshop2	workshop3	workshop4	workshop5
Fun	M = 49.61 SD = 13.044	M = 56.50 SD = 14.681	M = 49.81 SD = 2.676	M = 49.27 SD = 4.638	M = 53.15 SD = 4.086
Attitude	M = 2.71 SD = 0.663	M = 3.08 SD = 0.747	M = 2.74 SD = 0.521	M = 2.91 SD = 0.638	M = 2.84 SD = 0.549
Learning	M = 2.65 SD = 0.997	M = 2.78 SD = 1.060	M = 3.24 SD = 0.777	M = 3.16 SD = 0.653	M = 3.10 SD = 0.804
Self-regulation	M = 3.10 SD = 0.440	M = 3.23 SD = 0.582	M = 3.25 SD = 0.478	M = 3.30 SD = 0.406	M = 3.33 SD = 0.415
SES	M = 9.52 SD = 1.740	M = 8.67 SD = 1.426	M = 10.27 SD = 1.845	M = 10.38 SD = 1.544	M = 9.76 SD = 1.363

For assessing fun, we used FunQ, which we have already introduced in Chapter 3. The internal consistency of FunQ is satisfactory on our data set ($Cronbach's\ \alpha_{FunQ}$ = 0.780).

Additionally, both at the beginning and at the end of the workshop, we measured students' attitude towards the topic across seven items ('Do you think that AI is fun/easy to do/ easy to understand/pleasant/exciting/something I would like to do'; 'I think that AI is my thing'), which we adopted from earlier research [206, 210, 302, 305], and were evaluated on a five-step Likert-type scale. On our data set Cronbach's alpha indicates an adequate internal consistency of the seven attitude items ($\alpha_{pre\text{-}workshop}$ = 0.808, $\alpha_{post\text{-}workshop}$ = 0.811).

For the assessment of learning we used two measures that reflect two levels of learning according to Bloom's taxonomy [34]. Accordingly, we recorded a knowledge assessment test before and after the workshop and investigated students' perceived learning after the workshop. To calculate the *measured learning* (*knowledge* level of Bloom's taxonomy), we applied the earlier introduced calculation for Relative Learning Gain (see Chapter 9, section 9.3.5), which measure is more accurate than typical learning gain (i.e., difference between the post-test and pre-test scores), since it captures how much students learn beyond what they knew prior to the intervention. However, given that the participants were novices in the field of programming AI, they could mostly only finish the first of the three tasks. Since the knowledge test was designed to investigate learning from all three tasks/AI input modalities, it was not nuanced enough to capture learning reliably only from the first task, hence, we abandoned the use of it. For the assessment of the *perceived learning* (*evaluation* level of Bloom's taxonomy), we adopted four items from earlier research addressing the cognitive aspects of perceived learning (CPL; [24]), which we slightly adopted to fit the study purpose (e.g., instead of 'The game added to my knowledge' we used 'The workshop added to my knowledge'). The four items were additionally extended with the following two: 'I learnt new skills today.' and 'I have leant something new about AI today' and were evaluated on a 5-point Likert-type scale. The internal consistency of the perceived learning dimension (including all six items) is adequate on our data set (α_{PL} = 0.910).

For the assessment of self-regulation we used the Adolescent Self-Regulatory Inventory (ASRI; [192]), which is a validated scale that measures self-regulation across 27 items and two dimensions (i.e., short-term self-regulation (ASRI-ST) and long-term self-regulation (ASRI-LT)). While in Chapter 7 we used a different metric, we decided to use the ASRI in the current study especially due to its short-term and long-term dimensions, as in our earlier study (Chapter 7) we found that these might have an influence on the study outcomes. ASRI is evaluated on a 5-step Likert-type scale. The internal consistency of the long-term dimension appeared to be adequate ($\alpha_{ASRI\text{-}LT}$ = 0.738) on our data set, however, that of the short-term dimension was below the acceptable range ($\alpha_{ASRI\text{-}ST}$ = 0.348). Since we did not aim to examine the underlying reasons, we decided to exclude the short-term dimension (ASRI-ST) from our analysis and include only the long-term dimension (ASRI-LT).

Table 10.3 Study dimensions, their operational definition, and their respective measures. All items were evaluated on a 5-point scale (1 – Totally disagree /Never; 5 – Totally agree / All the time).

Component	Operational definition	Measure/Source
Fun	The degree to which students experienced fun during the activity.	FunQ [299]
Attitude	The degree to which students indicate their attitude towards the subject.	'I think that AI is my thing.' 'Do you think that AI is fun / easy to do / easy to understand / pleasant / exciting / something I want to do again?'
Learning	The degree to which students indicate their learning during the activity.	CPL [24] and 'I have learnt new skills today.' 'I have learnt something new about AI today.'
Self-regulation	The degree to which students can self-regulate.	ASRI [192]
Socioeconomic status	An indicator for students' socioeconomic status	R-FAS [308]

Since the two participating schools in the current study were from an approximately equally wealthy neighborhood, the use of the previously applied average yearly income in the neighborhood of the school (Chapter 8) was not enough informative. Hence, to address students' socioeconomic status and to capture individual differences, we used the Revised Family Affluence Scale (R-FAS; [308]), which consist of six items and is a validated instrument for children and adolescents, allowing them to self-report on their socioeconomic background without directly investigating their parents' yearly income or educational level (which often they are not aware of). Nevertheless, previous research has shown a high correlation between students' FAS index and their parents' yearly income (e.g., [116]), hence, the instrument can be used as a reliable proxy for students' socioeconomic status. We calculated the FAS index by summing the coded responses according to earlier research [116]. Accordingly, the FAS index could range from 0 to 13. On our sample the minimum value was 6 and the maximum value was 13, with a mean of 9.70 (SD = 1.686). This means that the students in our sample, in general, were coming from relatively wealthy families, but the sample still provided sufficient variability for the planned analysis. A summary of the study dimensions and their respective measures is displayed in Table 10.3.

10.3.5 Data Analysis

The descriptive analysis and the assessment of the internal consistency was done with SPSS Statistics Software version 27.0.0. For the path analysis RStudio 1.1.453 [252] software and the lavaan [248] and psych [240] packages were used.

10.4 Results

10.4.1 Descriptive Results

For the model elements the scale scores were calculated. Accordingly, for calculating the FunQ scores, after recoding the reversed items, we summed the values. This resulted in an average FunQ score of 51.79 (*SD* = 9.477) from the possible range of 18 − 90. This average aligns with our previous studies with secondary school students (Chapter 5 and 7), and is slightly lower than the given average for the activities introduced for primary school students (Chapter 4, 6, 8 and 9). Accordingly, we conclude that students had a moderate level of fun while learning to program with AI.

For the attitude score we calculated the mean of the seven post-workshop attitude items, which resulted in an average of 2.84 (*SD* = 0.626) on a five-point scale. Comparing the average pre-workshop attitude to the average post-workshop attitude we conclude that students' attitude about programming AI has increased ($M_{pre\text{-}workshop}$ = 2.78, *SD* = 0.663), however, this increase was not significant (p = 0.285, t = -1.074, *Cohen's d* = -0.107).

Regarding students' learning, the reported average score of the six items is 2.98 (*SD* = 0.887) on a five-point scale. In other words, students self-rated their learning as moderate.

For self-regulation, we calculated the mean of the scores given for the ASRI-LT items. This resulted in an average of 3.23 (*SD* = 0.469) on a five-step scale.

For addressing students' socioeconomic status, we calculated the score of the R-FAS as described in earlier research [116]. Accordingly, for each item 0 or 1 (or 2 or 3) points – depending on the number of response options - could have been collected. The item scores were then summed, resulting in an index score, for which the possible range was 0 to 13. The average R-FAS score on our sample is 9.70 (*SD* = 1.686), which means that the study participants were coming from a relatively wealthy environment regardless the school they were attending.

The descriptive statistics, the correlation between the variables, the skewness and kurtosis values are summarized in Table 10.4. Since the univariate skewness and kurtosis values did not exceed the 2.0 and 7.0 values respectively, multivariate normality was assumed [64].

Table 10.4 The descriptive statistics of the model elements and the correlation between them (n = 122). *p < 0.05

Element	Mean	SD	Skewness	Kurtosis	1	2	3	4
1. Fun	51.79	9.477	0.051	1.595	1			
2. Attitude	2.84	0.623	-0.084	0.257	0.571*	1		
3. Learning	2.98	0.887	-0.490	-0.443	0.286*	0.315*	1	
4. Self-reg.	3.23	0.469	-0.213	0.596	0.296*	0.287*	0.050	1
5. SES	9.70	1.686	0.030	-0.756	0.106	-0.169	0.020	0.88

10.4.2 Path Analysis

To address the hypotheses simultaneously, we conducted path analysis with MLR (maximum likelihood with robust standard errors) estimation, which method is robust to non-normality and non-independence. The tested model included both the direct and the indirect effects. The analysis revealed that the experienced fun has a significant and positive direct effect on students' attitude about the topic (H1a; $p < 0.001$, $\beta_{std} = 0.585$), that students' attitude about the topic influences positively and significantly their learning (H1b; $p = 0.035$, $\beta_{std} = 0.266$), and that the indirect effect of fun on learning across attitude is also significant (H1d; $p = 0.042$, $\beta_{std} = 0.156$). We did not find a significant direct effect between fun and learning (H1c; $p = 0.372$, $\beta_{std} = 0.138$).

Regarding the self-regulation aspect, we did not find any significant association between students' self-regulation and their learning (H2; $p = 0.489$, $\beta_{std} = -0.069$), the experienced fun (H3; $p = 0.197$, $\beta_{std} = 0.186$) and their attitude about the topic (H4; $p = 0.126$, $\beta_{std} = 0.188$).

As for students' socioeconomic status, we found a negative and significant effect on students' attitude about the topic (H5; $p = 0.007$, $\beta_{std} = -0.229$), however, no significant association was found between SES and students' learning (H6; $p = 0.628$, $\beta_{std} = 0.053$) and the experienced fun (H7; $p = 0.466$, $\beta_{std} = 0.016$). The results of the path analysis, including the standardized path coefficients are depicted in Figure 10.3. The path model yielded good fit indices ([117]; *CFI* = 0.998, *RMSEA* = 0.034, *SRMR* = 0.022) except for the $\chi 2$ value (1.111, $p = 0.292$), which is however, known to be affected by the sample size, and is mostly significant when N > 75.

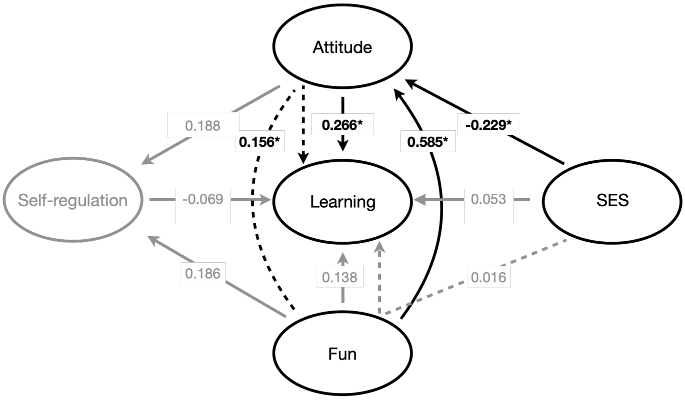


Figure 10.3 Path analysis results: standardized coefficients are added to the paths. Straight line indicates a direct relationship. Dashed line indicates an indirect relationship. * p < 0.05

10.5 Discussion

Since in the past decade it has become a worldwide pursuit to increase children's and adolescents' interest in STEM subjects in general, and in coding or programming in specific, more and more educational activities are targeting these goals. However, our knowledge is still limited about the influential factors on a) children's willingness to participate in such learning activities, b) their attitude towards the topic, and c) their learning outcomes. The FiL model introduced in Chapter 6 investigated this question and proposed that students' attitude about coding is influenced by the level of fun they experience while learning to code, and on top of that, we provided supportive evidence that having fun while learning has also a positive effect on students' learning outcomes. This study aimed to confirm the FiL model in the context of AI programming, and to extend it with two additional factors that earlier research has shown to have importance on students learning outcomes. Namely, with students' self-regulatory skills (as a personal factor) and their socioeconomic background (as an environmental factor). To test the hypothesized relationships simultaneously, we applied path analysis. Our results are in line with earlier research, however, only a part of our hypotheses was confirmed.

Regarding the FiL model, we conclude that it is applicable within the context of AI programming with secondary school students, as we found that fun has a positive effect on students' attitude about the topic (H1a), that student's attitude has a positive effect on their learning outcomes (H1b), and that fun also has an indirect effect on students' learning outcomes across their attitude about the topic (H1d). These results are perfectly aligned with that of our earlier research introduced in Chapter 6. However, we did not find a significant direct association between fun and learning (H1c), hence, the FiL model could not be extended in this direction.

Regarding the considered additional personal factor (i.e., students' self-regulation), we found no significant association between self-regulation and neither learning (H2), nor fun (H3), or attitude (H4).

Not finding a significant relationship between self-regulation and learning (H2) in our study contradicts the long-standing belief on the positive association between self-regulation and learning outcomes (or academic achievement). However, it is in line with previous research [316, 318], including our earlier findings introduced in Chapter 7, which we partially deemed to be due to the small sample size. However, in the current study, the issue of sample size can be omitted as we investigated the aforementioned relationships on a proper sample size for path analysis ($n > 100$).

Furthermore, possible issues related to the validity of the applied measures can be also excluded as in the current study we used a different (but also validated) measure for addressing students' self-regulatory skills than in the previous study discussed in Chapter 7. Moreover, the selected measure in this study [192], is especially designed for adolescents and addresses both the short-term (task-focused, e.g., 'During a dull class, I have trouble forcing myself to start paying attention'), and long-term (general and/or with focus on goals in the near future, e.g., 'I can find a way to stick with my plans and goals, even when it's tough') self-regulation. However, the internal consistency of the short-term dimension was unacceptably low on our sample, hence, we dropped this measure due to reliability issues.

As an explanation for these findings we propose that while previous studies are mostly focused on the relationship between self-regulation and learning outcomes on the long term (i.e., measured academic achievement mostly with course grades), in both the current study and our earlier study (see Chapter 7) the focus was on the short term learning (i.e., learning from a 90 min/120 min workshop), which might have influence the study outcomes. Therefore, we call on future research to better understand the relationship between short- and long-term self-regulation and short- and long-term learning (i.e., class level vs course level) as to our best knowledge, no previous studies exist that would focus on this issue.

Another possible explanation for these finding lies in the context of the study, namely, our study and the study of Villavicencio et al. [318] and Vestege et al. [316] with aligning results were all within the context of STEM learning (i.e., programming, cognitive learning, trigonometry and enzymology), while other, contradicting research findings were obtained from studies, which took place outside of STEM subjects (i.e., language learning [8], or mixed study disciplines [345]). We see this as a substantial difference, which can provide the explanation for our findings. In accordance, we propose, that the positive association between self-regulation and learning appears not to be valid within the context of STEM subjects. The underlying reasons, we argue, are related to the cognitive strategies students utilize while learning. It seems feasible that different strategies are required to learn STEM subjects than to, for example, learn language, and those strategies can have a different impact on self-regulation, which at the end results in different self-regulatory skills required to obtain good learning outcomes. Since it is well reflected in the aforementioned that the positive association between self-regulation and learning outcomes cannot be taken granted as it might vary across different study disciplines, we call on future research for a deeper understanding of this topic.

Regarding our results about self-regulation and fun, we find it important to mention that the expected positive association was based on earlier research [15, 215], which, however, pertained to the examination of enjoyment rather than fun, and did not examine the relationship between these two and students' learning outcomes. In our study, despite that we found a positive and significant correlation between fun and self-regulation (just as earlier research [15, 215] did), when looking at the model-level and testing multiple associations simultaneously, the association between self-regulation and fun were found non-significant. This is a normal phenomenon when complex models are analyzed and is due to partial redundancy among the predictor variables [56]. Thus, we conclude that in relation to learning, we did not find a significant association between fun and self-regulation, and hence, the FiL model could not be extended with self-regulation.

As for the considered environmental factor (i.e., students' socioeconomic background) our study results partially confirm our hypotheses. Namely, we found that students' SES influences their attitude about the topic significantly (H5), but we did not find a significant direct association between SES and students' learning outcomes (H6) or a significant indirect association between SES and learning across the level of fun students experienced while learning (H7). These findings broaden our understanding on how students' socioeconomic background can influence their learning outcomes, by investigating this relationship in the context of programming AI with secondary-school students. While numerous studies exists on the relationship between SES and learning outcomes in general (see e.g., the meta-review of [166, 191, 244, 276, 326], our knowledge is still limited in the specific field of STEM learning. Adding to this body of knowledge is crucial, especially in reflection of the worldwide pursuit to increase student's interest in STEM subject from early ages on. While we did not have a specific expectation whether the association between SES and attitude towards programming AI would be positive or negative, the negative association between these two might come as a surprise. What we found, in other words is that students from lower socioeconomic background reported on higher attitude scores in comparison with students from higher socioeconomic background. These results align with that of Yerdelen, Kahraman, and Tas [334], who found positive attitudes towards pursuing a STEM carrier among low-SES students. While we propose that this finding might be context-dependent (e.g., country, region, culture), as possible explanation we suggest that students from higher socioeconomic background are better exposed to STEM and/or programming-related activities, thus, they have a better-informed picture about what it involves. Therefore, their scores reflect a better-informed decision, meanwhile those from a less-wealthy environment report rather on their anticipation, without the same amount of personal experience. As a side note, however, we need to say that participating students in our study were all coming from a relatively wealthy environment, hence, when putting these results outside of the context of the study, we shall regard them as middle-SES and high-SES students.

As for the relationship between SES and learning, recent meta-reviews on the topic in general found only weak to moderate correlations between the two [244, 276], noting that academic achievement is more strongly related to other factors such as prior academic achievement or working status than to SES [244], and that the unit of measurement for SES has an influence on the strength of the observed correlations [326]. In our study we measured SES at the individual (student) level, which according to White [326] could have had consequences for the strength of the relationship with students' learning, as they propose that the relationship between these two are stronger at an aggregate level. Nevertheless, these findings add to the body of literature by focusing on the relationship between SES and learning within the specific field of programming. Since previous research is limited in this specific field but is very much needed as there is a worldwide pursuit to increase students' interest in programming, we call on future research for a better understanding of the underlying reasons for students' participation in programming-related activities – for which their SES is proposed to be an important factor.

In reflection of our results, we conclude that the FiL model can be extended with the considered environmental factor: students' socioeconomic background, but not with the examined personal factor (i.e., students' self-regulation).

10.6 Limitations and Future Work

To start with, despite our intentions, we could not address the relationship between fun, attitude, learning, SES and students' short-term self-regulation as the internal consistency of the applied measure was under the acceptable range on our sample. This issue, we propose, is sample related as the scale has been validated before. Since our understanding is limited on how short-term and long-term self-regulation affects short-term (i.e., class level) and long-term (i.e., course level) learning, what are the differences and similarities, we call on future research to investigate these questions.

Similarly, we intended to measure two levels of learning (i.e., measured and reported learning), and compare the relationship of those to the other study dimensions, but we could not address students' measured learning as students progressed slower with the workshop material than anticipated, and therefore, the knowledge test was not enough nuanced to capture students' learning gain. Therefore, future studies should address how the different levels of learning (e.g., measured learning, perceived learning, task-based performance) relate to students' attitude about the topic, self-regulation, SES, and their level of fun experience while learning.

Additionally, in the herein introduced study we assessed students' socioeconomic status by using the revised Family Affluence Scale (R-FAS, [308]. Despite being a validated scale for the assessment of students' SES, we could not verify reported values with data from parents due to anonymity and resource issues. Therefore, to complement our study findings, future research could consider assessing students' SES at various levels (e.g., individual level, family level, school level, neighborhood level) and could collect relevant data from multiple sources (e.g., self-report from children, data from parents, relevant neighborhood statistics etc.) as White [326] indicated that the level of association between academic achievement and SES is strongly dependent on the unit of measurement for SES.

Further, a contextual limitation of our study is that in the current study no low-SES students were present. This is because we recruited schools from socioeconomically similar neighborhoods (average yearly income of the neighborhood of the two schools are €29.900 and €28.400 respectively, compared with the €20.200 of the low-SES school in Chapter 8) and did not use the neighborhood as an indicator for students' SES but we investigated their SES by the use of the R-FAS, in comparison with our earlier study introduced in Chapter 8, where the students were recruited from schools which were located at three socioeconomically distinct neighborhoods. Therefore, to extend the scope and the generalizability of these results, we propose the study replication with a wider range of participants regarding their socioeconomic background and in less wealthy countries than the Netherlands. Additionally, since our results might be context-dependent, we call on future research not only in different countries, but with different age groups, cultures, schooling systems and ethnicity to capture a nuanced picture on what and how influences students' interest and learning in programming-related activities.

10.7 Conclusion

This study aimed to understand how students' attitude about programming, their self-regulation, their socioeconomic background, and the level of fun they experience while learning relates to their learning outcomes. Based on our study findings we conclude that having fun while learning to program contributes positively not only to students' attitude towards the topic, but also to their learning outcomes, which finding supports endeavors of educators who aim to pass their study material to their students in a playful, enjoyable, and engaging way. In line with these we conclude that adopting a playful approach to introduce STEM subjects to children and adolescents can contribute to the worldwide pursuit of getting more children interested in scientific topics. Further, our study contributes to the literature by providing insight into how students' socioeconomic background might influence their interest and participation in programming-related activities, and in accordance with that, we recommend that these sorts of activities are better targeted at, and more successful with middle (and/or low) income students than with high-income students. These findings provide clues for researchers, educators, and practitioners when designing and implementing STEM-related activities for children on how to create meaningful learning activities that are not only enjoyable, but also serve the purpose of being educative for the participants.

.

Part VII.

CONCLUSION

11 Discussion and Conclusion

11.1 Introduction

In this final part and chapter, we provide an overview of the insights from this thesis. We start with answering the research questions, followed by the summary of the research contributions, and finishing with the limitations of the introduced work and a set of directions for future studies. This thesis presents findings about how the experienced fun while learning influences learning outcomes in various learning settings, with focus on the field of STEM education. The findings are based on empirical evidence from questionnaires and multimodal data studies conducted between 2018 and 2022 with children and adolescents between age 8 and 16.

The introduced studies mostly took place in the non-formal learning environment (as defined in Chapter 1, see Table 1.1). Accordingly, the related activities were structured but were not part of the school curriculum, there were no pre-determined learning goals by the teacher and the researcher as the main aim of these activities was to introduce the topic to students in a playful and engaging way, however, the researcher, and in some cases also the teacher, were present and were available to guide the students if they needed support. Furthermore, when designing the activities we built on students' intrinsic motivation, as participation in the activities (i.e., staying in the classroom) was often compulsory because they took place during school hours, but *active* participation could only be reached if the designed activity evoked students' curiosity to follow the activity. One study took place in the formal learning setting. In contrast to the aforementioned studies, in that case there were predetermined learning goals by the teacher and the researcher, which aligned with the school curriculum, and therefore, participation was compulsory. A summary of the studies conducted during this research work is displayed in Table 11.1 below.

Table 11.1 Summary of studies conducted during this research work.

Chapter/ study	Type of learning activity and topic	Place	Setting	Age
3/ Initial item pool study	Interactive exhibition (science)	Out of school (Outdoor museum/ theme park)	Non-formal	11-13
3/ Think-aloud study	Playful robotics workshop	At school	Non-formal	11
3/ Validation study	Interactive exhibition (science)	Out of school (Science Museum)	Non-formal	10-15
4	Playful coding workshop (micro:bit)	At school	Non-formal	10-12
5	DGBL (biology)	Online	Formal	13-14
6	Playful coding workshop (micro:bit)	At school	Non-formal	9-12
7	DGBL (cognitive game)	At school	Non-formal	16-17
8	Playful coding workshop (micro:bit)	At school	Non-formal	8-12
9	Playful coding workshop (micro:bit)	At school	Non-formal	8-12
10	Playful coding workshop (AI)	At school	Non-formal	12-13

11.2 Answers to Research Questions

In Chapter 1 we raised four main research questions. Here, we answer them based on the findings of our studies introduced in this thesis. To be able to answer the driving question of this thesis (RQ1), first, we start with answering SQ1 and SQ2.

11.2.1 SQ1: What is fun?

In Chapter 2, we reviewed previous work related to the concept of fun and its possible assets. We found that despite the importance of fun in learning activities is getting widely acknowledged, there is no conceptual clarity and theoretical framework regarding the definition of fun; and that the term fun is often used interchangeably with synonyms, such as enjoyment. Based on the reviewed theories we established our initial definition for fun, which we empirically validated in the following step (described in Chapter 3) by applying a deductive scale development approach. Accordingly, in this work we defined fun as follows:

Fun is a positive emotional experience during which the level of challenge meets the level of skills, one feels in control, loses the perception of time and space, lets go of social inhibitions, experiences low levels of stress, and is intrinsically motivated for the participation in the experience.

While our conception of fun is focused on learning, we propose that it can be possibly extended to other domains (see further discussion on this matter in section 11.4 below).

The literature review allowed us to map the concept of fun with neighboring terms. Based on the review we concluded that while the most frequently used synonym enjoyment is a discreet emotion that is often used in relation to learning, and described by Pekrun et al. [215] as an 'academic emotion' (i.e., emotion that arise in different academic settings and is directly linked to learning, instruction and/or achievement), fun is a multidimensional concept, which describes a state and a complex emotional experience. In this sense, fun is more closely related to the concept of flow [61]. Despite having overlapping core elements between the two, we can still make a clear distinction, as flow (among others) has the characteristic of having a clear goal and immediate feedback, while fun is not associated with explicit goals, but with high levels of positive emotions, and low levels of stress. Another neighboring term discussed is game experience, which refers to one's subjective experiences associated with game play [225], therefore, it is context dependent while fun is not. Finally, we discussed the notion of play and concluded, that while in early childhood fun and play often concur, as children approach being a teenager the hedonistic character of fun that is purely present during play gives way to challenge, which is a well-established dimension of fun at later ages. We propose that researchers endorsing fun as an important element in learning environments could consider a) whether our definition of fun lends welcome precision to their discourse, and b) how the notion of fun they advocate or observe, relate to the definition provided in this thesis. Similarly, those relying on related but different concepts could find it useful to clearly articulate how their conception relate to our definition of fun.

11.2.2 SQ2: How can we measure fun?

While some previous attempts exist for the measurement of fun (mostly in the field of human-computer interaction), they rather measure product liking (acceptance or preference) with young, preliterate children, or focus on adults and game engagement or the gaming experience. We have thoroughly discussed the caveats of these instruments in Chapter 3 and Chapter 6. The main issue, is however, the lack of a commonly accepted theoretical framework for the assessment of fun. Since without knowing what we measure, measurement is complicated, it was crucial to first address SQ1 to be able to answer SQ2. After having a clear definition of fun, by applying a deductive scale development approach a questionnaire named FunQ was designed and validated in four consecutive steps. Chapter 3 presented the design and validation of FunQ. In Chapters 4-10 we introduced several studies that investigated the relationship between fun and learning from different angles, and where we consequently used FunQ for assessing the experienced fun while learning. In all introduced studies FunQ showed a high internal consistency, and when examined (e.g., Chapter 7) a high convergent validity with related measures, and significant associations with physiological markers (Chapter 9), hence FunQ is put forward as a highly reliable assessment tool for fun. Based on this work we conclude that we can measure the experienced fun while learning with FunQ, that FunQ is a reliable and much needed addition to the current palette of tools, which is specially designed for, and validated with adolescents. Additionally, we have shown how FunQ can be used in various learning settings (online-, offline-, formal-, non-formal, and gamified learning), with various STEM subjects and with a wide age range (8-16).

11.2.3 RQ1: What is the role of fun in learning?

Although all chapters introduced in this thesis ascertained the relationship between fun and learning, and hence, contributed to answering this overarching research question, there are two chapters that focuses exclusively on this question.

Chapter 6 introduced the initial FiL model. The FiL model describes how the experienced fun while learning has a positive influence on students' learning outcomes across its positive effect on their attitude about the topic (see Figure 6.5). Accordingly, Chapter 6 provides quantifiable evidence that fun can enhance learning within the context of STEM education. These findings have a special importance in relation to STEM education, as it has become a recent worldwide pursuit to increase children's interest toward scientific topics, however, empirical evidence was hard to come by to support the underlying assumption that making learning fun has a positive effect on the learning outcomes. According to our findings, designing enjoyable STEM learning activities not only contribute to children's increased attitude about the topic, but it also improves their learning outcomes, thereby fun has a crucial role on making children more interested in STEM subjects and in related learning activities.

In Chapter 7 and 8 we investigated whether the role of fun described by the FiL model is influenced by personal factors such as students' self-regulation, and by environmental factors such as students' socioeconomic background. The initial results indicated that the effect of fun in learning is independent of students' self-regulation, but their socioeconomic background might have an influence on it. In Chapter 10 we introduced a final study, which aimed to establish an extended model on the role of fun in learning, taking simultaneously into account the main study dimensions (i.e., fun, attitude, learning, self-regulation, and socioeconomic background) investigated during this dissertation research. Based on the results from this final study we conclude that the FiL model can be extended with students' socioeconomic background, but not with their self-regulation. Therefore, answering RQ1 we conclude that fun plays a vital role on learning, as it positively influences students' learning outcomes across its positive effect on their topic-related attitude. Further, we found that students' socioeconomic background has an influence on their attitude about the topic, but the socioeconomic background was neither directly, nor indirectly related to students' experienced fun while learning or their learning outcomes. Finally, in our studies we did not find a link between self-regulation and the other investigated dimensions (i.e., fun, attitude, learning and SES), which means that the herein discussed relationships are found to be independent of students' self-regulatory skills. The final model on the role of fun in learning within the scope of studies in this thesis is depicted on Figure 11.1. We propose this model as a working model for further use.

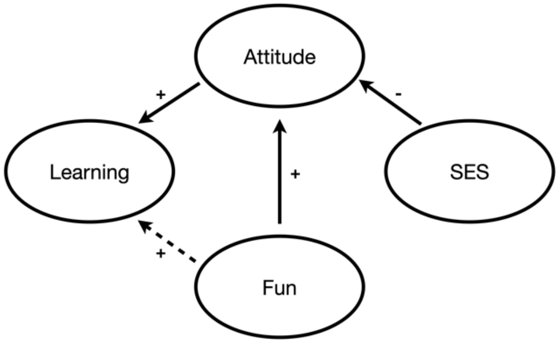


Figure 11.1 The role of fun in learning: the revealed relationships in this thesis between fun, attitude, learning, and socioeconomic background. Straight line indicates a direct relationship. Dashed line indicates an indirect relationship.

11.2.4 RQ2: Is the relationship between fun and learning affected by personal factors such as self-regulation, and environmental factors such as socioeconomic status?

The studies introduced in Chapter 7, 8, and 10 aimed to give an answer for this research question.

Based on the long-standing belief on the positive association between self-regulation and academic success we expected that students' self-regulation will have a link to the FiL model. However, when we investigated this relationship by means of path modeling (Chapter 7) we did not find a significant association. In other words, based on our study findings, students' self-regulatory skills have no direct or indirect influence on the level of fun they experience while learning, nor on their learning outcomes. This latter finding challenges the long-standing general belief that a positive relationship exists between self-regulation and learning, however, it also extends the current body of knowledge by nuancing the aforementioned relationship. While earlier research generally investigated the relationship between self-regulation and long-term learning (i.e., course level; [16]), our study focused on the short-term learning (i.e., class level), and investigated learning from a different angle. Namely, according to Bloom's taxonomy of learning, we investigated learning at the *Evaluation* level (i.e., judgements about the value of the material for a specific purpose); and *Application* level (i.e., use of abstractions in concrete situations), instead of the generally used *Knowledge* level (i.e., recall of specific material; measured by the course grade). Additionally, we see that recent studies found similar outcomes (no linear relationship between learning and self-regulation) when the topic was science-related [316, 318], while it is not true for non-scientific topics (e.g., language learning [8]). The study introduced in Chapter 10 further supported our earlier findings, as we did not find any direct or indirect link between students' self-regulation and the level of fun they experienced while learning, their topic-related attitude and their learning outcomes. In accordance, we conclude that based on our study findings, the relationship between fun and learning described by the FiL model is independent from students' self-regulatory skills.

Based on the theory of science capital [13], which theory is also positively associated with one's SES, we expected that students from high income families would hold more positive attitudes towards science and technology and would perform better in programming than students from lower income families based on their generally higher exposure to experiences involving computing technology, and hence, students from different socioeconomic background would benefit differently from the applied playful coding activity (Chapter 8). Indeed, our findings support the idea of students benefitting differently from the activity, but not exactly as it was expected. Namely, we found that students from the middle- and low-income school profited the most from the applied playful coding activity. Children's attitude about coding (which is related to their learning outcomes) changed positively only in the middle- and low-income school, and the results regarding the learning outcomes also suggested that students from the high-income school learnt the least. This finding has special importance for teaching students programming and computational skills, as it clearly indicates that tailoring the methods and study material to the audience (i.e., their socioeconomic background) is required to successfully implement the learning activities. The extended model introduced in Chapter 10 and the related study further supports these findings as we found a negative association between students' SES and their attitude about the topic. In other words, students from high socioeconomic background had lower attitudes about programming than students from lower socioeconomic background. An important sidenote is, however, that the students in our sample were all coming from a relatively wealthy environment, and hence, when putting these results outside of the study context, we shall consider them as middle- and high-SES students. Nevertheless, these findings also indicate that students from different socioeconomic background profit differently from the same playful (STEM) learning activity, and hence, adjusting the teaching methods and study material to the participating students (i.e., their socioeconomic background) is essential for delivering meaningful and successful (STEM) learning activities.

11.3 Summary of Contributions

The work presented in this thesis contributes to the intersection of child-computer interaction and learning sciences fields, particularly to the subfield focusing on STEM education as there is a worldwide pursuit to increase children's and adolescents' interest in scientific topics and related activities. This interest is especially emphasized in relation to computer science, as computational thinking and coding are frequently seen as the literacy skills of the 21st century [210]. The two main, non-traditional approaches to increase students' interest in scientific topics are gamification and non-formal learning. Both these approaches aim to make the learning experience fun. Despite this wide interest, there has been little systemic effort to provide quantifiable evidence regarding the role of fun in learning not least due to a lack of a common theoretical framework for defining the concept of fun and for supporting its measurement. Therefore, in this research work we conducted several studies in different learning settings to thoroughly examine whether and how fun influences students' learning, and to finally, quantify the relationship between the experienced fun while learning and students' learning outcomes. To summarize these findings, this thesis makes the following six main contributions.

11.3.1 It provides a theoretically grounded and empirically validated definition for fun.

While designing and implementing fun learning activities, especially in STEM education, is gaining its momentum, not having a commonly accepted definition for the notion of fun has been a striking issue that emerged from reviewing related literature. In Chapter 2 we thoroughly examined previous attempts for describing fun and scrutinized the differences and similarities with neighboring terms such as the academic emotion enjoyment [215], flow [61], game experience [225], and play [121]. Based on previous literature we concluded that fun is a multidimensional construct, and accordingly, we created the initial definition, which was refined during the empirical validation (Chapter 3). Therefore, this thesis extends the previous body of knowledge with a much-needed, theoretically grounded definition for fun.

The provided definition can serve as a common theoretical framework for researchers in the fields of child-computer interaction and learning sciences, and thereby support the investigation and a better understanding of the concept of fun, ultimately leading to a more adequate implementation of fun aspects into learning activities.

11.3.2 It provides a much needed, validated measurement tool for the assessment of fun, which fills a gap in the current palette of tools.

While some measurement tools existed previously, mostly within the field of human-computer interaction, they either measure product liking (acceptance of preference) with young, preliterate children (e.g., fun Toolkit [234]; fun Semantic Differential Scales [338]; This or That [340]), or address the gaming experience (e.g., EGameFlow [91]; UES [204]; GEQ [226]). Remarkably, there has been a gap in research with adolescents (age 11-18), and no tool existed previously that would have addressed the experienced fun while learning. FunQ is also different from earlier measurement tools in the sense that a) it focuses on the personal experience and not on the evaluation of a product or activity, b) it builds on a theoretically grounded definition of fun, making a clear distinction from the neighboring concepts (i.e., enjoyment, play, game experience, flow etc.) that in previous literature have been often used interchangeably with fun, and c) considers fun as a multidimensional construct. Additionally, during the design of FunQ special attention has been devoted to possible methodological issues when creating a self-report measure for children and adolescents (e.g., attention span, response bias, readability, negation, item length and phrasing, labelling of the scale steps etc.), and accordingly, we followed the recommendations of earlier scholars to optimize the questionnaire from both the methodological point of view (e.g., [163, 187]) and from the design perspective (e.g., [29]) to its target audience. In Chapter 3 and Chapter 6 we have thoroughly compared FunQ with existing measurement tools to address similarities, differences, and suitability for the purpose of addressing children's and adolescents' fun experience while learning. Throughout this thesis, FunQ has been proven to be a reliable measurement tool for the assessment of the experienced fun in various learning contexts (i.e., formal-, non-formal, online-, offline-, and gamified learning), with various STEM subjects (e.g., programming, biology), and with a wide age range (8-16 years).

While previous research indicated that non-formal science learning activities often aim to implement playful and fun elements [303], due to the lack of reliable measurement tools the assessment of the 'funness' of such activities was left for personal impressions, or had to be assumed. FunQ provides educators and researchers an easy and reliable way to address whether their activities reached the desired goal of making learning fun. This feedback then can serve the purpose of increasing the quality of the learning activity at various educational settings (e.g., museum visits, DGBL, non-formal science learning activities etc.), by providing a nuanced picture across the six dimensions of FunQ.

11.3.3 It introduces the original FiL model that describes the relationship between fun, attitude, and learning.

Despite the growing efforts amongst educators and researchers in gamification and non-formal learning spaces to ensure that learning environments are fun and enjoyable, the role of fun in learning has not yet been well understood, and quantifiable evidence was hard to come by. Certainly, one of the main underlying issues has been the lack of commonly accepted theoretical framework for defining the concept of fun and for supporting its measurement. Gamification builds on the idea that learning games resemble free-time activities, such as playing video games, and hence, they are fun and intrinsically motivating [205]. However, Yee [333] argued that learning games require players to do many tasks, making students feel tired and tedious. In case of non-formal learning, the activities often aim to implement playful elements [303] to make the learning experience fun and enjoyable, however, there existed no systematic research to assess whether besides students enjoying themselves during these activities learn something or not.

In this thesis, based on initial case studies we have proposed the FiL model that describes the relationship between fun and learning, and provided empirical evidence that fun has a positive influence on learning across having a positive effect on students' attitude on the topic they are learning about (Chapter 6). Therefore, our results support efforts of educational researchers and practitioners who try to make learning activities more fun for students in the belief that making (STEM) education fun enhances students' learning. On top of this, the FiL model also provides supportive evidence that making learning fun improves students' attitude about the topic. Therefore, for those who aim for attitude change about a certain (scientific) topic with their educational activities we suggest implementing fun elements into the learning activities. Additionally, researchers can also refer to the FiL model when designing future studies where the aim is to improve participants knowledge about a certain topic.

11.3.4 It nuances the effect of fun in learning by considering personal and environmental factors (i.e., self-regulation and socioeconomic status).

Researchers have argued that self-regulation is an essential component of academic success [221], and there is also substantial effort invested towards enhancing fun in learning to improve student engagement and learning outcomes. However, little has been known about the relationship between self-regulated learning, emotions and learning outcomes. Where all these aspects were simultaneously examined, previous research appeared to be inconclusive [8, 318]. To fill this research gap and to potentially extend the FiL model, in Chapter 7 and 10 we introduced two studies that scrutinized the relationship between self-regulation, fun, attitude and learning. While our research results indicated no significant relationship between self-regulation and the other investigated dimensions, and hence, the FiL model could not have been extended, the results aligned with earlier research findings [318], and extended those by looking outside of the scope of academic emotions and focusing on fun, while also zooming into the setting of a single learning activity in comparison with the generally studied course-level effects.

Regarding the effect of socioeconomic status (SES), the theoretical perspective of science capital [13] suggests that students from high income families would hold more positive attitudes towards science and technology and would perform better in programming than students from lower income families based on their generally higher exposure to experiences involving computing technology. To examine this assumption, and to potentially extend the FiL model, in Chapter 8 we introduced a study that compared low-, middle-, and high-SES students along their attitude about programming, the fun they have experienced while learning to program, and their learning outcomes. Our findings indicated that students' socioeconomic background influenced how they profited from such learning experiences, middle- and low-income students benefiting the most in terms of attitude change and learning outcomes. Additionally, the effect of SES was further investigated in Chapter 10, in which we have extended the FiL model with one's socioeconomic background, as an influential factor on one's attitude. These findings complement previous literature, as the effect of socioeconomic background has been understudied within the field of non-curricular STEM activities in general, and in programming activities in specific. Knowing more about what influences students' interest and willingness to participate in such activities can contribute to the worldwide pursuit of attracting more children and adolescents into programming-related fields while studying, and jobs later in their lives. In reflection of our findings, therefore, we propose that similar activities are best designed for, and implemented with students from middle and low socioeconomic background as they appeared to benefit the most in terms of attitude change and learning outcomes in comparison with students from high socioeconomic background.

The findings of these three chapters, on one hand, provide educators cues on tailoring their activities to different audiences, and on the other hand, provide possible explanations for the underlying reasons why learning activities are not equally well received and not equally positively perceived across students with varying backgrounds and cognitive skills. Additionally, while both the original and the extended FiL model is the first of its kind, they can serve as a starting point and an inspiration for researchers and educators for orchestrating meaningful fun learning activities for students that not only serve as entertainment, but also beneficial for achieving the learning goals.

11.3.5 It links fun to physiological markers, allowing for a momentarily investigation of the effect of fun in learning.

In the field of interaction design and children, evaluation of fun has been largely focused on self-reported data from children and adolescents, asking them to assess specific activities in single-item scales or to compare the experienced fun in relation to different elements of the design [234]. Despite being a wiely used approach, it suffers from being retrospective and coarse grained. To further contemporary approaches and understand students' affective processes while learning comprehensively, in Chapter 9 we introduced a study that adopted a multimodal data analysis approach. In particular, we used both objective automated measures coming from students' physiological response data (collected by wristbands and facial video recordings) and self-reports (i.e., surveys). By using multimodal data, we went a step further than earlier research as it either pertained to surveys or to physiological response data only. Using the combination of the two allowed us a deeper understanding on how fun occurs during learning, and which physio-affective states can be used as a predictor of fun.

Being able to predict fun from physiological signals can help assess different learning activities, but potentially can be developed further to support timely interventions to get disengaged students on track again, ultimately leading to better learning outcomes. Developing new tools or further improving existing ones that address the potential of using unobstructive physiological response data could support both teachers and students in their everyday life, by providing systems with affordances for reflective purposes, indicating students' disengagement or other features to support better classroom management.

11.4 Limitations and Future Work

While our studies contribute by revealing insights regarding the research questions we put forward, they also have certain limitations. Although these limitations are constraining on one hand, on the other hand, they highlight emerging opportunities for further research. We start this discussion with limitations concerning the thesis contributions, then we move on to more generic considerations.

In this work we provided a theoretically grounded and empirically validated definition for fun. While we propose that due to the nature of how the definition was established it is generally valid, we only tested it within educational contexts in general, and within STEM education in specific. Therefore, future research could investigate the validity of the definition in contexts outside of education. The same argument is true for the assessment tool, FunQ.

Additionally, the FiL model that describes the relationship between fun, attitude and learning has been also established based on data from mainly STEM learning and mostly in the non-formal context. We propose that due to the general nature of the model it is valid in other learning contexts as well (i.e., outside of STEM subjects and in formal- and informal setting, too), however, it is left for future researchers to verify this assumption. The same argument is valid for the age group.

In this thesis we also investigated whether fun can be linked to physiological markers, however, we used only a set of Control-Value Theory emotions and their related facial expressions to investigate this question. Since it is a novel perspective, the applied measures are not validated with students yet, hence, estimating the accuracy is difficult. Nevertheless, given that we were not looking at absolute numbers but at variations, we believe that this limitation did not hinder our study results. Still, future research could focus on the validation of these measures with children and teenagers. Furthermore, future studies could broaden our understanding and further the state of art by studying education specific emotions - such as boredom, frustration, confusion, and delight -, and related facial expressions to investigate whether the herein discussed associations can be transmitted to those other emotions, potentially contributing to computer science education. At last, future studies could use different sort of physiological sensors, such as eye-tracking or EEG caps to complement our findings as they provide a richer data than wristbands, but these devices are not dependent on computer vision like facial data, hence, using them will exclude vision-specific errors. Additionally, EEG caps and eye-tracking provide the opportunity to investigate cognitive processes beyond emotions (i.e., attention, cognitive load, mental effort etc), and eye-tracking data, specifically, also takes into account the context, therefore, it can help with removing contextual biases.

To nuance the effect of fun in learning, we considered one personal (i.e., self-regulation), and one environmental (i.e., socioeconomic background) factor in this work. While the effect of fun on learning described by the FiL model appeared to be independent from students' self-regulatory skills, future research could test whether other personal factors (e.g., personality traits) have an influence on the relationship between fun and learning. Additionally, while we found that students' attitude is influenced by their socioeconomic background (but their level of fun they experienced and their learning outcomes not), other environmental factors (e.g., location, (sub)culture) could be scrutinized as well.

Concerning the general limitations, first, and foremost, the herein introduced studies were mostly conducted in relation to STEM education, with a strong focus on learning to program. While it is an important topic that attracts a lot of attention nowadays, more research is required to assess whether the herein revealed associations can be generalized to different learning topics, and possibly to different age groups. Although we have no reason to assume the herein revealed associations to be different with adults, it is certainly interesting to investigate in the future whether the role of fun in learning is similar with younger children (i.e., below age 10), especially in reflection that at younger ages fun and play often concur.

Second, most of the activities were designed as a non-formal learning activity, however, they took place in the formal learning setting. While this specific setting could have affected the transferability of our results, it provided us the opportunity to investigate the average student, who not necessarily has a positive attitude about the topic or comes from a wealthy family (i.e., based on the theory of science capital [13], one can expect that students from lower socioeconomic background would be underrepresented at non-formal science learning activities due to e.g., low interest or financial obstacles). We propose that fun has a positive influence on students' attitude about the topic and their learning regardless the place or setting of the learning, but it is possible that the scale of these effects differs across the various learning contexts. Therefore, future research should investigate these differences to gain a deeper understanding on the role of fun in learning across various learning contexts and settings.

Related to the aforementioned, given that non-formal learning is less structured than formal learning, despite the nature of the tasks required individual work, collaboration was allowed, and hence, it took place to a certain extent. Accordingly, peer learning could have occurred as well. The investigation of collaboration and peer learning was outside of the scope of this research, nevertheless, future studies focusing on this question could investigate whether the FiL model could be extended in this direction.

Next, there can be cultural differences in the meaning, concept, and experience of fun. All studies reported in this thesis have been conducted within Western Europe. Therefore, further research should examine possible intercultural differences, which would lead to our increased understanding not only about the notion of fun, but also about its role on students' learning outcomes in different cultures.

Furthermore, concerns can be raised about the validity of survey data collected from students. While this is certainly a valid argument, we paid special attention to overcome possible challenges and prevent response bias. We wrote in detail about surveying children and adolescents in Chapter 3, and we followed to those considerations throughout this research work, ensuring high quality response data. Nevertheless, in the future, extending the herein described quantitative findings with qualitative methods could contribute to a deeper understanding of the underlying reasons for the described associations.

While this thesis examined learning according to different levels and aspects of learning (i.e., perceived learning, measured learning, relative learning gain, and task-based performance), we need to recognize that the assessment of learning is complicated, and hence, our study results are contingent upon the validity of the used measures. As for future directions of research, we call on future research focusing on the eventual differences and similarities between the different ways of learning assessment, not only to test further the validity of the introduced models, but also, to provide the learning sciences community with a deeper insight about the differences and similarities of various ways of learning assessment, including its potential effects on capturing *actual* learning.

Similarly to learning, assessing children's and adolescents' socioeconomic status can be a difficulty. To overcome this issue, we applied different approaches, such as using the average yearly income in the neighborhood of the school as a proxy for the attendings students' SES or we recorded the widely used Family Affluence Scale [308]. Nevertheless, both methods provided us with a proxy for students' SES. Another approach could have been to asks the parents of the children directly, however, it would have hindered the anonymity of the research, and hence, this option was abandoned.

Additionally, students' experience with different pedagogical qualities they are familiar with through their school (e.g., playful learning, game-based learning etc.) might have influenced our outcomes. While all studies were conducted in schools without any specific pedagogy, investigating whether the type of school students attend (e.g., traditional, Montessori, Waldorf) has an influence on the herein revealed relationships would add to our understanding on both the personal and the environmental influential factors on the effect of fun in learning.

At last, investigating non-linear relationships between the model components, including bi-directional effects in the future would contribute to a better understanding of the role that fun plays on learning.

11.5 The Future of Fun in Learning

Due to the rapid technological developments of the past decades, education (i.e., both teaching and learning) is currently under a previously has never seen pressure to adapt to the changing needs of the society. The consequences of the COVID-19 pandemic (i.e., online classes and distance learning) further accentuated this need. However, getting and keeping children and adolescents interested and motivated in learning (regardless the topic), especially if the learning takes place online can be difficult. We have shown that integrating fun elements into the educational materials can improve both students' attitude about the topic and their learning outcomes. Building on this idea to support the required changes in education, we call on future research into understanding what universal elements (e.g., learning setting, type of material used, teaching pedagogy applied) can make a learning activity fun. Having a thorough understanding on this question could contribute fundamentally to both the field of learning sciences and the gamification industry. However, we will also need further research, similar to what has been introduced in this thesis, which will provide theoretical grounding for those elements.

Additionally, a possible direction of future research is related to Multimodal Learning Analytics. This upcoming field of research aims to capture and access detailed data obtained from a combination of emerging technologies (e.g., advanced sensing and artificial intelligence) and classic data collection methods (e.g., surveys and interviews). In this thesis we have shown that it is possible to link both learning and fun to certain physiological markers. Extending our knowledge and linking fun and learning to further markers from various (everyday) sensors (e.g., eye tracking or smart watches) would enable a new pedagogy, which could closely follow the technological advances of the 21st century, and where interventions and learning approaches could be monitored and motivated in real time, leading to the maximization of each students' potential. While personalization has, among others, the aforementioned benefit, it has its downsides as well. Therefore, future research should also investigate how personalization can and will affect society at all levels, with specific focus on education. Questions such as whether personalization is possible without the further datafication of the individual (i.e., data collection on all levels of education, and about all processes of learning and teaching [132]), and without keeping one in their comfort zone should be addressed. Consequently, researchers in this field should focus on how personalization could promote open mindedness, and prompt students to move out of their comfort zone, by, for example, implementing fun element to make activities inviting. Since experiencing fun has a strong social aspect, we propose that having fun (and thus, creating fun learning activities) might become a key element that will help students connect with each other in the future society of personalized learning.

While it sounds promising that in the future we could be able to personalize and optimize the learning experience to a high-level of individual needs, it also carries potential ethical issues. Therefore, our final point for future research is the investigation of ethical considerations. For example, future research could address whether a high-level personalization of a learning activity would be ethical (Would not students become 'slaves' to AI that wants them to learn?), and what would be an ethically acceptable way of personalization. Related to this, applying various sensors to capture real-time data to support personalization also invites a debate on data protection issues, which should not be neglected by future work.

BIBLIOGRAPHY

[1] Abbasi, A.Z., Ting, D.H., Hlavacs, H., Costa, L.V. and Veloso, A.I. 2019. An empirical validation of consumer video game engagement: A playful-consumption experience approach. *Entertainment Computing*. 29, December 2018 (2019), 43–55. DOI:https://doi.org/10.1016/j.entcom.2018.12.002.

[2] Abeele, V. Vanden, Spiel, K., Nacke, L., Johnson, D. and Gerling, K. 2020. Development and validation of the player experience inventory: A scale to measure player experiences at the level of functional and psychosocial consequences. *International Journal of Human Computer Studies*. 135, January 2019 (2020), 102370. DOI:https://doi.org/10.1016/j.ijhcs.2019.102370.

[3] Adèr, H.J., Mellenbergh, G.J. and Hand, D.J. 2008. *Advising on Research Methods: A consultant's companion*. Johannes van Kessel Publishing.

[4] Afari, E., Aldridge, J.M., Fraser, B.J. and Khine, M.S. 2013. Students' perceptions of the learning environment and attitudes in game-based mathematics classrooms. *Learning Environments Research*. 16, 1 (Apr. 2013), 131–150. DOI:https://doi.org/10.1007/s10984-012-9122-6.

[5] Ainley, M. and Ainley, J. 2011. Student engagement with science in early adolescence : The contribution of enjoyment to students ' continuing interest in learning about science. *Contemporary Educational Psychology*. 36, 1 (2011), 4–12. DOI:https://doi.org/10.1016/j.cedpsych.2010.08.001.

[6] Akinsola, M.K. and Animasahun, I.A. 2007. The Effect of Simulation-Games Environment on Students Achievement in and Attitudes To Mathematics in Secondary Schools. *The Turkish Online Journal of Educational Technology – TOJET July*. 6, 3 (2007), 1303–6521.

[7] Amos, B., Ludwiczuk, B. and Satyanarayanan, M. 2016. *OpenFace: A general-purpose face recognition library with mobile applications. Technical Report. CMU-CS-16-118, CMU School of Computer Science.*

[8] An, Z., Wang, C., Li, S., Gan, Z. and Li, H. 2021. Technology-Assisted Self-Regulated English Language Learning: Associations With English Language Self-Efficacy, English Enjoyment, and Learning Outcomes. *Frontiers in Psychology*. 11, January (2021). DOI:https://doi.org/10.3389/fpsyg.2020.558466.

[9] Andrade, A., Danish, J.A. and Maltese, A. V. 2017. A Measurement Model of Gestures in an Embodied Learning Environment: Accounting for Temporal Dependencies. *Journal of Learning Analytics*. 4, 3 (Dec. 2017). DOI:https://doi.org/10.18608/jla.2017.43.3.

[10] Andrade, A., Delandshere, G. and Danish, J.A. 2016. Using Multimodal Learning Analytics to Model Student Behavior: A Systematic Analysis of Epistemological Framing. *Journal of Learning Analytics*. 3, 2 (Sep. 2016), 282–306. DOI:https://doi.org/10.18608/jla.2016.32.14.

[11] Anguera, J.A., Brandes-Aitken, A.N., Antovich, A.D., Rolle, C.E., Desai, S.S. and Marco, E.J. 2017. A pilot study to determine the feasibility of enhancing cognitive abilities in children with sensory processing dysfunction. *PLOS ONE*. 12, 4 (Apr. 2017), e0172616. DOI:https://doi.org/10.1371/journal.pone.0172616.

[12] Aoki, N., Ohta, S., Masuda, H., Naito, T., Sawai, T., Nishida, K., Okada, T., Oishi, M., Iwasawa, Y., Toyomasu, K., Hira, K. and Fukui, T. 2004. Edutainment tools for

initial education of type-1 diabetes mellitus: Initial diabetes education with fun. *Studies in Health Technology and Informatics*. 107, May 2014 (2004), 855–859. DOI:https://doi.org/10.3233/978-1-60750-949-3-855.

[13] Archer, L., Dawson, E., DeWitt, J., Seakins, A. and Wong, B. 2015. "Science capital": A conceptual, methodological, and empirical argument for extending bourdieusian notions of capital beyond the arts. *Journal of Research in Science Teaching*. 52, 7 (2015), 922–948. DOI:https://doi.org/10.1002/tea.21227.

[14] Archer, L., Moote, J., MacLeod, E., Francis, B. and DeWitt, J. 2020. *ASPIRES 2: Young people's science and career aspirations, age 10–19*.

[15] Artino, A.R. and Jones, K.D. 2012. Exploring the complex relations between achievement emotions and self-regulated learning behaviors in online learning. *Internet and Higher Education*. 15, 3 (2012), 170–175. DOI:https://doi.org/10.1016/j.iheduc.2012.01.006.

[16] Artino, A.R. and Stephens, J.M. 2009. Beyond Grades in Online Learning: Adaptive Profiles of Academic Self-Regulation Among Naval Academy Undergraduates. *Journal of Advanced Academics*. 20, 4 (Aug. 2009), 568–601. DOI:https://doi.org/10.1177/1932202X0902000402.

[17] Atwood-Blaine, D., Rule, A.C. and Walker, J. 2019. Creative self-efficacy of children aged 9-14 in a science center using a situated Mobile game. *Thinking Skills and Creativity*. 33, January (2019), 100580. DOI:https://doi.org/10.1016/j.tsc.2019.100580.

[18] Azevedo, R., Mudrick, N. V, Taub, M. and Bradbury, A.E. 2019. Self-Regulation in Computer-Assisted Learning Systems. 587–618.

[19] Bakar, K.A., Tarmizi, R.A., Mahyuddin, R., Elias, H., Luan, W.S. and Ayub, A.F.M. 2010. Relationships between university students' achievement motivation, attitude and academic performance in Malaysia. *Procedia - Social and Behavioral Sciences*. 2, 2 (2010), 4906–4910. DOI:https://doi.org/10.1016/j.sbspro.2010.03.793.

[20] Baños, R.M., Cebolla, A., Oliver, E., Alcañiz, M. and Botella, C. 2013. Efficacy and acceptability of an Internet platform to improve the learning of nutritional knowledge in children: The ETIOBE mates. *Health Education Research*. 28, 2 (2013), 234–248. DOI:https://doi.org/10.1093/her/cys044.

[21] Barnard, L., Lan, W.Y., To, Y.M., Paton, V.O. and Lai, S.-L. 2009. Measuring self-regulation in online and blended learning environments. *The Internet and Higher Education*. 12, 1 (Jan. 2009), 1–6. DOI:https://doi.org/10.1016/j.iheduc.2008.10.005.

[22] Barrett, L.F., Adolphs, R., Marsella, S., Martinez, A.M. and Pollak, S.D. 2019. Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements. *Psychological Science in the Public Interest*. 20, 3 (Dec. 2019), 165–166. DOI:https://doi.org/10.1177/1529100619889954.

[23] Barria-Pineda, J., Akhuseyinoglu, K., Brusilovsky, P., Pollari-Malmi, K., Sirkiä, T. and Malmi, L. 2020. Personalized Remedial Recommendations for SQL Programming Practice System. *UMAP 2020 Adjunct - Adjunct Publication of the 28th ACM Conference on User Modeling, Adaptation and Personalization*. Apcse (2020), 135–142. DOI:https://doi.org/10.1145/3386392.3399312.

[24] Barzilai, S. and Blau, I. 2014. Scaffolding game-based learning: Impact on learning achievements, perceived learning, and game experiences. *Computers and Education*. 70, (2014), 65–79. DOI:https://doi.org/10.1016/j.compedu.2013.08.003.

[25] Baser, M. 2013. Attitude, gender and achievement in computer programming.

Middle East Journal of Scientific Research. 14, 2 (2013), 248–255. DOI:https://doi.org/10.5829/idosi.mejsr.2013.14.2.2007.

[26] Beaton, D.E., Bombardier, C., Guillemin, F. and Ferraz, M.B. 2000. Guidelines for the Process of Cross-Cultural Adaptation of Self-Report Measures. *Spine*. 25, 24 (2000), 3186–3191.

[27] Bell, A. 2007. Designing and testing questionaires for children. *Journal of Research in Nursing*. 12, 5 (2007), 461–469. DOI:https://doi.org/10.1177/1744987107079616.

[28] Bergkvist, L. and Rossiter, J.R. 2007. The predictive validity of multiple-item versus single-item measures of the same constructs. *Journal of Marketing Research*. 44, 2 (2007), 175–184. DOI:https://doi.org/10.1509/jmkr.44.2.175.

[29] Bernard, M., Mills, M., Talissa, F. and McKown, J. 2001. Which Fonts Do Children Prefer to Read Online? *Usability News*. 3, 1 (2001).

[30] Bezruczko, N., Fatani, S.S. and Magari, N. 2016. Three Tales of Change: Ordinal Scores, Residualized Gains, and Rasch Logits—When Are They Interchangeable? *SAGE Open*. 6, 3 (2016). DOI:https://doi.org/10.1177/2158244016659905.

[31] Bidin, S., Jusoff, K., Aziz, N.A., Salleh, M.M. and Tajudin, T. 2009. Motivation and Attitude in Learning English among UiTM Students in the Northern Region of Malaysia. *English Language Teaching*. 2, 2 (2009), 16–20. DOI:https://doi.org/10.5539/elt.v2n2p16.

[32] Bisson, C. and Luckner, J. 1996. Fun in Learning: The Pedagogical Role of Fun in Adventure Education. *The Journal of Experiential Education*. 19, 2 (1996), 108–112. DOI:https://doi.org/10.1177/105382599601900208.

[33] Blikstein, P. and Worsley, M. 2016. Multimodal Learning Analytics and Education Data Mining: using computational technologies to measure complex learning tasks. *Journal of Learning Analytics*. 3, 2 (Sep. 2016), 220–238. DOI:https://doi.org/10.18608/jla.2016.32.11.

[34] Bloom, B. 1956. *Taxonomy of educational objectives: The classification of educational goals*. David McKay Co Inc.

[35] Blums, A., Belsky, J., Grimm, K. and Chen, Z. 2017. Building Links Between Early Socioeconomic Status, Cognitive Ability, and Math and Science Achievement. *Journal of Cognition and Development*. 18, 1 (2017), 16–40. DOI:https://doi.org/10.1080/15248372.2016.1228652.

[36] Bontchev, B. and Vassileva, D. 2016. Assessing engagement in an emotionally-adaptive applied game. *Proceedings of the Fourth International Conference on Technological Ecosystems for Enhancing Multiculturality* (New York, NY, USA, Nov. 2016), 747–754.

[37] Borgers, N. 2003. Response Quality in Survey Research with Children and Adolescents: The Effect of Labeled Response Options and Vague Quantifiers. *International Journal of Public Opinion Research*. 15, 1 (2003), 83–94. DOI:https://doi.org/10.1093/ijpor/15.1.83.

[38] Borgers, N., de Leeuw, E. and Hox, J. 2000. Children as Respondents in Survey Research: Cognitive Development and Response Quality 1. *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*. 66, 1 (Apr. 2000), 60–75. DOI:https://doi.org/10.1177/075910630006600106.

[39] Boulton, C.A., Kent, C. and Williams, H.T.P. 2018. Virtual learning environment engagement and learning outcomes at a ‘bricks-and-mortar’ university. *Computers & Education*. 126, (Nov. 2018), 129–142.

DOI:https://doi.org/10.1016/j.compedu.2018.06.031.

[40] Boyle, E.A., Hainey, T., Connolly, T.M., Gray, G., Earp, J., Ott, M., Lim, T., Ninaus, M., Ribeiro, C. and Pereira, J. 2016. An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games. *Computers and Education*. 94, (2016), 178–192. DOI:https://doi.org/10.1016/j.compedu.2015.11.003.

[41] Bradbury, N.A. 2016. Attention span during lectures: 8 seconds, 10 minutes, or more? *Advances in Physiology Education*. 40, 4 (2016), 509–513. DOI:https://doi.org/10.1152/advan.00109.2016.

[42] Bradley, M.M. and Lang, P.J. 1994. Measuring Emotion: The Self-Assessment Manikin and the Semantic Differential. *Journal of Behavior Therapy and Experimental Psychiatry*. 25, 1 (Mar. 1994), 49–59. DOI:https://doi.org/10.1016/0005-7916(94)90063-9.

[43] Briscoe, G. and Mulligan, C. 2014. Digital Innovation: The Hackathon Phenomenon. *Creativeworks London*. 6 (2014), 1–13.

[44] Burguillo, J.C. 2010. Using game theory and Competition-based Learning to stimulate student motivation and performance. *Computers and Education*. 55, 2 (2010), 566–575. DOI:https://doi.org/10.1016/j.compedu.2010.02.018.

[45] Byun, J. and Joung, E. 2018. Digital game-based learning for K-12 mathematics education: A meta-analysis. *School Science and Mathematics*. 118, 3–4 (Apr. 2018), 113–126. DOI:https://doi.org/10.1111/ssm.12271.

[46] Cacioppo, J.T. and Tassinary, L.G. 1990. Inferring psychological significance from physiological signals. *American Psychologist*. 45, 1 (1990), 16–28. DOI:https://doi.org/10.1037/0003-066X.45.1.16.

[47] Caine, R.N. and Caine, G. 1991. *Making connections: Teaching and the Human Brain*. Association for Supervision and Curriculum Development.

[48] Çankaya, S. and Karamete, A. 2009. The effects of educational computer games on students' attitudes towards mathematics course and educational computer games. *Procedia - Social and Behavioral Sciences*. 1, 1 (2009), 145–149. DOI:https://doi.org/10.1016/j.sbspro.2009.01.027.

[49] Caputo, A. 2017. A Brief Scale on Attitude Towards Learning of Scientific Subjects (ATLoSS) for Middle School Students. *Journal of Educational, Cultural and Psychological Studies*. 2017, 16 (2017), 57–76. DOI:https://doi.org/10.7358/ecps-2017-016-capu.

[50] Cetin, I. and Ozden, M.Y. 2015. Development of computer programming attitude scale for university students. *Computer Applications in Engineering Education*. 23, 5 (2015), 667–672. DOI:https://doi.org/10.1002/cae.21639.

[51] Chan, S.C.H., Wan, J.C.L. and Ko, S. 2019. Interactivity, active collaborative learning, and learning performance: The moderating role of perceived fun by using personal response systems. *International Journal of Management Education*. 17, 1 (2019), 94–102. DOI:https://doi.org/10.1016/j.ijme.2018.12.004.

[52] Chi, M.T.H. and Wylie, R. 2014. The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes. *Educational Psychologist*. 49, 4 (Oct. 2014), 219–243. DOI:https://doi.org/10.1080/00461520.2014.965823.

[53] Chu, S.L., Angello, G., Saenz, M. and Quek, F. 2017. Fun in Making: Understanding the experience of fun and learning through curriculum-based Making in the elementary school classroom. *Entertainment Computing*. 18, (2017),

31–40. DOI:https://doi.org/10.1016/j.entcom.2016.08.007.

[54] Clark, D.B., Tanner-Smith, E.E. and Killingsworth, S.S. 2016. Digital Games, Design, and Learning. *Review of Educational Research*. 86, 1 (Mar. 2016), 79–122. DOI:https://doi.org/10.3102/0034654315582065.

[55] Clarke, V. and Braun, V. 2017. Thematic analysis. *Journal of Positive Psychology*. 12, 3 (2017), 297–298. DOI:https://doi.org/10.1080/17439760.2016.1262613.

[56] Cohen, J., Cohen, P., West, S.G. and Aiken, L.S. 2014. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. Psychology Press.

[57] Cohn, J.F., Ambadar, Z. and Ekman, P. 2007. Observer-based measurement of facial expression with the Facial Action Coding System. e. *The handbook of emotion elicitation and assessment,*. 203–221.

[58] Connolly, T.M., Boyle, E.A., MacArthur, E., Hainey, T. and Boyle, J.M. 2012. A systematic literature review of empirical evidence on computer games and serious games. *Computers and Education*. 59, 2 (2012), 661–686. DOI:https://doi.org/10.1016/j.compedu.2012.03.004.

[59] Costa, C.J., Aparicio, M., Aparicio, S. and Aparicio, J.T. 2017. Gamification usage ecology. *SIGDOC 2017 - 35th ACM International Conference on the Design of Communication*. (2017). DOI:https://doi.org/10.1145/3121113.3121205.

[60] Craig, S., Graesser, A., Sullins, J. and Gholson, B. 2004. Affect and learning: An exploratory look into the role of affect in learning with AutoTutor. *Journal of Educational Media*. 29, 3 (2004), 241–250. DOI:https://doi.org/10.1080/1358165042000283101.

[61] Csikszentmihalyi, M. 1990. *Flow: The psychology of optimal experience*. Harper & Row.

[62] Cukurova, M., Giannakos, M. and Martinez-Maldonado, R. 2020. The promise and challenges of multimodal learning analytics. *British Journal of Educational Technology*. 51, 5 (Sep. 2020), 1441–1449. DOI:https://doi.org/10.1111/bjet.13015.

[63] Cukurova, M., Kent, C. and Luckin, R. 2019. Artificial intelligence and multimodal data in the service of human decision-making: A case study in debate tutoring. *British Journal of Educational Technology*. 50, 6 (Nov. 2019), 3032–3046. DOI:https://doi.org/10.1111/bjet.12829.

[64] Curran, P.J., West, S.G. and Finch, J.F. 1996. The robustness of test statistics to nonnormality and specification error in confirmatory factory analysis. *Psychological Methods*. 1, 1 (1996), 16–29.

[65] D'Mello, S. and Graesser, A. 2012. Dynamics of affective states during complex learning. *Learning and Instruction*. 22, 2 (Apr. 2012), 145–157. DOI:https://doi.org/10.1016/j.learninstruc.2011.10.001.

[66] Davis, B.G. 2009. *Tools for Teaching*. Jossey-Bass.

[67] Deci, E.L. and Ryan, R.M. 1985. *Intrinsic Motivation and Self-Determination in Human Behavior*. Springer US.

[68] Deterding, S., Dixon, D., Khaled, R. and Nacke, L. 2011. From game design elements to gamefulness: Defining "gamification." *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments, MindTrek 2011*. March 2014 (2011), 9–15. DOI:https://doi.org/10.1145/2181037.2181040.

[69] DeWitt, J. and Archer, L. 2015. Who Aspires to a Science Career? A comparison of survey responses from primary and secondary school students. *International*

Journal of Science Education. 37, 13 (Sep. 2015), 2170–2192. DOI:https://doi.org/10.1080/09500693.2015.1071899.

[70] Dickey, M.D. 2011. Murder on Grimm Isle: The impact of game narrative design in an educational game-based learning environment. *British Journal of Educational Technology*. 42, 3 (May 2011), 456–469. DOI:https://doi.org/10.1111/j.1467-8535.2009.01032.x.

[71] Dillon, R. 2010. *On the way to fun: An emotion-based approach to successful game design*.

[72] Dismore, H. and Bailey, R. 2011. Fun and enjoyment in physical education: Young people's attitudes. *Research Papers in Education*. 26, 4 (2011), 499–516. DOI:https://doi.org/10.1080/02671522.2010.484866.

[73] Dunn, T.J., Baguley, T. and Brunsden, V. 2014. From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*. 105, 3 (2014), 399–412. DOI:https://doi.org/10.1111/bjop.12046.

[74] Ekman, P. 1992. An argument for basic emotions. *Cognition and Emotion*. 6, 3–4 (May 1992), 169–200. DOI:https://doi.org/10.1080/02699939208411068.

[75] Ekman, P. and Friesen, W. V. 1978. *The facial action coding system: a technique for the measurement of facial movement*. Consulting Psychologists Press.

[76] Ekman, P., Friesen, W. V. and Hager, J.C. 2002. *Facial Action Coding System. The Manual On CD ROM*.

[77] Ekman, P., Levenson, R.W. and Friesen, W. V. 1983. Autonomic Nervous System Activity Distinguishes Among Emotions. *Science*. 221, 4616 (Sep. 1983), 1208–1210. DOI:https://doi.org/10.1126/science.6612338.

[78] Elton-Chalcraft, S. and Mills, K. 2015. Measuring challenge, fun and sterility on a 'phunometre' scale: evaluating creative teaching and learning with children and their student teachers in the primary school. *Education 3-13*. 43, 5 (2015), 482–497. DOI:https://doi.org/10.1080/03004279.2013.822904.

[79] Erhel, S. and Jamet, E. 2013. Digital game-based learning: Impact of instructions and feedback on motivation and learning effectiveness. *Computers and Education*. 67, (2013), 156–167. DOI:https://doi.org/10.1016/j.compedu.2013.02.019.

[80] Ericsson, K.A. and Simon, H.A. 1980. Verbal reports as data. *Psychological Review*. 87, 3 (May 1980), 215–251. DOI:https://doi.org/10.1037/0033-295X.87.3.215.

[81] Erikson, E.H. 1950. *Childhood and Society*. Norton & Company.

[82] Erkman, F., Caner, A., Hande Sart, Z., Börkan, B. and Şahan, K. 2010. Influence of Perceived Teacher Acceptance, Self-Concept, and School Attitude on the Academic Achievement of School-Age Children in Turkey. *Cross-Cultural Research*. 44, 3 (Aug. 2010), 295–309. DOI:https://doi.org/10.1177/1069397110366670.

[83] Eshach, H. 2007. Bridging In-school and Out-of-school Learning: Formal, Non-formal, and Infromal Education. *Journal of Science Education and Technology*. 16, 2 (2007), 171–190. DOI:https://doi.org/10.1007/s10956-006-9027-1.

[84] Fedon, J.P. 1958. The Role of Attitude in Learning Arithmetic. *The Arithmetic Teacher*. 5, 6 (1958), 304–310.

[85] FitzGerald, E., Kucirkova, N., Jones, A., Cross, S., Ferguson, R., Herodotou, C., Hillaire, G. and Scanlon, E. 2018. Dimensions of personalisation in technology-enhanced learning: A framework and implications for design. *British Journal of*

Educational Technology. 49, 1 (2018), 165–181. DOI:https://doi.org/10.1111/bjet.12534.

[86] Flavell, J.H. 1979. Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*. 34, 10 (1979), 906–911. DOI:https://doi.org/10.1037/0003-066X.34.10.906.

[87] Flesch, R. 1948. A new readability yardstick. *Journal of Applied Psychology*. 32, 3 (1948), 221–233.

[88] Fowler, A. 2016. Informal STEM learning in game jams, ackathons and game creation events. *Proceedings of the International Conference on Game Jams, Hackathons, and Game Creation Events, GJH and GC 2016*. (2016), 38–41. DOI:https://doi.org/10.1145/2897167.2897179.

[89] Fowler, A. 2013. Measuring learning and fun in video games for young children: A proposed method. *Proceedings of the 12th International Conference on Interaction Design and Children - IDC '13* (New York, New York, USA, 2013), 639.

[90] Frenzel, A.C., Pekrun, R. and Goetz, T. 2007. Perceived learning environment and students' emotional experiences: A multilevel analysis of mathematics classrooms. *Learning and Instruction*. 17, 5 (2007), 478–493. DOI:https://doi.org/10.1016/j.learninstruc.2007.09.001.

[91] Fu, F.L., Su, R.C. and Yu, S.C. 2009. EGameFlow: A scale to measure learners' enjoyment of e-learning games. *Computers and Education*. 52, 1 (2009), 101–112. DOI:https://doi.org/10.1016/j.compedu.2008.07.004.

[92] Gajadhar, B.J., de Kort, Y.A.W. and IJsselsteijn, W.A. 2008. Shared Fun is Doubled Fun: Player Enjoyment as a Function of Social Setting. *Proceedings of Fun and Games Second International COnference Eindhoven*. 106–117.

[93] Gao, F., Li, L. and Sun, Y. 2020. A systematic review of mobile game-based learning in STEM education. *Educational Technology Research and Development*. 68, 4 (Aug. 2020), 1791–1827. DOI:https://doi.org/10.1007/s11423-020-09787-0.

[94] Garn, A.C. and Cothran, D.J. 2006. The Fun Factor in Physical Education. *Journal of Teaching in Physical Education*. 25, 3 (2006), 281–297.

[95] Giannakos, M.N., Sharma, K., Papavlasopoulou, S., Pappas, I.O. and Kostakos, V. 2020. Fitbit for learning: Towards capturing the learning experience using wearable sensing. *International Journal of Human Computer Studies*. 136, November 2019 (2020). DOI:https://doi.org/10.1016/j.ijhcs.2019.102384.

[96] Giannakos, M.N., Sharma, K., Pappas, I.O., Kostakos, V. and Velloso, E. 2019. Multimodal data as a means to understand the learning experience. *International Journal of Information Management*. 48, (Oct. 2019), 108–119. DOI:https://doi.org/10.1016/j.ijinfomgt.2019.02.003.

[97] Girls Who Code 2019. *Advocacy Report 2019 - The State of Girls in K-12 Computer Science Classrooms: Making the Case for Gender-Specific Education Policies.*

[98] Glasser, W. 1986. *Control theory in the classroom*. Perennial Library/Harper & Row Publishers.

[99] Godec, S., King, H. and Arthur, L. 2017. The Science Capital Teaching Approach. *University College London*. (2017).

[100] Goss Lucas, S. and Bernstein, D.A. 2005. *Teaching Psychology: A Step by Step Guide*. Lawrence Erlbaum Associates.

[101] Graesser, A.C. 2019. Emotions are the experiential glue of learning environments in the 21st century. *Learning and Instruction*. May (2019), 101212.

DOI:https://doi.org/10.1016/j.learninstruc.2019.05.009.
[102] Graesser, A.C., Lu, S., Olde, B.A., Cooper-Pye, E. and Whitten, S. 2005. Question asking and eye tracking during cognitive disequilibrium: Comprehending illustrated texts on devices when the devices break down. *Memory & Cognition*. 33, 7 (Oct. 2005), 1235–1247. DOI:https://doi.org/10.3758/BF03193225.
[103] Grosshandler, D.J. and Niswander Grosshandler, E.N. 2000. Constructing fun: Self-determination and learning at an afterschool design lab. *Computers in Human Behavior*. 16, 3 (2000), 227–240. DOI:https://doi.org/10.1016/S0747-5632(00)00003-0.
[104] Gunbatar, M.S. and Karalar, H. 2018. Gender differences in middle school students' attitudes and self-efficacy perceptions towards MBlock programming. *European Journal of Educational Research*. 7, 4 (2018), 925–933. DOI:https://doi.org/10.12973/eu-jer.7.4.923.
[105] Guo, J. 2018. Building bridges to student learning: Perceptions of the learning environment, engagement, and learning outcomes among Chinese undergraduates. *Studies in Educational Evaluation*. 59, (Dec. 2018), 195–208. DOI:https://doi.org/10.1016/j.stueduc.2018.08.002.
[106] Hainey, T., Connolly, T.M., Boyle, E.A., Wilson, A. and Razak, A. 2016. A systematic literature review of games-based learning empirical evidence in primary education. *Computers and Education*. 102, January 2004 (2016), 202–223. DOI:https://doi.org/10.1016/j.compedu.2016.09.001.
[107] Hair Jr., J.F., Black, W.C., Babin, B.J. and Anderson, R.E. 2014. *Multivariate Data Analysis*. Pearson.
[108] Hall, L., Hume, C. and Tazzyman, S. 2016. Five Degrees of Happiness: Effective Smiley Face Likert Scales for Evaluating with Children. *Proceedings of the The 15th International Conference on Interaction Design and Children*. (2016), 311–321. DOI:https://doi.org/10.1145/2930674.2930719.
[109] Hard Fun: *http://www.papert.org/articles/HardFun.html*.
[110] Harris, D. 2008. *A comparative study of the effect of collaborative problem solving in a massively multiplayer online game (MMOG) on individual achievement*.
[111] Harris, K. and Reid, D. 2005. The influence of virtual reality play on children's motivation. *Canadian Journal of Occupational Therapy*. 72, 1 (2005), 21–29. DOI:https://doi.org/10.1177/000841740507200107.
[112] Hascher, T. 2010. Learning and emotion: Perspectives for theory and research. *European Educational Research Journal*. 9, 1 (2010), 13–28. DOI:https://doi.org/10.2304/eerj.2010.9.1.13.
[113] Hatlevik, O.E. and Christophersen, K.A. 2013. Digital competence at the beginning of upper secondary school: Identifying factors explaining digital inclusion. *Computers and Education*. 63, (2013), 240–247. DOI:https://doi.org/10.1016/j.compedu.2012.11.015.
[114] Heaven, D. 2020. Why faces don't always tell the truth about feelings. *Nature*. 578, 7796 (Feb. 2020), 502–504. DOI:https://doi.org/10.1038/d41586-020-00507-5.
[115] Hilbert, S., Bruckmaier, G., Binder, K., Krauss, S. and Bühner, M. 2019. Prediction of elementary mathematics grades by cognitive abilities. *European Journal of Psychology of Education*. 34, 3 (Jul. 2019), 665–683. DOI:https://doi.org/10.1007/s10212-018-0394-9.
[116] Hobza, V., Hamrik, Z., Bucksch, J. and De Clercq, B. 2017. The family affluence

scale as an indicator for socioeconomic status: Validation on regional income differences in the Czech Republic. *International Journal of Environmental Research and Public Health*. 14, 12 (2017). DOI:https://doi.org/10.3390/ijerph14121540.

[117] Hu, L. and Bentler, P.M. 1999. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*. 6, 1 (1999), 1–55. DOI:https://doi.org/10.1080/10705519909540118.

[118] Hu, Z. and Li, J. 2015. The Integration of EFA and CFA: One Method of Evaluating the Construct Validity. *Global Journal of Human-Social Science: Arts & Humanities -Psychology*. 15, 6 (2015), 15–19.

[119] Huang, W.-H., Huang, W.-Y. and Tschopp, J. 2010. Sustaining iterative game playing processes in DGBL: The relationship between motivational processing and outcome processing. *Computers & Education*. 55, 2 (Sep. 2010), 789–797. DOI:https://doi.org/10.1016/j.compedu.2010.03.011.

[120] Huizenga, J., Admiraal, W., Akkerman, S. and Ten Dam, G. 2009. Mobile game-based learning in secondary education: engagement, motivation and learning in a mobile city game: Original article. *Journal of Computer Assisted Learning*. 25, 4 (2009), 332–344. DOI:https://doi.org/10.1111/j.1365-2729.2009.00316.x.

[121] Huizinga, J. 1949. *Homo Ludens - A Study of the Play-element in Culture*. Routledge & Kegan Paul.

[122] Hung, C.-M., Huang, I. and Hwang, G.-J. 2014. Effects of digital game-based learning on students' self-efficacy, motivation, anxiety, and achievements in learning mathematics. *Journal of Computers in Education*. 1, 2–3 (2014), 151–166. DOI:https://doi.org/10.1007/s40692-014-0008-8.

[123] Hung, H.T., Yang, J.C., Hwang, G.J., Chu, H.C. and Wang, C.C. 2018. A scoping review of research on digital game-based language learning. *Computers and Education*. 126, July (2018), 89–104. DOI:https://doi.org/10.1016/j.compedu.2018.07.001.

[124] Hussein, M.H., Ow, S.H., Cheong, L.S., Thong, M.-K. and Ale Ebrahim, N. 2019. Effects of Digital Game-Based Learning on Elementary Science Learning: A Systematic Review. *IEEE Access*. 7, (2019), 62465–62478. DOI:https://doi.org/10.1109/ACCESS.2019.2916324.

[125] Hwang, G.-J. and Wu, P.-H. 2012. Advancements and trends in digital game-based learning research: a review of publications in selected journals from 2001 to 2010. *British Journal of Educational Technology*. 43, 1 (Jan. 2012), E6–E10. DOI:https://doi.org/10.1111/j.1467-8535.2011.01242.x.

[126] IJsselsteijn, W., Hoogen, W. Van Den, Klimmt, C., Kort, Y. de, Lindley, C., Mathiak, K., Poels, K., Ravaja, N., Turpeinen, M. and Vorderer, P. 2008. Measuring the Experience of Digital Game Enjoyment. *Proceedings of Measuring Behavior 2008*. 2008, (2008), 88–89. DOI:https://doi.org/10.3758/BRM.41.3.717.

[127] Iten, N. and Petko, D. 2016. Learning with serious games: Is fun playing the game a predictor of learning success? *British Journal of Educational Technology*. 47, 1 (2016), 151–163. DOI:https://doi.org/10.1111/bjet.12226.

[128] Jackson, D.L., Gillaspy, J.A. and Purc-Stephenson, R. 2009. Reporting Practices in Confirmatory Factor Analysis: An Overview and Some Recommendations. *Psychological Methods*. 14, 1 (2009), 6–23. DOI:https://doi.org/10.1037/a0014694.

[129] Jackson, P.W. and Getzels, J.W. 1959. Psychological health and classroom

functioning: A study of dissatisfaction with school among adolescents. *Journal of Educational Psychology*. 50, 6 (1959), 295–300. DOI:https://doi.org/10.1037/h0039656.

[130] Jackson, P.W. and Lahaderne, H.M. 1967. Scholastic success and attitude toward school in a population of sixth graders. *Journal of Educational Psychology*. 58, 1 (1967), 15–18. DOI:https://doi.org/10.1037/h0024233.

[131] Jackson, S.A. and Marsh, H.W. 1996. Development and validation of a scale to measure optimal experience: The flow state scale. *Journal of Sport and Exercise Psychology*. 18, 1 (1996), 17–35. DOI:https://doi.org/10.1123/jsep.18.1.17.

[132] Jarke, J. and Breiter, A. 2019. Editorial: the datafication of education. *Learning, Media and Technology*. 44, 1 (2019), 1–6. DOI:https://doi.org/10.1080/17439884.2019.1573833.

[133] Jeno, L.M., Vandvik, V., Eliassen, S. and Grytnes, J.-A. 2019. Testing the novelty effect of an m-learning tool on internalization and achievement: A Self-Determination Theory approach. *Computers & Education*. 128, (Jan. 2019), 398–413. DOI:https://doi.org/10.1016/j.compedu.2018.10.008.

[134] Joëls, M., Pu, Z., Wiegert, O., Oitzl, M.S. and Krugers, H.J. 2006. Learning under stress: how does it work? *Trends in Cognitive Sciences*. 10, 4 (Apr. 2006), 152–158. DOI:https://doi.org/10.1016/j.tics.2006.02.002.

[135] Johnson, D., Gardner, M.J. and Perry, R. 2018. Validation of two game experience scales: The Player Experience of Need Satisfaction (PENS) and Game Experience Questionnaire (GEQ). *International Journal of Human Computer Studies*. 118, February (2018), 38–46. DOI:https://doi.org/10.1016/j.ijhcs.2018.05.003.

[136] Jones, A., Bull, S. and Castellano, G. 2018. "I Know That Now, I'm Going to Learn This Next" Promoting Self-regulated Learning with a Robotic Tutor. *International Journal of Social Robotics*. 10, 4 (2018), 439–454. DOI:https://doi.org/10.1007/s12369-017-0430-y.

[137] Jöreskog, K.G. 1999. *How large can a standardized coefficient be? Unplublished report*.

[138] Jorgensen, T.D., Pornprasertmanit, S., Schoemann, A.M. and Rosseel, Y. 2018. semTools: Useful tools for structural equation modeling. R package version 0.5-1.

[139] Kalelioğlu, F. 2015. A new way of teaching programming skills to K-12 students: Code.org. *Computers in Human Behavior*. 52, (2015), 200–210. DOI:https://doi.org/10.1016/j.chb.2015.05.047.

[140] Kalelioğlu, F. and Gülbahar, Y. 2014. The effects of teaching programming via Scratch on problem solving skills: A discussion from learners' perspective. *Informatics in Education*. 13, 1 (2014), 33–50.

[141] Kalogiannakis, M., Papadakis, S. and Zourmpakis, A.I. 2021. Gamification in science education. A systematic review of the literature. *Education Sciences*. 11, 1 (2021), 1–36. DOI:https://doi.org/10.3390/educsci11010022.

[142] Kangas, M., Siklander, P., Randolph, J. and Ruokamo, H. 2017. Teachers' engagement and students' satisfaction with a playful learning environment. *Teaching and Teacher Education*. 63, (2017), 274–284. DOI:https://doi.org/10.1016/j.tate.2016.12.018.

[143] Kankaanranta, M., Koivula, M., Laakso, M.-L. and Mustola, M. 2017. Digital Games in Early Childhood: Broadening Definitions of Learning, Literacy, and Play. *Serious Games and Edutainment Applications*. Springer International

Publishing. 349–367.

[144] Ke, F. 2011. A Qualitative Meta-Analysis of Computer Games as Learning Tools. *Gaming and Simulations*. (2011). DOI:https://doi.org/10.4018/9781609601959.ch701.

[145] Kendzierski, D. and DeCarlo, K.J. 1991. Physical Activity Enjoyment Scale: Two Validation Studies. *Journal of Sport and Exercise Psychology*. 13, 1 (Mar. 1991), 50–64. DOI:https://doi.org/10.1123/jsep.13.1.50.

[146] Killen, H., Weintrop, D. and Garvin, M. 2019. AP Computer Science Principles' Impact on the Landscape of High School Computer Science using Maryland as a Model. *Proceedings of the 50th ACM Technical Symposium on Computer Science Education* (New York, NY, USA, Feb. 2019), 1060–1066.

[147] Kim, C. and Pekrun, R. 2014. Emotions and Motivation in Learning and Performance. *Handbook of Research on Educational Communications and Technology*. J.M. Spector, M.D. Merrill, J. Elen, and M.J. Bishop, eds. Springer New York.

[148] Kincaid, J.P., Fishburne Jr., R.P., Rogers, R.L. and Chissom, B.S. 1975. Derivation of New Readability Formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy Enlisted Personnel. (1975). DOI:https://doi.org/10.21236/ADA006655.

[149] Kline, R.B. 2015. *Principles and Practice of Structural Equation Modeling*. The Guilford Press.

[150] Knowles, C., Harris, A. and Van Norman, R. 2017. Family Fun Nights: Collaborative Parent Education Accessible for Diverse Learning Abilities. *Early Childhood Education Journal*. 45, 3 (2017), 393–401. DOI:https://doi.org/10.1007/s10643-016-0801-2.

[151] Kong, S.C., Chiu, M.M. and Lai, M. 2018. A study of primary school students' interest, collaboration attitude, and programming empowerment in computational thinking education. *Computers and Education*. 127, August (2018), 178–189. DOI:https://doi.org/10.1016/j.compedu.2018.08.026.

[152] Koriat, A. and Bjork, R.A. 2005. Illusions of Competence in Monitoring One's Knowledge During Study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 31, 2 (2005), 187–194. DOI:https://doi.org/10.1037/0278-7393.31.2.187.

[153] Korkmaz, Ö. and Altun, H. 2013. Engineering and CEIT student's attitude towards learning computer programming. *Journal of Academic Social Science Studies*. 6, 2 (2013), 1169–1185.

[154] Kosmas, P., Ioannou, A. and Retalis, S. 2018. Moving Bodies to Moving Minds: A Study of the Use of Motion-Based Games in Special Education. *TechTrends*. 62, 6 (Nov. 2018), 594–601. DOI:https://doi.org/10.1007/s11528-018-0294-5.

[155] Kousar, S., Mehmood, N. and Ahmed, S. 2019. Serious Games for Autism Children: A Comparative Study. *University of Sindh Journal of Information and Communication Technology*. 3, 3 (2019), 162–170.

[156] Kucirkova, N. 2019. Children's agency by design: Design parameters for personalization in story-making apps. *International Journal of Child-Computer Interaction*. 21, (2019), 112–120. DOI:https://doi.org/10.1016/j.ijcci.2019.06.003.

[157] Kuyvenhoven, J. and Boterman, W.R. 2021. Neighbourhood and school effects on educational inequalities in the transition from primary to secondary education in Amsterdam. *Urban Studies*. 58, 13 (2021), 2660–2682.

DOI:https://doi.org/10.1177/0042098020959011.

[158] Lara, M. and Lockwood, K. 2016. Hackathons as Community-Based Learning: a Case Study. *TechTrends*. 60, 5 (2016), 486–495. DOI:https://doi.org/10.1007/s11528-016-0101-0.

[159] Di Lascio, E., Gashi, S. and Santini, S. 2018. Unobtrusive Assessment of Students' Emotional Engagement during Lectures Using Electrodermal Activity Sensors. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*. 2, 3 (Sep. 2018), 1–21. DOI:https://doi.org/10.1145/3264913.

[160] Lee-Cultura, S., Sharma, K. and Giannakos, M. 2021. Children's play and problem-solving in motion-based learning technologies using a multi-modal mixed methods approach. *International Journal of Child-Computer Interaction*. (Jul. 2021), 100355. DOI:https://doi.org/10.1016/j.ijcci.2021.100355.

[161] Lee-Cultura, S., Sharma, K., Papavlasopoulou, S., Retalis, S. and Giannakos, M. 2020. Using sensing technologies to explain children's self-representation in motion-based educational games. *Proceedings of the Interaction Design and Children Conference, IDC 2020*. June (2020), 541–555. DOI:https://doi.org/10.1145/3392063.3394419.

[162] Lee, J. 2016. Attitude toward school does not predict academic achievement. *Learning and Individual Differences*. 52, (2016), 1–9. DOI:https://doi.org/10.1016/j.lindif.2016.09.009.

[163] de Leeuw, E.D. 2011. Improving Data Quality when Surveying Children and Adolescents: Cognitive and Social Development and its Role in Questionnaire Construction and Pretesting. (2011).

[164] Lin, C.H., Hsiao, C. and Chen, W. 1999. Development of Sustained Attention Assessed Using the Continuous Performance Test among Children 6-15 Years of Age. *Journal of abnormal child psychology*. 27, 5 (1999), 403–412. DOI:https://doi.org/10.1023/A:1021932119311.

[165] Liu, J., Peng, P. and Luo, L. 2019. The Relationship Between Family Socioeconomic Status and Academic Achievement in China: A Meta-analysis. *Springer Science+Business Media*. (2019), 49–76.

[166] Liu, J., Peng, P., Zhao, B. and Luo, L. 2022. Socioeconomic Status and Academic Achievement in Primary and Secondary Education: a Meta-analytic Review. *Educational Psychology Review*. (Jul. 2022). DOI:https://doi.org/10.1007/s10648-022-09689-y.

[167] Liu, Y. 2014. Motivation and Attitude: Two Important Non-Intelligence Factors to Arouse Students' Potentialities in Learning English. *Creative Education*. 05, 14 (2014), 1249–1253. DOI:https://doi.org/10.4236/ce.2014.514140.

[168] Loderer, K., Pekrun, R. and Lester, J.C. 2020. Beyond cold technology: A systematic review and meta-analysis on emotions in technology-based learning environments. *Learning and Instruction*. 70, (Dec. 2020), 101162. DOI:https://doi.org/10.1016/j.learninstruc.2018.08.002.

[169] Long, J. 2007. Just For Fun: Using Programming Games in Software Programming Training and Education - A Field Study of IBM Robocode Community. *Journal of Information Technology Education*. 6, (2007), 279–290. DOI:https://doi.org/10.28945/216.

[170] Lu, M.-H., Lin, W. and Yueh, H.-P. 2017. Development and Evaluation of a Cognitive Training Game for Older People: A Design-based Approach. *Frontiers*

in Psychology. 8, (Oct. 2017). DOI:https://doi.org/10.3389/fpsyg.2017.01837.

[171] Lucardie, D. 2014. The Impact of Fun and Enjoyment on Adult's Learning. *Procedia - Social and Behavioral Sciences*. 142, (2014), 439–446. DOI:https://doi.org/10.1016/j.sbspro.2014.07.696.

[172] Malone, T.W. 1981. Toward a Theory of Intrinsically Instruction Motivating. *Cognitive Science*. 5, 4 (1981), 333–369. DOI:https://doi.org/10.1207/s15516709cog0504_2.

[173] Malone, T.W. and Lepper, M.R. 1987. Making learning fun: A taxonomy of intrinsic motivations for learning. *Aptitude learning and instruction*.

[174] Maloney, J., Resnick, M., Rusk, N., Silverman, B. and Eastmond, E. 2010. The scratch programming language and environment. *ACM Transactions on Computing Education*. 10, 4 (2010), 1–15. DOI:https://doi.org/10.1145/1868358.1868363.

[175] Di Malta, G., Evans, C. and Cooper, M. 2020. Development and validation of the relational depth frequency scale. *Psychotherapy Research*. 30, 2 (2020), 213–227. DOI:https://doi.org/10.1080/10503307.2019.1585590.

[176] Markopoulos, P., Read, J.C., MacFarlane, S. and Hoysniemi, J. 2008. *Evaluating Children's Interactive Products: Principles and Practices for Interaction Designers*. Morgan-Kafumann.

[177] Mason, S.L. and Rich, P.J. 2020. Development and analysis of the Elementary Student Coding Attitudes Survey. *Computers and Education*. 153, August 2019 (2020), 103898. DOI:https://doi.org/10.1016/j.compedu.2020.103898.

[178] Master, A., Cheryan, S., Moscatelli, A. and Meltzoff, A.N. 2017. Programming experience promotes higher STEM motivation among first-grade girls. *Journal of Experimental Child Psychology*. 160, (2017), 92–106. DOI:https://doi.org/10.1016/j.jecp.2017.03.013.

[179] Mayer, R.E. 2012. Cognitive Learning. *Encyclopedia of the Sciences of Learning*. Springer US. 594–596.

[180] Mayer, R.E. 2019. Searching for the role of emotions in e-learning. *Learning and Instruction*. May (2019), 101213. DOI:https://doi.org/10.1016/j.learninstruc.2019.05.010.

[181] McClelland, M.M., Acock, A.C., Piccinin, A., Rhea, S.A. and Stallings, M.C. 2013. Relations between preschool attention span-persistence and age 25 educational outcomes. *Early Childhood Research Quarterly*. 28, 2 (2013), 314–324. DOI:https://doi.org/10.1016/j.ecresq.2012.07.008.

[182] McKeachie, W.J. and Svinicki, M. 2006. *McKeachie's teaching tips: Strategies, research, and theory for college and university teachers*. Houghton-Mifflin.

[183] McKee, A. 2016. *FUN!*. Palgrave Macmillan UK.

[184] McManus, I.C. and Furnham, A. 2010. "Fun , Fun , Fun : Types of Fun , Attitudes to Fun , and their Relation to Personality and Biographical Factors. *Psychology*. 1, August (2010), 159–168. DOI:https://doi.org/10.4236/psych.2010.13021.

[185] Mehrabian, A. and Russell, J.A. 1974. *An approach to environmental psychology*. The MIT Press.

[186] Mellecker, R., Lyons, E.J. and Baranowski, T. 2013. Disentangling Fun and Enjoyment in Exergames Using an Expanded Design, Play, Experience Framework: A Narrative Review. *Games for Health Journal*. 2, 3 (2013), 142–149. DOI:https://doi.org/10.1089/g4h.2013.0022.

[187] Mellor, D. and Moore, K.A. 2014. The use of likert scales with children. *Journal of Pediatric Psychology*. 39, 3 (2014), 369–379. DOI:https://doi.org/10.1093/jpepsy/jst079.

[188] Meluso, A., Zheng, M., Spires, H.A. and Lester, J. 2012. Enhancing 5th graders' science content knowledge and self-efficacy through game-based learning. *Computers and Education*. 59, 2 (2012), 497–504. DOI:https://doi.org/10.1016/j.compedu.2011.12.019.

[189] Mewton, L., Hodge, A., Gates, N., Visontay, R., Lees, B. and Teesson, M. 2020. A randomised double-blind trial of cognitive training for the prevention of psychopathology in at-risk youth. *Behaviour Research and Therapy*. 132, (Sep. 2020), 103672. DOI:https://doi.org/10.1016/j.brat.2020.103672.

[190] Microsoft Canada 2015. Attention spans. *Consumer insights*. (2015), 1–52.

[191] Mirjafari, S. et al. 2019. Differentiating Higher and Lower Job Performers in the Workplace Using Mobile Sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*. 3, 2 (Jun. 2019), 1–24. DOI:https://doi.org/10.1145/3328908.

[192] Moilanen, K.L. 2006. The adolescent Self-Regulatory inventory: The development and validation of a questionnaire of short-Term and long-term self-Regulation. *Journal of Youth and Adolescence*. 36, 6 (2006), 835–848. DOI:https://doi.org/10.1007/s10964-006-9107-9.

[193] Moote, J., Archer, L., DeWitt, J. and MacLeod, E. 2020. Comparing students' engineering and science aspirations from age 10 to 16: Investigating the role of gender, ethnicity, cultural capital, and attitudinal factors. *Journal of Engineering Education*. 109, 1 (Jan. 2020), 34–51. DOI:https://doi.org/10.1002/jee.20302.

[194] Morgan, D.L. 2019. Commentary—After Triangulation, What Next? *Journal of Mixed Methods Research*. 13, 1 (2019), 6–11. DOI:https://doi.org/10.1177/1558689818780596.

[195] Morrison, T.G., Morrison, M.A. and McCutcheon, J.M. 2017. Best Practice Recommendations for Using Structural Equation Modelling in Psychological Research. *Psychology*. 08, 09 (2017), 1326–1341. DOI:https://doi.org/10.4236/psych.2017.89086.

[196] Moyer, K.E. and von Haller Gilmer, B. 1954. The Concept of Attention Spans in Children. *The Elementary School Journal*. 54, 8 (Apr. 1954), 464–466. DOI:https://doi.org/10.1086/458623.

[197] Munro, D. 2018. CODING THE FUTURE : What Canadian youth and their parents think about coding. (2018).

[198] Nandi, A. and Mandernach, M. 2016. Hackathons as an informal learning platform. *SIGCSE 2016 - Proceedings of the 47th ACM Technical Symposium on Computing Science Education*. (2016), 346–351. DOI:https://doi.org/10.1145/2839509.2844590.

[199] Narmadha, U. and Chamundeswari, S. 2013. Attitude towards Learning of Science and Academic Achievement in Science among Students at the Secondary Level. *Journal of Sociological Research*. 4, 2 (2013), 114–124. DOI:https://doi.org/10.5296/jsr.v4i2.3910.

[200] Nieuwenhuis, J. and Hooimeijer, P. 2016. The association between neighbourhoods and educational achievement, a systematic review and meta-analysis. *Journal of Housing and the Built Environment*. 31, 2 (2016), 321–347.

DOI:https://doi.org/10.1007/s10901-015-9460-7.
[201] Niu, L. 2017. Family Socioeconomic Status and Choice of STEM Major in College: An Analysis of a National Sample. *College Student Journal*. 51, 2 (2017), 298–312.
[202] Nomikou, E., Archer, L. and King, H. 2017. Building “Science Capital” in the Classroom. *School Science Review*. 98, 265 (2017), 118–124.
[203] Nte, S. and Stephens, R. 2008. Videogame Aesthetics and e-Learning: a retro-looking computer game to explain the normal distribution in statistics teaching. *Proceedings of the 2nd European conference on games-based learning (ECGBL)* (Barcelona, Spain., 2008).
[204] O’Brien, H., Cairns, P. and Hall, M. 2018. A Practical Approach to Measuring User Engagement with the Refined User Engagement Scale (UES) and New UES Short Form. *International Journal of Human - Computer Studies*. 112, (2018), 28–39. DOI:https://doi.org/10.1016/j.ijhcs.2018.01.004.
[205] Papastergiou, M. 2009. Digital Game-Based Learning in high school Computer Science education: Impact on educational effectiveness and student motivation. *Computers and Education*. 52, 1 (2009), 1–12. DOI:https://doi.org/10.1016/j.compedu.2008.06.004.
[206] Papavlasopoulou, S., Giannakos, M.N. and Jaccheri, L. 2016. Creative Programing Experiences for Teenagers : Attitudes, Performance and Gender Differences. *IDC Extended Abstracts* (2016).
[207] Papavlasopoulou, S., Giannakos, M.N. and Jaccheri, L. 2019. Exploring children’s learning experience in constructionism-based coding activities through design-based research. *Computers in Human Behavior*. (Jan. 2019). DOI:https://doi.org/10.1016/j.chb.2019.01.008.
[208] Papavlasopoulou, S., Sharma, K., Giannakos, M. and Jaccheri, L. 2017. Using Eye-Tracking to Unveil Differences Between Kids and Teens in Coding Activities. *Proceedings of the 2017 Conference on Interaction Design and Children* (New York, NY, USA, Jun. 2017), 171–181.
[209] Papavlasopoulou, S., Sharma, K. and Giannakos, M.N. 2019. Coding activities for children: Coupling eye-tracking with qualitative data to investigate gender differences. *Computers in Human Behavior*. 7491 (2019), 1–11. DOI:https://doi.org/10.1016/j.chb.2019.03.003.
[210] Papavlasopoulou, S., Sharma, K. and Giannakos, M.N. 2018. How do you feel about learning to code? Investigating the effect of children’s attitudes towards coding using eye-tracking. *International Journal of Child-Computer Interaction*. 17, (2018), 50–60. DOI:https://doi.org/10.1016/j.ijcci.2018.01.004.
[211] Papert, S. 1980. *Mindstorms, Children, Computers and Powerful Ideas*. Basic Books.
[212] Papert, S. 1988. The conservation of Piaget: The computer as a grist to the constructivist mill,. *Constructivism in the Computer Age*. G. Forman and P. Pufall, eds. Lawrence Erlbaum Associates Publishers. 3–13.
[213] Pekrun, R. 2014. *Educational Practices Series 24: Emotion and Learning*.
[214] Pekrun, R. 2006. The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*. 18, 4 (2006), 315–341. DOI:https://doi.org/10.1007/s10648-006-9029-9.
[215] Pekrun, R., Goetz, T., Titz, W. and Perry, R.P. 2002. Academic emotions in students’ self-regulated learning and achievement: A program of qualitative and

quantitative research. *Educational Psychologist*. 37, 2 (2002), 91–105. DOI:https://doi.org/10.1207/S15326985EP3702_4.

[216] Pellas, N. and Mystakidis, S. 2020. A Systematic Review of Research about Game-based Learning in Virtual Worlds. *Journal of Universal Computer Science*. 26, 8 (2020), 1017–1042.

[217] Peters, G.-J.Y. 2014. The alpha and the omega of scale reliability and validity: Why and how to abandon Cronbach's alpha and the route towards more comprehensive assessment of scale quality. *The European Health Psychologist*. 16, 2 (2014), 56–69.

[218] Piaget, J. 1964. Part I: Cognitive development in children: Development and learning. *Journal of Research in Science Teaching*. 2, 3 (Sep. 1964), 176–186. DOI:https://doi.org/10.1002/tea.3660020306.

[219] Piaget, J. 1954. *The Construction of Reality in the Child*. Basic Books.

[220] Pienimäki, M., Kinnula, M. and Iivari, N. 2021. Finding fun in non-formal technology education. *International Journal of Child-Computer Interaction*. 29, (2021), 100283. DOI:https://doi.org/10.1016/j.ijcci.2021.100283.

[221] Pintrich, P.R. 2003. A Motivational Science Perspective on the Role of Student Motivation in Learning and Teaching Contexts. *Journal of Educational Psychology*. 95, 4 (2003), 667–686. DOI:https://doi.org/10.1037/0022-0663.95.4.667.

[222] 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 (1990), 33–40. DOI:https://doi.org/10.1037/0022-0663.82.1.33.

[223] Pintrich, P.R., Smith, D.A.F., Garcia, T. and McKeachie, W.J. 1991. *A Manual for the Use of the Motivated Startegies for Learning Questionnaire (MSLQ)*.

[224] Plass, J.L., Homer, B.D. and Kinzer, C.K. 2015. Foundations of Game-Based Learning. *Educational Psychologist*. 50, 4 (2015), 258–283. DOI:https://doi.org/10.1080/00461520.2015.1122533.

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

[226] Poels, K., de Kort, Y.A.W. and IJsselsteijn, W.A. 2013. The Game Experience Questionnaire. (2013).

[227] Polo, F., Cervai, S. and Kantola, J. 2018. Training culture. *Journal of Workplace Learning*. 30, 3 (2018), 162–173. DOI:https://doi.org/10.1108/jwl-01-2018-0024.

[228] Poole, F.J. and Clarke-Midura, J. 2020. A Systematic Review of Digital Games in Second Language Learning Studies. *International Journal of Game-Based Learning*. 10, 3 (Jul. 2020), 1–15. DOI:https://doi.org/10.4018/IJGBL.2020070101.

[229] Prensky, M. 2001. Fun , Play and Games : What Makes Games Engaging. *Digital Game-Based Learning*. McGraw-Hill. 05-1-05–31.

[230] Prieto, L.P., Sharma, K. and Dillenbourg, P. 2015. Studying Teacher Orchestration Load in Technology-Enhanced Classrooms. G. Conole, T. Klobučar, C. Rensing, J. Konert, and E. Lavoué, eds. Springer International Publishing. 268–281.

[231] Ragosa, D.R. and Willett, J.B. 1983. Demonstrating the reliability of the difference score in the measurement of change. *Journal of Educational Measurement*. 20, 4 (1983), 335–343. DOI:https://doi.org/10.1111/j.1745-3984.1983.tb00211.x.

[232] Rambli, D.R.A., Matcha, W. and Sulaiman, S. 2013. Fun learning with AR alphabet book for preschool children. *Procedia Computer Science*. 25, (2013), 211–

219. DOI:https://doi.org/10.1016/j.procs.2013.11.026.
[233] Read, J.C. 2012. Evaluating Artefacts with Children: Age and Technology Effects in the Reporting of Expected and Experienced Fun. *Icmi 2012*.
[234] Read, J.C. 2008. Validating the Fun Toolkit: an instrument for measuring children's opinions of technology. *Cognition, Technology & Work*. 10, 2 (Apr. 2008), 119–128. DOI:https://doi.org/10.1007/s10111-007-0069-9.
[235] Read, J.C. and MacFarlane, S. 2006. Using the fun toolkit and other survey methods to gather opinions in child computer interaction. *Proceeding of the 2006 conference on Interaction design and children - IDC '06* (2006), 81.
[236] Read, J.C., Macfarlane, S. and Casey, C. 2002. Endurability , Engagement and Expectations : Measuring Children ' s Fun. *Interaction Design and Children*. 2, (2002), 1–23. DOI:https://doi.org/10.1.1.100.9319.
[237] Resnick, M. 2007. All I really need to know (about creative thinking) I learned (by studying how children learn) in kindergarten. *Proceedings of the 6th ACM SIGCHI conference on Creativity & cognition - C&C '07* (New York, New York, USA, 2007), 1–6.
[238] Resnick, M. 1998. Technologies for lifelong kindergarten. *Educational Technology Research and Development*. 46, 4 (1998), 43–55. DOI:https://doi.org/10.1007/BF02299672.
[239] Resnick, M., Kafai, Y.., Maeda, J., Rusk, N. and Maloney, J. 2003. *A networked, media-rich programming envi- ronment to enhance technological fluency at after-school centers in economically-disadvantaged communities. Proposal to the National Science Foundation (project funded 2003–2007).*
[240] Revelle, W. 2018. psych: Procedures for Personality and Psychological Research. Northwestern University.
[241] Revelle, W. and Zinbarg, R.E. 2009. Coefficients Alpha, Beta, Omega, and the glb: Comments on Sijtsma. *Psychometrika*. 74, 1 (Mar. 2009), 145–154. DOI:https://doi.org/10.1007/s11336-008-9102-z.
[242] Richard, G.T., Kafai, Y.B., Adleberg, B. and Telhan, O. 2015. StitchFest: Diversifying a college Hackathon to Broaden participation and perceptions in computing. *SIGCSE 2015 - Proceedings of the 46th ACM Technical Symposium on Computer Science Education*. (2015), 114–119. DOI:https://doi.org/10.1145/2676723.2677310.
[243] Rodríguez-Ardura, I. and Meseguer-Artola, A. 2017. Flow in e-learning: What drives it and why it matters. *British Journal of Educational Technology*. 48, 4 (2017), 899–915. DOI:https://doi.org/10.1111/bjet.12480.
[244] Rodríguez-Hernández, C.F., Cascallar, E. and Kyndt, E. 2020. Socio-economic status and academic performance in higher education: A systematic review. *Educational Research Review*. 29, September 2019 (2020), 100305. DOI:https://doi.org/10.1016/j.edurev.2019.100305.
[245] Rohnke, K. 1993. Fun. *Zip Lines*. 23, 1 (1993), 12–17.
[246] Romero, V. 2014. Children's experiences: Enjoyment and fun as additional encouragement for walking to school. *Journal of Transport and Health*. 2, 2 (2014), 230–237. DOI:https://doi.org/10.1016/j.jth.2015.01.002.
[247] Ronimus, M., Kujala, J., Tolvanen, A. and Lyytinen, H. 2014. Children's engagement during digital game-based learning of reading: The effects of time, rewards, and challenge. *Computers and Education*. 71, (2014), 237–246.

DOI:https://doi.org/10.1016/j.compedu.2013.10.008.
[248] Rosseel, Y. 2012. lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*. 48, 2 (2012), 1–36.
[249] Rossiou, E. and Papadakis, S. 2008. Applying Online Multiplayer Educational Games Based on Generic Shells to Enhance Learning of Recursive Algorithms: Students' Preliminary Results. *European Conference on Games Based Learning* (2008), 373–382.
[250] van Roy, R. and Zaman, B. 2018. Need-supporting gamification in education: An assessment of motivational effects over time. *Computers and Education*. 127, August (2018), 283–297. DOI:https://doi.org/10.1016/j.compedu.2018.08.018.
[251] van Roy, R. and Zaman, B. 2017. Why Gamification Fails in Education and How to Make It Successful: Introducing Nine Gamification Heuristics Based on Self-Determination Theory. *Serious Games and Edutainment Applications*. Springer International Publishing. 485–509.
[252] RStudio Team 2016. RStudio: Integrated Development Environment for R.
[253] Rubio, M.A., Romero-Zaliz, R., Mañoso, C. and De Madrid, A.P. 2015. Closing the gender gap in an introductory programming course. *Computers and Education*. 82, (2015), 409–420. DOI:https://doi.org/10.1016/j.compedu.2014.12.003.
[254] Ruiz-Garcia, A., Subirats, L. and Freire, A. 2016. Lessons Learned in Promoting New Technologies and Engineering in Girls Through a Girls Hackathon and Mentoring. *EDULEARN16 Proceedings*. 1, July (2016), 248–256. DOI:https://doi.org/10.21125/edulearn.2016.1042.
[255] Rushton, E.A.C. and King, H. 2020. Play as a pedagogical vehicle for supporting gender inclusive engagement in informal STEM education. *International Journal of Science Education, Part B: Communication and Public Engagement*. 10, 4 (2020), 376–389. DOI:https://doi.org/10.1080/21548455.2020.1853270.
[256] Ryan, R.M. 1982. Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *Journal of Personality and Social Psychology*. 43, 3 (1982), 450–461. DOI:https://doi.org/10.1037/0022-3514.43.3.450.
[257] Ryan, R.M. and Deci, E.L. 2000. Self-determination theory and the facilitation of intrinsic motivation. *American Psychologist*. 55, 1 (2000), 68–78. DOI:https://doi.org/10.1037/0003-066X.55.1.68.
[258] Ryan, R.M., Mims, V. and Koestner, R. 1983. Relation of reward contingency and interpersonal context to intrinsic motivation: A review and test using cognitive evaluation theory. *Journal of Personality and Social Psychology*. 45, 4 (1983), 736–750. DOI:https://doi.org/10.1037/0022-3514.45.4.736.
[259] Ryan, R.M., Rigby, C.S. and Przybylski, A. 2006. The motivational pull of video games: A self-determination theory approach. *Motivation and Emotion*. 30, 4 (2006), 347–363. DOI:https://doi.org/10.1007/s11031-006-9051-8.
[260] Sáez-López, J.M., Román-González, M. and Vázquez-Cano, E. 2016. Visual programming languages integrated across the curriculum in elementary school: A two year case study using "scratch" in five schools. *Computers and Education*. 97, (2016), 129–141. DOI:https://doi.org/10.1016/j.compedu.2016.03.003.
[261] Sanders, M. 2009. STEM,STEMEducation,STEMmania. *The Technology Teacher*. 20, (2009), 20–27.
[262] Sayette, M.A., Cohn, J.F., Wertz, J.M., Perrott, M.A. and Parrott, D.J. 2001. A psychometric evaluation of the facial action coding system for assessing

spontaneous expression. *Journal of Nonverbal Behavior*. 25, 3 (2001), 167–185.
[263] Schepers, S., Dreessen, K. and Zaman, B. 2018. Fun as a user gain in participatory design processes involving children: A case study. *IDC 2018 - Proceedings of the 2018 ACM Conference on Interaction Design and Children*. (2018), 396–404. DOI:https://doi.org/10.1145/3202185.3202763.
[264] Scherer, R. and Siddiq, F. 2019. The relation between students' socioeconomic status and ICT literacy: Findings from a meta-analysis. *Computers and Education*. 138, 0317 (2019), 13–32. DOI:https://doi.org/10.1016/j.compedu.2019.04.011.
[265] Schunk, D.H. 2012. *Learning Theories - An Educational Perspective*. Pearson.
[266] Schunk, D.H. 1985. Self-efficacy and classroom learning. *Psychology in the Schools*. 22, 2 (Apr. 1985), 208–223. DOI:https://doi.org/10.1002/1520-6807(198504)22:2<208::AID-PITS2310220215>3.0.CO;2-7.
[267] Schunk, D.H. and Ertmer, P.A. 2000. Self-Regulation and Academic Learning. *Handbook of Self-Regulation*. 631–649.
[268] Senkbeil, M., Ihme, J.M. and Wittwer, J. 2013. The Test of Technological and Information Literacy (TILT) in the National Educational Panel Study: Development, Empirical Testing, and Evidence for validity/Test Zur Erfassung Technologischer Und Informationsbezogener Literacy (TILT) Im Nationalen Bildung. *Journal for Educational Research Online / Journal für Bildungsforschung Online*. 5, 2 (2013), 139–161. DOI:https://doi.org/10.25656/01:8428.
[269] Sharma, K. and Giannakos, M. 2020. Multimodal data capabilities for learning: What can multimodal data tell us about learning? *British Journal of Educational Technology*. 51, 5 (Sep. 2020), 1450–1484. DOI:https://doi.org/10.1111/bjet.12993.
[270] Sharma, K., Lee-Cultura, S. and Giannakos, M. 2022. Keep Calm and Don't Carry-Forward: Towards sensor-data driven AI agent to enhance human learning. *Frontiers in Artificial Intelligence*. 198, (2022).
[271] Sharma, K., Niforatos, E., Giannakos, M. and Kostakos, V. 2020. Assessing Cognitive Performance Using Physiological and Facial Features. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*. 4, 3 (2020). DOI:https://doi.org/10.1145/3411811.
[272] Sharma, K., Papamitsiou, Z. and Giannakos, M. 2019. Building pipelines for educational data using AI and multimodal analytics: A "grey-box" approach. *British Journal of Educational Technology*. June (2019), bjet.12854. DOI:https://doi.org/10.1111/bjet.12854.
[273] Sharma, K., Papavlasopoulou, S. and Giannakos, M. 2019. Coding games and robots to enhance computational thinking: How collaboration and engagement moderate children's attitudes? *International Journal of Child-Computer Interaction*. 21, (2019), 65–76. DOI:https://doi.org/10.1016/j.ijcci.2019.04.004.
[274] Shortt, M., Tilak, S., Kuznetcova, I., Martens, B. and Akinkuolie, B. 2021. Gamification in mobile-assisted language learning: a systematic review of Duolingo literature from public release of 2012 to early 2020. *Computer Assisted Language Learning*. 0, 0 (2021), 1–38. DOI:https://doi.org/10.1080/09588221.2021.1933540.
[275] Sim, G., MacFarlane, S. and Read, J.C. 2006. All work and no play: Measuring fun, usability, and learning in software for children. *Computers & Education*. 46, 3 (Apr. 2006), 235–248. DOI:https://doi.org/10.1016/j.compedu.2005.11.021.
[276] Sirin, S.R. 2005. Socioeconomic Status and Academic Achievement: A Meta-

Analytic Review of Research. *Review of Educational Research*. 75, 3 (Sep. 2005), 417–453. DOI:https://doi.org/10.3102/00346543075003417.

[277] Snow, R.E., Corno, L. and Jackson, D.N. 1996. Individual differences in affective and conative functions. *Handbook of Educational Psychology*.

[278] Sousa, D.A. 2011. *How the Brain Learns*. Thousand Oaks.

[279] Spikol, D., Avramides, K. and Cukurova, M. 2016. Exploring the interplay between human and machine annotated multimodal analytics in hands-on STEM activities. *Proceedings of the Sixth International Learning Analytics and Knowledge Conference* (New York, NY, 2016), 522–523.

[280] Spikol, D., Ruffaldi, E., Dabisias, G. and Cukurova, M. 2018. Supervised machine learning in multimodal learning analytics for estimating success in project-based learning. *Journal of Computer Assisted Learning*. 34, 4 (Aug. 2018), 366–377. DOI:https://doi.org/10.1111/jcal.12263.

[281] Stephanidis, C. et al. 2019. Seven HCI Grand Challenges. *International Journal of Human–Computer Interaction*. 35, 14 (Aug. 2019), 1229–1269. DOI:https://doi.org/10.1080/10447318.2019.1619259.

[282] Streiner, D.L. 2005. Finding our way: An introduction to path analysis. *Canadian Journal of Psychiatry*. 50, 2 (2005), 115–122. DOI:https://doi.org/10.1177/070674370505000207.

[283] Su, A.Y.S., Yang, S.J.H., Hwang, W.Y., Huang, C.S.J. and Tern, M.Y. 2014. Investigating the role of computer-supported annotation in problem-solving-based teaching: An empirical study of a Scratch programming pedagogy. *British Journal of Educational Technology*. 45, 4 (2014), 647–665. DOI:https://doi.org/10.1111/bjet.12058.

[284] Subhash, S. and Cudney, E.A. 2018. Gamified learning in higher education: A systematic review of the literature. *Computers in Human Behavior*. 87, February (2018), 192–206. DOI:https://doi.org/10.1016/j.chb.2018.05.028.

[285] Suhr, D. 2006. Exploratory or Confirmatory Factor Analysis ? *Proceedings of the 31st Annual SAS Users Group International Conference* (Cary, NC, 2006), Paper number 200-31.

[286] Sung, H.Y. and Hwang, G.J. 2013. A collaborative game-based learning approach to improving students' learning performance in science courses. *Computers and Education*. 63, (2013), 43–51. DOI:https://doi.org/10.1016/j.compedu.2012.11.019.

[287] Sutton-Smith, B. 2011. *The Ambiguity of Play*. Harward University Press.

[288] Sykes, B. and Musterd, S. 2011. Examining neighbourhood and school effects simultaneously: What does the Dutch evidence show? *Urban Studies*. 48, 7 (2011), 1307–1331. DOI:https://doi.org/10.1177/0042098010371393.

[289] Szegedy, C., Ioffe, S. and Vanhoucke, V. 2016. *Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. arXiv:1602.07261.*

[290] Taber, K.S. 2018. The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. *Research in Science Education*. 48, 6 (2018), 1273–1296. DOI:https://doi.org/10.1007/s11165-016-9602-2.

[291] Taelman, J., Vandeput, S., Spaepen, A. and Van Huffel, S. 2009. Influence of Mental Stress on Heart Rate and Heart Rate Variability. 1366–1369.

[292] Tapia, M. and Marsh, G.E. 2002. Confirmatory Factor Analysis of the Attitudes toward Mathematics Inventory. *The Annual Meeting of the Mid-South Educational Research Association*. (2002), 12.

[293] Tasci, A.D.A. and Ko, Y.J. 2016. A FUN-SCALE for Understanding the Hedonic Value of a Product: The Destination Context. *Journal of Travel and Tourism Marketing*. 33, 2 (2016), 162–183. DOI:https://doi.org/10.1080/10548408.2015.1038421.

[294] Tay, L. and Jebb, A.T. 2017. Scale Development. *The SAGE Encyclopedia of Industrial and Organizational Psychology*. S. Rogelberg, ed. Thousand Oaks.

[295] Tews, M.J., Michel, J.W. and Noe, R.A. 2017. Does fun promote learning? The relationship between fun in the workplace and informal learning. *Journal of Vocational Behavior*. 98, (Feb. 2017), 46–55. DOI:https://doi.org/10.1016/j.jvb.2016.09.006.

[296] Tews, M.J. and Noe, R.A. 2019. Does training have to be fun? A review and conceptual model of the role of fun in workplace training. *Human Resource Management Review*.

[297] Thomas, D.R. and Zumbo, B.D. 2012. Difference scores from the point of view of reliability and repeated-measures ANOVA: In defense of difference scores for data analysis. *Educational and Psychological Measurement*. 72, 1 (2012), 37–43. DOI:https://doi.org/10.1177/0013164411409929.

[298] Tisza, G., Gollerizo, A. and Markopoulos, P. 2019. Measuring fun with adolescents: Introducing the Spanish and Dutch adaptation of the funq. *CHI PLAY 2019 - Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play* (2019), 715–722.

[299] Tisza, G. and Markopoulos, P. 2021. FunQ: Measuring the fun experience of a learning activity with adolescents. *Current Psychology*. (Mar. 2021). DOI:https://doi.org/10.1007/s12144-021-01484-2.

[300] Tisza, G. and Markopoulos, P. 2021. Understanding the role of fun in learning to code. *International Journal of Child-Computer Interaction*. 28, (2021), 100270. DOI:https://doi.org/10.1016/j.ijcci.2021.100270.

[301] Tisza, G., Markopoulos, P. and Bekker, T.M. 2020. Learning to code: Interplay of attitude, emotions and fun. *Manuscript submitted for publication*. (2020).

[302] Tisza, G., Markopoulos, P. and King, H. 2022. Socioeconomic background to influence children's attitude and learning in a creative programming workshop. *Education and Information Technologies*. (2022).

[303] Tisza, G., Papavlasopoulou, S., Christidou, D., Iivari, N., Kinnula, M. and Voulgari, I. 2020. Patterns in informal and non-formal science learning activities for children–A Europe-wide survey study. *International Journal of Child-Computer Interaction*. 25, (2020), 100184. DOI:https://doi.org/10.1016/j.ijcci.2020.100184.

[304] Tisza, G., Papavlasopoulou, S., Christidou, D., Voulgari, I., Iivari, N., Giannakos, M.N., Kinnula, M. and Markopoulos, P. 2019. The role of age and gender on implementing informal and non-formal science learning activities for children. *ACM International Conference Proceeding Series*. (2019). DOI:https://doi.org/10.1145/3335055.3335065.

[305] Tisza, G., Tsiakas, K. and Markopoulos, P. 2022. Exploring the relationship between self-regulation and fun in learning. *Manuscript submitted for publication*. (2022).

[306] Tisza, G., Zhu, S. and Markopoulos, P. 2021. Fun to Enhance Learning, Motivation, Self-efficacy, and Intention to Play in DGBL. *IFIP ICEC 2021*. 28–45.

[307] Tokuhisa, S., Kamiyama, Y. and Tokiwa, T. 2015. Personal, Physical, Social, and

Creative Contextual Design for Art Education. *Computers in Entertainment*. 11, 4 (2015), 1–20. DOI:https://doi.org/10.1145/2582193.2633443.

[308] Torsheim, T., Cavallo, F., Levin, K.A., Schnohr, C., Mazur, J., Niclasen, B. and Currie, C. 2016. Psychometric Validation of the Revised Family Affluence Scale: a Latent Variable Approach. *Child Indicators Research*. 9, 3 (2016), 771–784. DOI:https://doi.org/10.1007/s12187-015-9339-x.

[309] Tsai, T.-W., Lo, H.Y. and Chen, K.-S. 2012. An affective computing approach to develop the game-based adaptive learning material for the elementary students. *Proceedings of the 2012 Joint International Conference on Human-Centered Computer Environments - HCCE '12* (New York, New York, USA, 2012), 8.

[310] Tsiakas, K., Barakova, E., Khan, J.V. and Markopoulos, P. 2020. BrainHood: Designing a cognitive training system that supports self-regulated learning skills in children. *Technology and Disability*. 32, 4 (2020), 219–228. DOI:https://doi.org/10.3233/TAD-200294.

[311] Tsiakas, K., Barakova, E., Khan, J.V. and Markopoulos, P. 2020. BrainHood: Towards an explainable recommendation system for self-regulated cognitive training in children. *ACM International Conference Proceeding Series*. (2020), 521–526. DOI:https://doi.org/10.1145/3389189.3398004.

[312] Tüzün, H., Yılmaz-Soylu, M., Karakuş, T., İnal, Y. and Kızılkaya, G. 2009. The effects of computer games on primary school students' achievement and motivation in geography learning. *Computers & Education*. 52, 1 (Jan. 2009), 68–77. DOI:https://doi.org/10.1016/j.compedu.2008.06.008.

[313] Valiente, C., Swanson, J. and Eisenberg, N. 2012. Linking Students' Emotions and Academic Achievement: When and Why Emotions Matter. *Child Development Perspectives*. 6, 2 (2012), 129–135. DOI:https://doi.org/10.1111/j.1750-8606.2011.00192.x.

[314] Vandevelde, S., Van Keer, H. and Rosseel, Y. 2013. Measuring the complexity of upper primary school children's self-regulated learning: A multi-component approach. *Contemporary Educational Psychology*. 38, 4 (2013), 407–425. DOI:https://doi.org/10.1016/j.cedpsych.2013.09.002.

[315] Vansteenkiste, M., Simons, J., Lens, W., Soenens, B. and Matos, L. 2005. Examining the Motivational Impact of Intrinsic Versus Extrinsic Goal Framing and Autonomy-Supportive Versus Internally Controlling Communication Style on Early Adolescents' Academic Achievement. *Child Development*. 76, 2 (Mar. 2005), 483–501. DOI:https://doi.org/10.1111/j.1467-8624.2005.00858.x.

[316] Verstege, S., Pijeira-Díaz, H.J., Noroozi, O., Biemans, H. and Diederen, J. 2019. Relations between students' perceived levels of self-regulation and their corresponding learning behavior and outcomes in a virtual experiment environment. *Computers in Human Behavior*. 100, September 2018 (2019), 325–334. DOI:https://doi.org/10.1016/j.chb.2019.02.020.

[317] Vieira, L.C. and da Silva, F.S.C. 2017. Assessment of fun in interactive systems: A survey. *Cognitive Systems Research*. 41, (2017), 130–143. DOI:https://doi.org/10.1016/j.cogsys.2016.09.007.

[318] Villavicencio, F.T. and Bernardo, A.B.I. 2013. Positive academic emotions moderate the relationship between self-regulation and academic achievement. *British Journal of Educational Psychology*. 83, 2 (2013), 329–340. DOI:https://doi.org/10.1111/j.2044-8279.2012.02064.x.

[319] Vogel, J.J., Greenwood-Ericksen, A., Cannon-Bowers, J. and Bowers, C.A. 2006. Using virtual reality with and without gaming attributes for academic achievement. *Journal of Research on Technology in Education*. 39, 1 (2006), 105–118. DOI:https://doi.org/10.1080/15391523.2006.10782475.

[320] Vrikki, M., Wheatley, L., Howe, C., Hennessy, S. and Mercer, N. 2019. Dialogic practices in primary school classrooms. *Language and Education*. 33, 1 (Jan. 2019), 85–100. DOI:https://doi.org/10.1080/09500782.2018.1509988.

[321] Vygotsky, L.S. 1967. Play and Its Role in the Mental Development of the Child. *Soviet Psychology*. 5, 3 (1967), 6–18. DOI:https://doi.org/10.2753/rpo1061-040505036.

[322] Wang, M. and Zheng, X. 2020. Using Game-Based Learning to Support Learning Science: A Study with Middle School Students. *The Asia-Pacific Education Researcher*. (Jul. 2020). DOI:https://doi.org/10.1007/s40299-020-00523-z.

[323] Warschauer, M., Knobel, M. and Stone, L. 2004. Technology and Equity in Schooling: Deconstructing the Digital Divide. *Educational Policy*. 18, 4 (Sep. 2004), 562–588. DOI:https://doi.org/10.1177/0895904804266469.

[324] Waterman, A.S. 1993. Two conceptions of happiness: Contrasts of personal expressiveness (eudaimonia) and hedonic enjoyment. *Journal of Personality and Social Psychology*. 64, 4 (Apr. 1993), 678–691. DOI:https://doi.org/10.1037/0022-3514.64.4.678.

[325] Wexler, B.E., Iseli, M., Leon, S., Zaggle, W., Rush, C., Goodman, A., Esat Imal, A. and Bo, E. 2016. Cognitive Priming and Cognitive Training: Immediate and Far Transfer to Academic Skills in Children. *Scientific Reports*. 6, 1 (Dec. 2016), 32859. DOI:https://doi.org/10.1038/srep32859.

[326] White, K.R. 1982. The relation between socioeconomic status and academic achievement. *Psychological Bulletin*. 91, 3 (1982), 461–481. DOI:https://doi.org/10.1037/0033-2909.91.3.461.

[327] Willis, J. 2007. The Neuroscience of Joyful Education. *Educational Leadership*. 64, (2007), 1–4.

[328] Winoto, P., Chen, J., Guo, H. and Tang, T.Y. 2018. A Mathematical and Cognitive Training Application for Children with Autism: A System Prototype. 114–119.

[329] Wolvin, A.D. 1983. Improving Listening Skills. *Improving Speaking and Listening Skills. New Directions for College Learning As- sistance*. R.B. Rubin, ed. Jossey-Bass.

[330] Women in STEM, Women in Computer Science: We're Looking at It Incorrectly: 2014. *https://cacm.acm.org/blogs/blog-cacm/180850-women-in-stem-women-in-computer-science-were-looking-at-it-incorrectly/fulltext*.

[331] Worsley, M. and Blikstein, P. 2018. A Multimodal Analysis of Making. *International Journal of Artificial Intelligence in Education*. 28, 3 (Sep. 2018), 385–419. DOI:https://doi.org/10.1007/s40593-017-0160-1.

[332] Wrzesien, M. and Alcañiz Raya, M. 2010. Learning in serious virtual worlds: Evaluation of learning effectiveness and appeal to students in the E-Junior project. *Computers & Education*. 55, 1 (Aug. 2010), 178–187. DOI:https://doi.org/10.1016/j.compedu.2010.01.003.

[333] Yee, N. 2006. The labor of fun: How video games blur the boundaries of work and play. *Games and Culture*. 1, 1 (2006), 68–71. DOI:https://doi.org/10.1177/1555412005281819.

[334] Yerdelen, S., Kahraman, N. and Taş, Y. 2016. Low socioeconomic status students' STEM career interest in relation to gender, grade level, and stem attitude. *Journal of Turkish Science Education*. 13, Specialissue (2016), 59–74. DOI:https://doi.org/10.12973/tused.10171a.

[335] Yilmaz, E. 2014. Analiysıs of students' success in the exam for transition to futher education through some of the variables. *International Journal of Academic Research*. 6, 1 (Jan. 2014), 57–63. DOI:https://doi.org/10.7813/2075-4124.2014/6-1/B.8.

[336] Yip, F.W.M. and Kwan, A.C.M. 2006. Online vocabulary games as a tool for teaching and learning English vocabulary. *Educational Media International*. 43, 3 (Sep. 2006), 233–249. DOI:https://doi.org/10.1080/09523980600641445.

[337] Yücel, Y. and Rızvanoğlu, K. 2019. Battling gender stereotypes: A user study of a code-learning game, "Code Combat," with middle school children. *Computers in Human Behavior*. 99, May 2018 (2019), 352–365. DOI:https://doi.org/10.1016/j.chb.2019.05.029.

[338] Yusoff, Y.M., Ruthven, I. and Landoni, M. 2011. The fun semantic differential scales. *Proceedings of the 10th International Conference on Interaction Design and Children - IDC '11* (2011), 221–224.

[339] Zaman, B. 2009. Introduction and validation of a pairwise comparison scale for UX evaluations and benchmarking with preschoolers. *User Experience Evaluation Methods in Product Development (UXEM'09)-Workshop*. (2009).

[340] Zaman, B., Abeele, V. Vanden and De Grooff, D. 2013. Measuring product liking in preschool children: An evaluation of the Smileyometer and This or That methods. *International Journal of Child-Computer Interaction*. 1, 2 (May 2013), 61–70. DOI:https://doi.org/10.1016/j.ijcci.2012.12.001.

[341] Zayeni, D., Raynaud, J.-P. and Revet, A. 2020. Therapeutic and Preventive Use of Video Games in Child and Adolescent Psychiatry: A Systematic Review. *Frontiers in Psychiatry*. 11, (Feb. 2020). DOI:https://doi.org/10.3389/fpsyt.2020.00036.

[342] Zhang, F., Markopoulos, P. and Bekker, T. 2020. Children's Emotions in Design-Based Learning: a Systematic Review. *Journal of Science Education and Technology*. 29, 4 (2020), 459–481. DOI:https://doi.org/10.1007/s10956-020-09830-y.

[343] Zhang, F., Markopoulos, P. and Bekker, T. 2018. The Role of Children's Emotions during Design-based Learning Activity - A Case Study at a Dutch High School. *Proceedings of the 10th International Conference on Computer Supported Education* (2018), 198–205.

[344] Zhang, F., Markopoulos, P., Bekker, T., Schüll, M. and Paule-Ruíz, M. 2019. EmoForm. *Proceedings of FabLearn 2019* (New York, NY, USA, Mar. 2019), 18–25.

[345] Zhou, X., Chai, C.S., Jong, M.S.Y. and Xiong, X.B. 2021. Does Relatedness Matter for Online Self-regulated Learning to Promote Perceived Learning Gains and Satisfaction? *Asia-Pacific Education Researcher*. 30, 3 (2021), 205–215. DOI:https://doi.org/10.1007/s40299-021-00579-5.

[346] Zimmerman, B.J. 2000. Chapter 2: Attening self-regulation A social cognitive perspective. *Handbook of Self-Regulation*. (2000), 13–39.

[347] Zimmerman, B.J. 2000. Self-Efficacy: An Essential Motive to Learn. *Contemporary Educational Psychology*. 25, 1 (2000), 82–91. DOI:https://doi.org/10.1006/ceps.1999.1016.

[348] Zimmerman, D.W. and Williams, R.H. 1982. Gain Scores in Research Can Be

Highly Reliable. *Journal of Educational Measurement*. 19, 2 (1982), 149–154. DOI:https://doi.org/10.1111/j.1745-3984.1982.tb00124.x.

[349] Zuckerman, O., Blau, I. and Monroy-Hernández, A. 2009. Children's Participation Patterns in Online Communities: An Analysis of Israeli Learners in the Scratch Online Community. *Interdisciplinary Journal of E-Learning and Learning Objects*. 5, (2009), 263–274.

APPENDIX

APPENDIX A The design of the FunQ (original 50-item version)

Gender: boy / girl Date: ________ Nr: ______
Age: ______ Activity code: ________
(indicated by administrator)

Hi there,

I would like to ask you to evaluate the activity you've just participated in. There are no good or bad answers. Just let me know how did you experience it. If you have any questions, feel free to ask!

Thank you!

Please indicate with an X

	never	rarely	some-times	often	all the time
I want to do the activity again.	o	o	o	o	o
This was an activity that I couldn't do very well.	o	o	o	o	o
I did this activity because I had to.	o	o	o	o	o
I liked the activity.	o	o	o	o	o
The activity was difficult for me.	o	o	o	o	o
I enjoyed doing the activity.	o	o	o	o	o
I did this activity because I had no choice.	o	o	o	o	o
The activity was easy for me.	o	o	o	o	o
I did this activity because I wanted to.	o	o	o	o	o

(GO TO NEXT PAGE)

Gender: boy / girl Date: ________ Nr: ______
Age: ______ Activity code: ________
(indicated by administrator)

During the activity…	never	rarely	some-times	often	all the time
I thought I was good at the activity.	o	o	o	o	o
I talked to others easier than usual.	o	o	o	o	o
I was scared of breaking something.	o	o	o	o	o
I felt smart.	o	o	o	o	o
I laughed a lot.	o	o	o	o	o
I felt that time flew.	o	o	o	o	o
I could make some choices about the activity.	o	o	o	o	o
I felt good.	o	o	o	o	o
I felt challenged.	o	o	o	o	o
I had the sense of controlling the activity.	o	o	o	o	o
I forgot about troubles.	o	o	o	o	o
I was afraid of making mistakes.	o	o	o	o	o
I forgot everything around me.	o	o	o	o	o
I felt angry.	o	o	o	o	o
I was bored.	o	o	o	o	o
I felt like I had to do the activity.	o	o	o	o	o
I knew what to do.	o	o	o	o	o
I made new friends.	o	o	o	o	o
I was happy.	o	o	o	o	o
I was anxious.	o	o	o	o	o
I had a lot of energy.	o	o	o	o	o
I felt irritated.	o	o	o	o	o
I forgot where I was.	o	o	o	o	o

(GO TO NEXT PAGE)

Gender: boy / girl Date: ________ Nr: ______
Age: ______ Activity code: ________
(indicated by administrator)

During the activity…	never	rarely	some-times	often	all the time
I was excited.	o	o	o	o	o
I forgot about school.	o	o	o	o	o
I was afraid of hurting someone.	o	o	o	o	o
I could do what I wanted to.	o	o	o	o	o
I was afraid of damaging something.	o	o	o	o	o
I felt sad.	o	o	o	o	o
I felt it was not my own choice to do the activity.	o	o	o	o	o
I did something I'd never done before.	o	o	o	o	o
I had fun.	o	o	o	o	o
I talked with others to whom I had never before.	o	o	o	o	o
I was only thinking about the activity.	o	o	o	o	o
I felt bad.	o	o	o	o	o
I felt closer to others more than usual.	o	o	o	o	o
I had to concentrate hard.	o	o	o	o	o
I forgot about my daily routine.	o	o	o	o	o
I did something new.	o	o	o	o	o
I was curious.	o	o	o	o	o
I forgot about homework.	o	o	o	o	o

THANK YOU! ☺

APPENDIX B The FunQ questionnaire sorted by the factorial structure and with standardized factor loadings of the concerning factors of the final 18-item model on the second data set. Gray font color indicates the initial items that were removed for the final version.

Nr.	Fact.	Init. label	Item	β_{std}	p
1	Autonomy (ω = 0.653)	A1	During the activity, I had the sense of controlling the activity.		
2		A2	During the activity, I knew what to do.	0.329	-
3		A3	During the activity, I could do what I wanted.		
4		A4	During the activity, I could make some choices about the activity.		
5		P1 †	During the activity, I felt it was not my own choice to do the activity. (R) ‡		
6		P2	During the activity, I felt like I had to do the activity. (R)		
7		P3	I did this activity because I had no choice. (R)		
8		P4	I did this activity because I had to. (R)	0.557	0.009
9		P5	I did this activity because I wanted to.	0.891	0.003
10	*Challenge* (ω = 0.990)	C1	The activity was easy for me.		
11		C2	During the activity, I felt I was good at this activity.	0.426	-
12		C3	During the activity, I had to concentrate hard.		
13		C4	The activity was difficult for me. (R)		
14		C5	During the activity, I felt challenged.		
15		C7	During the activity, I did something new.	0.523	0.036
16		C8	During the activity, I did something I'd never done before.		
17		C9	During the activity, I was curious.	0.394	0.017
18		C10	During the activity, I felt smart.		
19	*Delight* (ω = 0.993)	E1	During the activity, I had fun.	0.801	-
20		E2	I liked the activity.		
21		E3	I enjoyed doing the activity.		
22		E4	I want to do something like this again.	0.569	< 0.001
23		E5	During the activity, I laughed a lot.		

24		E6	During the activity, I was happy.	0.771	< 0.001
25		E7	During the activity, I had a lot of energy.		
26		E8	During the activity I was excited.		
28		C6	During the activity, I was bored. (R)		
29	*Fear of Damage*	D1	During the activity, I was scared of breaking something. (R)		
30		D2	During the activity, I was afraid of damaging something. (R)		
31		D3	During the activity, I was afraid of making mistakes. (R)		
32		D4	During the activity, I was afraid of hurting someone. (R)		
33	*Immersion* (ω = 0.885)	I1	During the activity, I was only thinking about the activity.		
34		I2	During the activity, I forgot everything around me.		
35		I3	During the activity, I felt that time flew.	0.645	-
36		I4	During the activity, I forgot where I was.		
37		I5	During the activity, I forgot about school.	0.357	< 0.001
38		I6	During the activity, I forgot about homework.		
39		I7	During the activity, I forgot about troubles.		
40		I8	During the activity, I forgot about my daily routine.		
27		D9	During the activity, I felt good.	0.880	< 0.001
41	*Loss of Social Barriers* (ω = 0.633)	SB1	During the activity, I made new friends.	0.386	-
42		SB2	During the activity, I talked to others easier than usual.	0.689	0.014
43		SB3	During the activity, I felt closer to others more than usual.	0.674	0.001
44		SB4	During the activity, I talked with others to whom I had never before.		
45	*Stress* (ω = 0.791)	S1	During the activity, I felt bad. (R)	0.838	-
46		S2	During the activity, I felt irritated. (R)		
47		S3	During the activity, I felt angry. (R)	0.712	< 0.001
48		S4	During the activity, I felt sad. (R)	0.700	< 0.001
49		S5	During the activity, I was anxious. (R)		
50		S6	This was an activity that I couldn't do very well. (R)		

† Label P indicates the initial factor Pressure
‡ (R) indicates a reverse statement

APPENDIX C Descriptive statistics and test of normalify of FunQ items on the second data set (N=128)

Item	Mean†	Standard deviation	Skewness	Kurtosis	Shapiro-Wilk test of normality
During the activity...					
I knew what to do	3.94	1.091	-0.705	-0.371	0.000
I felt I was good at this activity.	3.96	0.991	-0.660	-0.172	0.000
I did something new	3.86	1.293	-0.909	-0.249	0.000
I was curious.	3.54	1.407	-0.602	-0.861	0.000
I had fun.	4.22	1.023	-1.107	0.346	0.000
I was happy.	4.02	0.963	-0.469	-0.976	0.000
I felt that time flew.	3.64	1.329	-0.604	-0.729	0.000
I forgot about school.	3.09	1.635	-0.077	-1.606	0.000
I felt good	4.04	1.053	-1.066	0.769	0.000
I made new friends.	1.72	1.147	1.570	1.468	0.000
I talked to others easier than usual.	3.12	1.462	-0.085	-1.358	0.000
I felt closer to others more than usual.	2.44	1.384	0.389	-1.207	0.000
I felt bad. (R)	1.39	0.858	2.401	5.472	0.000
I felt angry. (R)	1.30	0.772	3.021	9.528	0.000
I felt sad. (R)	1.50	1.005	2.193	4.281	0.000
I did this activity because I had to. (R)	2.42	1.347	0.511	-0.966	0.000
I did this activity because I wanted to.	3.66	1.250	-0.594	-0.566	0.000
I want to do something like this again.	3.70	0.929	-0.212	-0.258	0.000

† 1 = never, 5 = all the time

APPENDIX D The Dutch FunQ

Item	Factor
Gedurende de activiteit...	
wist ik wat ik moest doen.	Autonomy
voelde ik dat ik er goed in was.	Challenge
deed ik iets nieuws.	Challenge
was ik nieuwsgierig.	Challenge
had ik het leuk.	Delight
was ik gelukkig.	Delight
vlog de tijd voorbij.	Immersion
vergat ik school.	Immersion
voelde ik me goed.	Immersion
heb ik nieuwe vrienden gemaakt.	Loss of Social Barriers
praatte ik gemakkelijker met anderen dan normaal.	Loss of Social Barriers
voelde ik me meer verbonden met anderen dan normaal.	Loss of Social Barriers
voelde ik me slecht. (R)	Stress
voelde ik me boos. (R)	Stress
voelde ik me geïrriteerd. (R)	Stress
Ik deed deze activiteit omdat het moest. (R)	Autonomy
Ik deed deze activiteit omdat ik het wilde.	Autonomy
Ik wil zoeits als dit nog wel een keer doen.	Delight

APPENDIX E The knowledge assessment test

1. What is a micro:bit?

a) a little piece of snack
b) a microcomputer
c) a natural phenomena
d) a microwave oven brand

2. Which device does have a microcomputer? (mark all that apply)

a) mobile phone
b) ballpoint pen
c) coffee machine
d) earphones

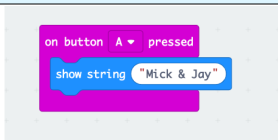
3. What/Who is an editor?

a) a program or app with which you can make digital things
b) a person who writes stories
c) the scientific nick-name for a famous researcher
d) a falling rock in the Earth's atmosphere from the space

4. What/Who is a variable?

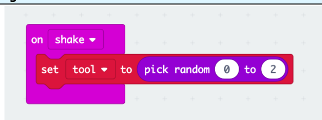
a) someone who cannot decide what to do
b) an atomic molecule
c) a value or information that is saved in the memory of the computer
d) a French cookie stuffed with vanilla pudding

5. What does the following code do?



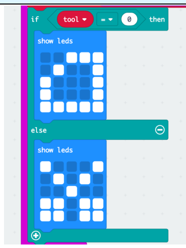
a) it shows the blue string of Mick & Jay
b) Mick and Jay will be written on button A
c) if button A is pressed Mick & Jay is shown
d) a blue picture will be shown with purple background

6. What does the following code do?



a) when purple the code picks a random tool
b) when the code is red and is between 0 and 2, then it sets the tool purple
c) when shaken, the code choses a random number from 0 to 2 and saves it in the tool
d) the code makes the tool shake

7. What does the following code do?



a) when the background is blue it can be both figures
b) when the tool is red it shows the upper figure, otherwise the bottom figure
c) when the tool is 0 it shows the upper figure, otherwise it shows the bottom figure
d) when the green block appear one can chose which figure to see

APPENDIX F Descriptive statistics of the knowledge assessment test

Table F1 Pre-workshop. The correct answer is indicated with bold typesetting.

	Response frequency				
item	**a**	**b**	**c**	**d**	**missing**
1	0	**87.0%**	4.3%	0	8.7%
3	**69.6%**	8.7%	17.4%	0	4.3%
4	13.0%	13.0%	**60.9%**	8.7%	4.3%
5	21.7%	0	**52.2%**	4.3%	21.7%
6	8.7%	13.0%	**60.9%**	0	17.4%
7	4.3%	17.4%	**56.5%**	4.3%	17.4%

	Response frequency						
item	**a**	**b**	**c**	**d**	**ac**	**ad**	**acd**	**missing**
2	34.8%	0	4.3%	4.3%	**21.7%**	4.3%	26.1%	4.3%

Table F2 Post-workshop. The correct answer is indicated with bold typesetting.

	Response frequency				
item	**a**	**b**	**c**	**d**	**missing**
1	0	**95.7%**	0	0	4.3%
3	**69.6%**	4.3%	13.0%	4.3%	8.7%
4	4.3%	13.0%	**60.9%**	0	21.7%
5	13.0%	17.4%	**65.2%**	0	4.3%
6	17.4%	4.3%	**65.2%**	8.7%	4.3%
7	4.3%	26.1%	**65.2%**	0	4.3%

	Response frequency					
item	**a**	**b**	**c**	**d**	**ac**	**missing**
2	8.7%	4.3	4.3%	4.3%	**73.9%**	4.3%

APPENDIX G Descriptive statistics of FunQ items (N=86)

Item	Mean†	SD	Factor	Standardized loading
During the activity…				
I knew what to do	4.22	0.918	Autonomy	0.342
I felt I was good at this activity.	3.97	0.960	Challenge	0.393
I did something new	4.00	1.184	Challenge	0.277
I was curious.	3.48	1.447	Challenge	0.284
I had fun.	4.56	0.859	Delight	0.298
I was happy.	4.51	0.714	Delight	0.445
I felt that time flew.	4.32	1.141	Immersion	0.333
I forgot about school.	2.82	1.615	Immersion	0.271
I felt good	4.59	0.689	Immersion	0.704
I made new friends.	1.82	1.412	Loss of Social Barriers	0.597
I talked to others easier than usual.	3.19	1.406	Loss of Social Barriers	0.362
I felt closer to others more than usual.	2.33	1.408	Loss of Social Barriers	0.334
I felt bad. (R)	1.27	0.812	Stress	0.502
I felt angry. (R)	1.14	0.476	Stress	0.842
I felt sad. (R)	1.14	0.687	Stress	0.554
I did this activity because I had to. (R)	2.14	1.402	Autonomy	0.464
I did this activity because I wanted to.	4.04	1.178	Autonomy	0.885
I want to do something like this again.	4.25	0.808	Delight	0.673

† 1 = never, 5 = all the time

APPENDIX H All statistical results regarding the pre- and post-workshop attitude items and the effect of school and gender.

Table H1 All statistical results of the effect of school on the pre-workshop attitude items.

Attitude dimension	F (df)	p	partial η^2
Boring - fun	3.803(2)	0.025	0.055
Difficult to do - Easy to do	1.336(2)	0.267	0.020
Difficult to understand - Easy to understand	5.217(2)	0.007	0.074
Unpleasant - Pleasant	4.292(2)	0.016	0.062
Uninteresting - Exciting	6.233(2)	0.003	0.088
I don't want to do - I want to do	2.232(2)	0.111	0.033

Table H2 All statistical results of the effect of gender on the pre-workshop attitude items.

Attitude dimension	F (df)	p	partial η^2
Boring - fun	0.647(1)	0.423	0.005
Difficult to do - Easy to do	0.500(1)	0.481	0.004
Difficult to understand - Easy to understand	0.527(1)	0.469	0.004
Unpleasant - Pleasant	0.000(1)	0.991	< 0.000
Uninteresting - Exciting	0.016(1)	0.899	< 0.000
I don't want to do - I want to do	0.597(1)	0.441	0.005

Table H3 All statistical results of the effect of school on the post-workshop attitude items.

Attitude dimension	F (df)	p	partial η^2
Boring - fun	3.342(2)	0.038	0.049
Difficult to do - Easy to do	11.094(2)	< 0.000	0.147
Difficult to understand - Easy to understand	7.132(2)	0.001	0.100
Unpleasant - Pleasant	4.308(2)	0.015	0.063
Uninteresting - Exciting	3.804(2)	0.025	0.056
I don't want to do - I want to do	2.411(2)	0.094	0.036
I think that programming is my thing	3.853(2)	0.024	0.056

Table H4 All statistical results of the effect of gender on the post-workshop attitude items.

Attitude dimension	F (df)	p	partial η^2
Boring - fun	0.529(1)	0.468	0.004
Difficult to do - Easy to do	1.849(1)	0.176	0.014
Difficult to understand - Easy to understand	1.807(1)	0.181	0.014
Unpleasant - Pleasant	0.062(1)	0.804	0.000
Uninteresting - Exciting	0.875(1)	0.351	0.007
I don't want to do - I want to do	1.945(1)	0.166	0.015
I think that programming is my thing	0.023(1)	0.880	0.000

APPENDIX I Full model details for predicting RLG and FunQ dimensions.

Table I1 Detailed results for the model predicting the RLG using affective states, the transitions among them and the physiological measurements. The boldface coefficients are significant. The predictor model accounts for 72% variance in students' RLG

Variable	β	Error	T-value	P-value
Intercept	0.27	0.69	0.52	> 0.05
Happiness (hap)	0.59	0.88	0.49	> 0.05
Sadness (sad)	**-1.37**	**0.009**	**-2.38**	**0.01**
Anger (ang)	**-1.56**	**0.003**	**-3.12**	**0.001**
Surprise (sup)	0.42	0.94	0.36	> 0.05
Trans:hap <-> sad	0.64	0.74	0.23	> 0.05
Trans:hap <-> ang	-0.48	0.84	-0.52	> 0.05
Trans:hap <-> sup	**1.13**	**0.04**	**2.03**	**0.02**
Trans:sad <-> ang	**-1.88**	**0.004**	**-3.42**	**0.0006**
Trans:sad <-> sup	0.11	0.88	0.16	> 0.05
Trans:ang <-> sup	0.08	0.63	0.12	> 0.05
Stress	**-1.03**	**0.01**	**-1.78**	**0.04**
Arousal	**1.19**	**0.02**	**1.71**	**0.04**

Table I2 Model for the FunQ Total Score using the affective states, the transitions among them and the physiological measurements. The boldface coefficients are significant. The predictor model accounts for 26% variance in students' FunQ Total score.

	β	Error	T-value	P-value
Intercept	0.57	0.56	0.75	> 0.05
Happiness (hap)	0.58	0.90	0.12	> 0.05
Sadness (sad)	-0.79	0.57	-0.61	> 0.05
Anger (ang)	-0.67	0.49	-0.46	> 0.05
Surprise (sup)	0.77	0.43	0.75	> 0.05
Trans:hap <-> sad	0.65	0.70	0.69	> 0.05
Trans:hap <-> ang	0.81	1.01	0.48	> 0.05
Trans:hap <-> sup	0.72	0.92	0.33	> 0.05
Trans:sad <-> ang	-0.71	0.46	-0.32	> 0.05
Trans:sad <-> sup	-0.58	0.68	-0.39	> 0.05
Trans:ang <-> sup	0.63	0.46	0.47	> 0.05
Stress	-0.51	0.78	-0.40	> 0.05
Arousal	0.70	0.84	0.37	> 0.05

Table I3 Model for the FunQ Autonomy using the affective states, the transitions among them and the physiological measurements. The boldface coefficients are significant. The predictor model accounts for 32% variance in students' FunQ Autonomy

	β	**Error**	**T-value**	**P-value**
Intercept	0.04	0.02	0.34	> 0.05
Happiness (hap)	-0.24	0.17	-0.82	> 0.05
Sadness (sad)	-0.72	0.58	-0.74	> 0.05
Anger (ang)	0.88	0.75	0.35	> 0.05
Surprise (sup)	0.96	0.45	0.34	> 0.05
Trans:hap <-> sad	0.44	0.35	0.06	> 0.05
Trans:hap <-> ang	0.89	0.80	0.06	> 0.05
Trans:hap <-> sup	0.29	0.18	0.61	> 0.05
Trans:sad <-> ang	-0.35	0.34	-0.30	> 0.05
Trans:sad <-> sup	-0.79	0.64	-0.37	> 0.05
Trans:ang <-> sup	0.95	0.87	0.99	> 0.05
Stress	0.28	0.32	0.73	> 0.05
Arousal	0.10	0.21	0.17	> 0.05

Table I4 Model for the FunQ Challenge using the affective states, the transitions among them and the physiological measurements. The boldface coefficients are significant. The predictor model accounts for 71% variance in students' FunQ Challenge.

	β	Error	T-value	P-value
Intercept	0.31	0.12	0.43	> 0.05
Happiness (hap)	**1.41**	**0.005**	**2.92**	**0.002**
Sadness (sad)	**-1.32**	**0.012**	**-2.10**	**0.02**
Anger (ang)	**1.02**	**0.004**	**3.13**	**0.001**
Surprise (sup)	**-1.35**	**0.003**	**-2.44**	**0.009**
Trans:hap <-> sad	**0.99**	**0.001**	**3.38**	**0.0007**
Trans:hap <-> ang	0.31	0.30	0.35	> 0.05
Trans:hap <-> sup	0.42	0.33	0.33	> 0.05
Trans:sad <-> ang	0.47	0.43	0.48	> 0.05
Trans:sad <-> sup	**-1.45**	**0.001**	**-3.04**	**0.001**
Trans:ang <-> sup	0.33	0.46	0.39	> 0.05
Stress	0.32	0.33	0.34	> 0.05
Arousal	**1.23**	**0.014**	**2.23**	**0.01**

Table I5 Model for the FunQ Delight using the affective states, the transitions among them and the physiological measurements. The boldface coefficients are significant. The predictor model accounts for 65% variance in students' FunQ Delight.

	β	**Error**	**T-value**	**P-value**
Intercept	0.44	0.24	0.54	> 0.05
Happiness (hap)	**1.34**	**0.051**	**2.13**	**0.02**
Sadness (sad)	-0.42	0.11	-0.97	> 0.05
Anger (ang)	-0.33	0.09	-1.44	> 0.05
Surprise (sup)	**1.93**	**0.009**	**3.19**	**0.001**
Trans:hap <-> sad	-0.39	0.42	-0.66	> 0.05
Trans:hap <-> ang	**-0.93**	**0.001**	**-4.18**	**0.00006**
Trans:hap <-> sup	0.53	1.12	0.20	> 0.05
Trans:sad <-> ang	-0.99	0.12	-1.35	> 0.05
Trans:sad <-> sup	0.33	0.23	0.44	> 0.05
Trans:ang <-> sup	-0.59	0.62	-0.43	> 0.05
Stress	0.43	0.91	0.12	> 0.05
Arousal	**-1.34**	**0.003**	**-3.53**	**0.0004**

Table I6 Model for the FunQ Immersion using the affective states, the transitions among them and the physiological measurements. The boldface coefficients are significant. The predictor model accounts for 63% variance in students' FunQ Immersion.

	β	**Error**	**T-value**	**P-value**
Intercept	0.19	0.89	0.52	> 0.05
Happiness (hap)	**1.46**	**0.001**	**4.34**	**0.00003**
Sadness (sad)	0.22	0.45	0.38	> 0.05
Anger (ang)	-0.62	0.26	-0.34	> 0.05
Surprise (sup)	0.44	0.61	0.22	> 0.05
Trans:hap <-> sad	0.27	0.77	0.40	> 0.05
Trans:hap <-> ang	-0.83	0.49	-0.36	> 0.05
Trans:hap <-> sup	0.32	0.59	0.22	> 0.05
Trans:sad <-> ang	**-1.73**	**0.003**	**-3.28**	**0.009**
Trans:sad <-> sup	0.31	0.82	0.37	> 0.05
Trans:ang <-> sup	0.62	0.77	0.20	> 0.05
Stress	0.83	0.65	0.20	> 0.05
Arousal	**2.01**	**0.001**	**4.26**	**0.00005**

Table I7 Model for the FunQ Social Barrier using the affective states, the transitions among them and the physiological measurements. The boldface coefficients are significant. The predictor model accounts for 17% variance in students' FunQ Social barrier.

	β	**Error**	**T-value**	**P-value**
Intercept	0.47	0.66	0.40	> 0.05
Happiness (hap)	0.57	0.48	0.45	> 0.05
Sadness (sad)	-0.77	0.67	-0.33	> 0.05
Anger (ang)	-0.80	0.49	-0.51	> 0.05
Surprise (sup)	0.50	0.92	0.30	> 0.05
Trans:hap <-> sad	-0.55	0.90	-0.56	> 0.05
Trans:hap <-> ang	-0.47	0.73	-0.57	> 0.05
Trans:hap <-> sup	0.44	0.54	0.48	> 0.05
Trans:sad <-> ang	-0.77	0.91	-0.44	> 0.05
Trans:sad <-> sup	-0.39	0.69	-0.56	> 0.05
Trans:ang <-> sup	-0.81	0.69	-0.55	> 0.05
Stress	0.62	0.90	0.47	> 0.05
Arousal	0.59	0.91	0.31	> 0.05

Table I8 Model for the FunQ Stress using the affective states, the transitions among them and the physiological measurements. The boldface coefficients are significant. The predictor model accounts for 70% variance in students' FunQ Stress.

	β	Error	T-value	P-value
Intercept	0.16	0.21	0.89	> 0.05
Happiness (hap)	-0.70	0.78	-0.25	> 0.05
Sadness (sad)	**1.28**	**0.005**	**2.48**	**0.008**
Anger (ang)	**0.94**	**0.001**	**3.32**	**0.0009**
Surprise (sup)	-0.87	0.86	-0.21	> 0.05
Trans:hap <-> sad	0.58	0.72	0.21	> 0.05
Trans:hap <-> ang	0.63	0.83	0.34	> 0.05
Trans:hap <-> sup	**-0.89**	**0.006**	**3.29**	**0.0009**
Trans:sad <-> ang	**1.27**	**0.017**	**2.70**	**0.005**
Trans:sad <-> sup	-0.88	0.74	-0.26	> 0.05
Trans:ang <-> sup	0.59	0.82	0.34	> 0.05
Stress	**1.39**	**0.001**	**3.88**	**0.0001**
Arousal	0.70	0.62	0.32	> 0.05

ACKNOWLEDGEMENTS

I would like to express my heartfelt gratitude to all of the people whom with I was lucky to share my doctoral research journey.

First and foremost, I would like to express my eternal gratitude to my supervisor, **Panos Makopoulos**. I cannot even put into words how grateful I am for you for being there. You have been the best 'boss' I ever had, and I have learnt a lot from you not only as an academic and as a researcher, but as a person as well. You always provided me with the accurate amount of challenge and with great ideas for research, from which I often could not choose so I persuaded both. Thank you for leading me in this journey, of which I have enjoyed every minutes truly. I hope that our ways will not diverge for a while.

My sincerer thanks go to my supervisor, **Tilde Bekker**, with whom I did not have a daily contact, but when we met, we had inspiring conversations and I gained new perspectives on my research. Thank you for sharing with me your valuable insights, and for helping me out when there was little hope left for conducting my study on time.

Esteemed committee members **Renate de Groot**, **Perry den Brok**, **Ruud Taconis**, **Bieke Zaman** and **Janet Read**: I am grateful to you for taking the time to read this work and providing insightful remarks and points for further discussion.

Many thanks to my collaborators from the **Com n Play Science** project, whom with I was dropped in deep waters at the very first day of my PhD, but who welcomed me as a family member. Special thanks to the NTNU team: **Sofia Papavlasopoulou**, with whom I not only collaborated on many occasions during the past years, but who, as a fellow PhD candidate, has been a role model for me. I also wish to thank **Kshitij Sharma** and **Michalis Giannakos** for all your support, for the many fruitful, inspiring conversations, and for sharing your expertise and ways of thinking with me.

I also thank all the schools that have been involved in my research. In particular, I wish to thank all the students and teachers who have been willing to participate in my research. Special thanks for **Robert Stewart**, **Corinne Kemperman**, **Annemarie Haket**, and **Onno Tuijten** who made their colleagues curious, ensured smooth visits, and prepared the students with great enthusiasm on several occasions.

This research would have been very different if it was not for the open access video materials on teaching programming for children and teenagers in a playful manner. Therefore, my heartfelt thanks go to **Jeroen Storm** for the brainstorming sessions on the optimal use of his videos, and for all the provided support for my research.

I would like to thank my TU/e colleagues as well. My fellow PhD candidates, especially **Baisong**, **Bruna, Debargha**, **Feiran**, **Huan**, **Jesus**, **Jingrui**, **Kristin**, **Maggie**, **Ruud**, and **Svetlana**: thank you for the stimulating conversations, encouragements, and fun time we spent together. My heartfelt thanks go to **Rosalinde Kennis** for being such a fantastic person, for answering all my questions and queries from the very beginning, and for all the emotional support you tirelessly provided. **Melanie Swallow**, despite you joined the team a bit later, with your positivity and radiating happiness you quickly became an essential team member and cheered us all up no matter what.

This section would not be complete without thanking all amazing people in my personal life. I am thankful for my parents, **Ferenc Farkas** and **Marianna Megyesi**, who injected curiosity towards science into me from early ages on and brought me up in a spirit of 'question everything'. My heartfelt gratitude also goes to my husband, **Ádám Bálint Tisza**, whom with we started our research journey together long ago by diving into methodological questions, which later became one of my main interests. Despite he does not work for academia, he remained my main supporter to persuade my dream of becoming an academic researcher. I am thankful for my daughters, **Lilien** and **Anabell Tisza**, who both arrived at the best time possible, and thought me a lot of things, among others, to focus even better, and to perfect my time-management skills. They both joined me for data collections, meetings, and conferences, and made the writing up of my research results and ultimately, this thesis a less lonely endeavor, first in my belly, then wrapped around me in the carrier. Köszönöm mindannyiótoknak, hogy vagytok nekem!

I am thankful for the presence of **Tania Kotiuk**, whom I learnt to know under special circumstances just before corona, and who has since become one of my best friends.

Besides the ones I have already mentioned, there are many other valuable individuals who has directly or indirectly affected me during my journey. Since a few pages cannot do justice to all people, know that you have my gratitude. I hope that our ways will cross each other again. My best to all of you.

Gabriella

BIOGRAPHY

Gabriella Tisza (maiden name: Gabriella Farkas) was born on the 18th of December 1988 in Budapest, Hungary. After obtaining her master's degree in 2013 in Education and Counseling Psychology from the University of Szeged, Hungary, she worked for a few years in education. In 2018 she obtained a master's degree in Methodology and Statistics from the University of Leiden, the Netherlands. From June 2018 she started a Ph.D. project in the Department of Industrial Design at Eindhoven University of Technology, under the guidance of Prof. Dr. Panos Markopoulos and Prof. Dr. Tilde Bekker, the results of which are presented in this dissertation. During her Ph.D. years, she has become a mother of two daughters with her husband, Adam: Lilien was born on the 1st of July in 2020, and Anabell was born on the 3rd of October in 2022.

LIST OF PUBLICATIONS

Presented in the dissertation

Tisza, G., & Markopoulos, P. (2021a). FunQ: Measuring the fun experience of a learning activity with adolescents. *Current Psychology*. https://doi.org/10.1007/s12144-021-01484-2

Tisza, G., & Markopoulos, P. (2021b). Understanding the role of fun in learning to code. *International Journal of Child-Computer Interaction*, *28*, 100270. https://doi.org/10.1016/j.ijcci.2021.100270

Tisza, G., Zhu, S., & Markopoulos, P. (2021). Fun to Enhance Learning, Motivation, Self-efficacy, and Intention to Play in DGBL. In *IFIP ICEC 2021* (pp. 28–45). https://doi.org/10.1007/978-3-030-89394-1_3

Tisza, G., Sharma, K., Papavlasopoulou, S., Markopoulos, P., & Giannakos, M. (2022). Understanding fun in learning to code: a multi-modal data approach. *IDC '22: Interaction Design and Children*, June 2022, p. 274-284. https://doi.org/10.1145/3501712.3529716

Tisza, G., Markopoulos, P., & King, H. (2022). Socioeconomic background influences children's attitude and learning in creative coding workshop. *Education and Information Technologies*. https://doi.org/10.1007/s10639-022-11467-w

Tisza, G., Markopoulos, P., & Bekker, T. M. (2020). Learning to code: Interplay of attitude, emotions and fun. *Manuscript submitted for publication*.

Tisza, G., Tsiakas, K., & Markopoulos, P. (2022). Exploring the relationship between self-regulation and fun in learning. *Manuscript submitted for publication*.

Tisza, G., & Markopoulos, P. (2022). Path analysis to understand better how self-regulation, socioeconomic background, and having fun while learning influence students' attitude towards programming and their learning outcomes. *Manuscript submitted for publication*

Further publications

Tisza, G., Papavlasopoulou, S., Christidou, D., Voulgari, I., Iivari, N., Giannakos, M. N., Kinnula, M., & Markopoulos, P. (2019). The role of age and gender on implementing informal and non-formal science learning activities for children. *ACM International Conference Proceeding Series*. https://doi.org/10.1145/3335055.3335065

Tisza, G., Gollerizo, A., & Markopoulos, P. (2019). Measuring fun with Adolescents: Introducing the Spanish and Dutch Adaptation of the FunQ. *CHI PLAY EA '19*. https://doi.org/https://doi.org/10.1145/3341215.3356265

Tisza, G., Papavlasopoulou, S., Christidou, D., Iivari, N., Kinnula, M. & Voulgari, I. (2020). Patterns in informal and non-formal science learning activities for children: A Europe-wide survey study. *International Journal of Child-Computer Interaction*. 25. https://doi.org/10.1016/j.ijcci.2020.100184

Tisza, G. (2021). Assessing the experienced fun with FunQ. 13 Oct 2021, *2021 IEEE/ACIS 21st International Fall Conference on Computer and Information Science (ICIS 2021–Fall)*. IEEE Computer Society, p. 144-148

Christidou, D., Voulgari, I., Tisza, G., Norouzi, B., Kinnula, M., Iivari, N., Papavlasopoulou,

S., Gollerizo, A., Lozano González, J.M., & Konstantinidi Sofrona, D. (2022) Obstacles and challenges identified by practitioners of non-formal science learning activities in Europe, *International Journal of Science Education*, 44 (3), 514-533, https://doi.org/10.1080/09500693.2022.203546

Editor of a Journal Special Issue

Tisza, G., Papavlasopoulou, S., & Sim, G. R. (Eds.). (2023). Fun in learning [Special Issue]. International Journal of Child-Computer Interaction. *Special Issue in preparation.*

Best paper awards

Tisza, G. (2021). Assessing the experienced fun with FunQ. 13 Oct 2021, *2021 IEEE/ACIS 21st International Fall Conference on Computer and Information Science (ICIS 2021–Fall).* IEEE Computer Society, p. 144-148

Tisza, G., Sharma, K., Papavlasopoulou, S., Markopoulos, P., & Giannakos, M. (2022). Understanding fun in learning to code: a multi-modal data approach. *IDC '22: Interaction Design and Children,* June 2022, p. 274-284, https://doi.org/10.1145/3501712.3529716

Best student paper award

Tisza, G. (2021). Assessing the experienced fun with FunQ. 13 Oct 2021, *2021 IEEE/ACIS 21st International Fall Conference on Computer and Information Science (ICIS 2021–Fall).* IEEE Computer Society, p. 144-148

Workshop

Tisza, G., Sim, G., Grammenos, D., & Papavlasopoulou, S., (2022). Fun in learning. IDC '22: Interaction Design and Children, June 2022, p. 709-711. https://doi.org/10.1145/3501712.3536385