Computers in Human Behavior 105 (2020) 105939

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Full length article

Coding activities for children: Coupling eye-tracking with qualitative data to investigate gender differences

Sofia Papavlasopoulou, Kshitij Sharma, Michail N. Giannakos*

Department of Computer Science (IDI), Norwegian University of Science and Technology (NTNU), Trondheim, Norway

ARTICLE INFO

Keywords:
Coding
Computational thinking
Eye-tracking
Gender differences
Learning strategies

ABSTRACT

Computational thinking and coding are becoming an integral part of K-12 education, with female students being underrepresented in such subjects. The proliferation of technological tools and programming environments offers the opportunity for creative coding activities for children and increases the need for appropriate instructional practices. In this study, we design and evaluate a coding workshop for children. Our goal is to examine differences between boys and girls using eye-tracking as an objective measure and triangulating the findings with qualitative data coming from children's interviews. The results show no statistically significant difference between female and male gaze and learning gain during the coding activity; interestingly, the qualitative data show differences in the strategies and implemented practices during coding, and in perceptions about those coding activities. Our results highlight that further studies need to utilize objective measures and unveil necessary differences in the design and implementation of coding activities. Furthermore, our results provide objective evidence that female students do not lack in competences compared to boys, but simply that they have a different approach during coding activities and different perspectives about coding, an approach that needs to be cultivated and nurtured.

1. Introduction

Increasing attention has been given to children's acquisition of 21st-century skills and digital competences. In accordance with this need, computational thinking and coding have, in recent years, become an integral part of school curricula in many countries. Estonia, Israel, Finland and the United Kingdom are only a few examples of the growing efforts of governments to integrate coding as a new literacy and to support students in creative problem-solving tasks (Hubwieser, Armoni, Giannakos, & Mittermeir, 2014). Similarly, organizations such as “code.org”, “codeacademy.com” offer fruitful learning environments to promote coding activities. In addition, ACM, the Computer Science Teachers Association, National Math and Science Initiative and K-12 Computer Science Framework provide guidelines for informing and building communities for the teaching of computer science. While ever more people believe that coding skills are as important as math and writing (Horizon, 2015), there is an acute need for evidence about the design of effective and engaging learning activities for children (Guzdial, 2017; Lye & Koh, 2014).

Since the first appearance of Papert's Logo programming environment introduced in the 1960s, many other programming tools have emerged. Today, there is a lot of child-friendly software that offers an intuitive and pleasant experience while coding. Examples of such software are Scratch, Alice and Kodu. By participating in coding activities, children are exposed to computational thinking (Wing, 2006), which involves, but is not limited to, critical thinking, problem solving and creativity. These activities are apparent in both formal and informal settings and are characterized by different designs, technologies and approaches (Papavlasopoulou, Giannakos, & Jaccheri, 2017). Previous research shows that different approaches can combine physical fabrication and coding (Kafai & Vasudevan, 2015), while others, such as Buechley, Eisenberg, Catchen, and Crockett (2008), used LilyPad Arduino to make coding attractive to girls. By using Logo-based environment and an interactive white board from kindergarten age, children developed mathematical concepts and social interaction, at the same time as enjoying the learning activity (Fessakis, Gouli, & Mavroudi, 2013). Different benefits arise from grounding coding activities on constructionism (Papert, 1980); children are given the opportunity to enhance their understanding of programming concepts, to promote collaboration with friends, and to change their attitudes towards coding (Fessakis et al., 2013; Maloney, Peppler, Kafai, Resnick, & Rusk, 2008; Papavlasopoulou, Giannakos, & Jaccheri, 2019).

Despite the growing research and the many possibilities offered by learning environments to design constructionism-based coding

* Corresponding author. Sem Sælands vei 9, NO-7491 Trondheim, Norway.
E-mail address: michailg@ntnu.no (M.N. Giannakos).

https://doi.org/10.1016/j.chb.2019.03.003
Received 16 July 2018; Received in revised form 15 January 2019; Accepted 3 March 2019
Available online 04 March 2019

activities for children (Kafai & Burke, 2015), there are relatively few studies focusing on gender issues in making and coding activities for children (Papavlasopoulou, Giannakos, et al., 2017). Gender discrepancy in science, technology, engineering and mathematics (STEM) exist, with women more underrepresented in the field of computer science (Cheryan, Ziegler, Montoya, & Jiang, 2017). In terms of interest, the gender gap starts at elementary school (Ceci & Williams, 2010). Among the different factors that impact on women following computer science paths are the lack of positive educational experiences in their childhood (Adya & Kaiser, 2005), their fear of being involved in very technical coding courses, and stereotypes and misconceptions around careers in computing (Teague, 2002). Girls' interest in computer science from a young age possibly fades because of a gendered or non-appropriate pedagogical approach (Schulte & Knobelsdorf, 2007). With respect to decreasing the gender gap in participation and to attracting more girls to computing, several studies on coding workshops have focused on differences in girls' competences compared to boys' (Kalelioğlu, 2015), while others have explored increasing girls' self-efficacy, interest, attitudes and confidence (Cheryan et al., 2017; Çakır, Gass, Foster, & Lee, 2017). However, the methods used in these studies were traditional qualitative and/or quantitative instruments like surveys, tests and interviews (Papavlasopoulou, Giannakos, et al., 2017). In order to gain knowledge on how to design the coding activities, it is necessary to use new objective methods to investigate the existence of gender differences in aspects like learning performance in coding activities and to discover the main differences between boys' and girls' practices (Papavlasopoulou, Sharma, & Giannakos, 2018; Papavlasopoulou, Sharma, Giannakos, & Jaccheri, 2017).

In this study we designed and evaluated a coding workshop for children aged 8–17 years old. The aim was to investigate if gender differences exist in children's coding behavior. We used objective measures (children's gaze) and triangulated them with qualitative data (interviews with the children) in order to acquire a deeper understanding of children's perspectives and practices. Eye-tracking is a method widely used in computer programming (Obaidellah, Al Haek, & Cheng, 2018), but studies with children are very limited and, to the best of our knowledge, eye-tracking has not been used to discover gender differences in children. Our study addressed the following research questions:

(1) Is there a difference in girls' and boys' gaze patterns in coding activities?
(2) Is there a difference in learning gain among boys and girls in coding activities?
(3) What are the differences in boys' and girls' strategies and implemented practices during the coding activity?

The rest of paper is structured as follows: in the following section, we present the related work and background theories; the third section describes our constructionism-based coding activity and the methodology used in our study; the fourth section presents the research findings; and the fifth and final section discusses the results in relation to the relevant literature, presents the research limitations, and suggests directions for future research.

2. Related work and background theory

2.1. Gender differences and characteristics in children's coding

The gender gap in STEM-related contexts has been examined in recent years (Cheryan et al., 2017). The profile of a computer scientist seems still to be stereotyped, and women show less interest in computer science and less likelihood to consider it properly as a possible future career (Wang & Degol, 2017). As a way to attract more females to computer science and to increase gender equality, educators have focused on offering diverse coding experiences specially for girls (Kelleher, Pausch, & Kiesler, 2007). The aim is to increase their interest in coding, enhance their attitudes, and examine the causes of gender differences (Denner, Werner, & Ortiz, 2012; Robertson, 2012). Çakır et al. (2017) conducted a game-design workshop for girls, showing that, at the end of the workshop, the girls had better attitudes towards computer science, higher confidence and self-reported competence with computers. However, a study by Robertson (2013) that investigated the influence of a game-development project on students' attitudes revealed that the level of enjoyment of the project was higher in boys than in girls, and that the project did not increase the possibility of them studying computer science in the future.

Bruckman, Jensen, and DeBonte (2002) showed that gender did not affect children's performance level in coding. Similarly, in a study of a game-development task for fourth-grade students, Owston, Wideman, Ronda, and Brown (2009) demonstrated that there were no gender differences in the learning outcomes. No significant gender differences were found in elementary school students' competence, interest at school and the use of deep learning strategies while constructing a "drag and drop" game (Vos, Van Der Meijden, & Denessen, 2011). Another study involving game-making showed that girls focused more on trying to improve their games following their peers' recommendations, and that overall they achieved a higher game quality (Robertson, 2012). In addition, in a study of the use of the code.org website to teach coding to primary school students, it was shown that girls' means of reflective thinking skills towards problems solving were higher than boys, although the results showed no statistically significant difference (Kalelioğlu, 2015).

Concerning children's approaches to and practices of coding activities, studies have reported differences depending on the gender of the participants. Robertson (2012) found that girls approached the game-making process differently when using a software called Adventure Author. For example, girls were spending more time than boys in writing dialogues for their games. In addition, girls' greater interest in narration was reflected in the use of Alice software, which is specifically designed for storytelling; nevertheless, this resulted in equal gains to those achieved by the girls who used the generic version (Kelleher et al., 2007). Denner et al. (2012), in an analysis of 108 games created by 59 middle-school girls, found that they were facing difficulties in organizing the design of their game and in handling their code when many pieces were involved. In another study of girls creating games, the results showed that they were very focused and collaborated well in the debugging process; as a result, they were trying to work more on their own before asking for help from the instructors (Denner, 2007). On the other hand, when boys were dealing with needlework, they were nervous when engaging with craft practices and they considered them to be "women's work". However, they were committed and realized how challenging and demanding it was to complete their task. Further, in the same study, they were able to see their actions in a tangible way and were testing different codes until they managed to accomplish the desired outcome (Searle & Kafai, 2015). A study with young participants aged 10–12 showed that girls spent more time on aesthetics and put more effort into having a good technical functionality (Lee, Kafai, Vasudevan, & Davis, 2014).

2.2. Gaze behavior and gender differences

Various eye-tracking studies in the past have shown results based on gender differences in a variety of contexts such as usability (Pan et al., 2004), Google searches (Lorigo et al., 2008), web design (Djamasbi et al., 2007), advertisement (Hewig, Trippe, Hecht, Straube, & Miltner, 2008), visuo-spatial planning (Cazzato, Basso, Cutini, & Bisiacchi, 2010), visual toy preferences (Alexander, 2006), facial emotion recognition (Schmid, Mast, Bombari, Mast, & Lobmaier, 2011) and color preferences (Moss & Colman, 2001).

In a study in which the participants were asked to observe a set of preselected gaze behavior, Pan et al. (2004) found that the average fixation duration for men was significantly higher than that for the women in the experiment. This depicted a deeper observational gaze behavior by men than by women in relation to webpages. In a comparison of search tasks (informational vs. navigational), Lorigo et al. (2008) did not find any gender differences based on engagement (pupil dilation). In a web-usability study, Djamasbi et al. (2007) found that the color of a specific part of the webpage influences the gaze behavior of women more than men; similar color preferences were found by Moss and Colman (2001). In an eye-tracking study to examine the role of models' gender in an advertisement on the ratings given by men and women, Hewig et al. (2008) found that the gender of the model has more impact on the ratings of men than of women; however, the authors did not find any other gender differences in gaze behavior. Cazzato et al. (2010) conducted an eye-tracking study in which the participants were required to solve visuo-spatial problems by finding the shortest path between a source and a destination. The authors found that females used more cognitive resources than males. However, in terms of gaze behavior, their study found differences between males and females. When Alexander (2006) presented “masculine” and “feminine” toys to participants, the visual behavior of men and women did not reveal any differences, even though there was clear gender bias in the preferences for toys. In another study in which the participants were asked to identify the facial emotions depicted in pictures, Schmid et al. (2011) found that women performed better than men. However, the authors did not find any gender-based gaze differences in the interfeatural saccades.[1] The only difference was that women processed information less locally than men did (Schmid et al., 2011).

Most of these studies show that gender differences are at the preference and/or performance levels. There are only a few gender differences as far as the gaze behavior is concerned. This leads to a working hypothesis that *there are not many gaze behavioral differences between men (boys) and women (girls)*; the differences occur at the attitudinal/strategic level.

2.3. Constructionism and computational thinking framework

Papert's theory of constructionism argues that learning experiences are more powerful when learners are actively involved in the learning process by creating their own projects (Papert, 1980). With the experience of developing a project, children build on their previous knowledge and discover new knowledge without receiving it passively. The core element in constructionism is an “object-to-think-with”; this is what will provide the opportunity for learners to interact and support their own thinking. However, constructionism is more like a synthesis of characteristics that will result in effective learning. Thus, together with the core element to stimulate an individual's thinking, it is necessary for there to be active involvement of the learners, socially meaningful elements and social interaction (Kafai, 2006).

Computer game programming represents an educational strategy of constructionist learning. During the process of making a game, children try to achieve a goal and to master their own ways of learning and thinking. The use of a programming environment offers the possibility of creating socially meaningful artifacts, of communicating with others and of having a pleasant and engaging experience (Robertson & Howells, 2008). Many studies have used constructionism as a support for coding activities in both formal and informal settings in order to promote coding, problem solving, critical thinking and collaborative skills (Papavlasopoulou, Giannakos, et al., 2017). In this study, which is based on constructionism, *we design a coding activity for children of all ages that does not demand any previous experience in coding.*

[1] While recognizing the emotions from the facial images, it is necessary to look at the specific points on the face, such as eyebrows, shape of lips and the extension of the eye-opening. The gaze shifts from these features are called interfeatural saccades.

Computational thinking can be traced to Papert's strong support of the idea that children who use the Logo programming language develop algorithmic thinking (Papert, 1980). However, the term “computational thinking” was made popular by Wing (2006), who argued that “computational thinking represents a universally applicable attitude and skill set everyone, not just computer scientists, would be eager to learn and use” (p. 33). Since then, different efforts to define computational thinking have appeared, with the aim of supporting the importance of research on making computational thinking a 21st-century literacy accessible to all (Guzdial, 2008). Examples include the Computer Science Teachers Association and the International Society for Technology in Education framework (Barr & Stephenson, 2011), and the National Research Council's “Framework for K-12 Science Education” (NRC, 2012). In our study, we adopted Brennan and Resnick’s (2012) computational thinking framework. With respect to Scratch, their framework suggests three key dimensions to portray computational thinking: computational concepts (concepts the users engage with when they program, such as parallelism and variables); computational practices (practices that users develop, such as abstraction and debugging); and computational perspectives (perspectives users develop for computation, themselves and the world around them). Brennan and Resnick's computational thinking framework enables the researcher to monitor the coding activity and to understand how children use the different constructs and deal with the concepts, how they focus on learning and adopt different thinking practices, and, finally, how their perspectives evolve in relation to themselves, others and the technological world. In our constructionism-based coding activity using Scratch, these three dimensions were utilized to explore and gain insights into children's experience of coding.

2.4. Selectivity theory and gender schema theory

In terms of information processing in task-related circumstances, two theories have been used to shed light on gender differences, selectivity theory and gender schema theory. First, selectivity theory aims to explain gender perceptual differences (Meyers-Levy, 1986; Meyers-Levy & Maheswaran, 1991; Meyers-Levy & Sternthal, 1991). This theory implies that gender perceptual differences rely on how males' and females' brains function. According to this theory, males rely on their right hemisphere, which indicates a “selective” way of processing information. More specifically, selective processing shows a more heuristic approach, focusing on the most prominent signs and visual representations instead of the details, which requires less cognitive effort (Meyers-Levy, 1989). In contrast, females are more likely to employ their left hemisphere, which results in what is named “comprehensive processing” (Goodrich, 2014). Females' approach shows their tension in a detailed analysis of all available information related to the specific task. Females assess and examine all the factors in a given task, involving themselves in more extensive elaboration compared to men.

Second, gender schema theory argues that there are gender differences in the way males and females use schema for cognitive processing (Martin, Ruble, & Szkrybalo, 2002). According to this theory, a schema relates to cognitive structures that we apply to form our perceptions, and this differs depending on gender. Males' schema associates with success, having as a result an attitude more strongly related to risk-taking, ambition and competition than that of females (Noble, Griffith, & Adjei, 2006). On the other hand, females focus more than males on collective actions and tend to care about relationships, sharing information frequently (Putrevu, 2001).

Previous empirical studies have used the theories mentioned above to examine gender differences in contexts of information processing. For example, in a study on the use of websites in e-commerce, Simon (2001) found that males' and females' preferences differ. Regarding the use of programming environments in the industry, Burnett et al. (2010) revealed significant gender differences in using and exploring software features and in users' confidence. In addition, other studies have used eye-tracking data to examine gender differences. Hwang and Lee (2018) found that gender differences exist in terms of visual attention and attitudes towards the presented products in online shopping environments. Exploring females' and males' characteristics of identifier style in source code reading, Sharafi, Soh, Guéhéneuc, and Antoniol (2012) presented mixed results, with no significant differences in accuracy, time and effort, but gender differences in strategies used in males' and females' approaches.

In our study, we aimed to investigate gender differences in coding activities for children. Given our coding task, children's activity required cognitive processes to successfully complete their goal, using Scratch and social interaction to collaboratively create a game. Therefore, we assumed that there are gender differences in how girls and boys behave during coding activities. In order to investigate the impact of gender, we used eye-tracking measures to generate objective quantitative data and qualitative data to examine different aspects of gender characteristics.

3. Methodology

In this section, we present the details of the two experiments, data collection, variables and analysis. We ran two studies, one in autumn 2016 and one in autumn 2017. For both experiments, the duration of the workshop for all groups of students was the same, as an out-of-school one-day activity. The implementation of the coding workshop was an intervention over two years, with few differences in design of the activity; the main differences were in the research design, the description of evidence and the results from different instruments used for data collection.

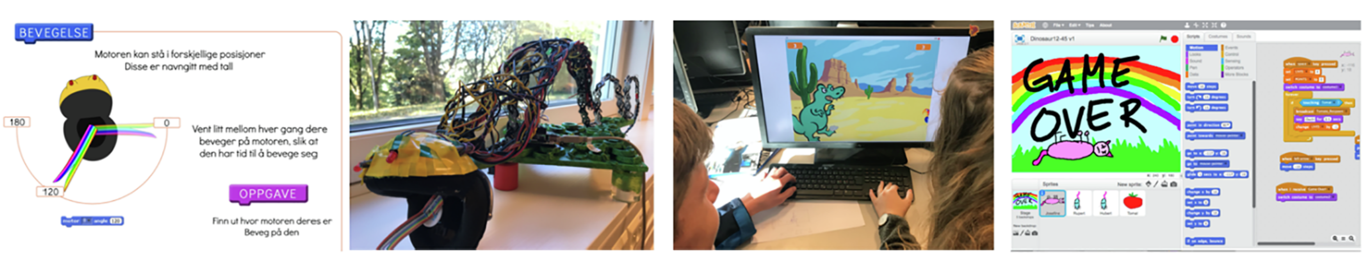
3.1. Description of coding activity

Based on the constructionist approach and its main principle, learning by doing (Resnick et al., 2009), we designed and implemented a coding workshop at the Norwegian University of Science and Technology (NTNU), in Trondheim, Norway. Our coding workshop was an out-of-school activity, in which children, novices to coding, from 8 to 17 years old interacted with digital robots, using Scratch for Arduino (S4A), and then coded their own game using the Scratch programming language. At each workshop, the children worked for approximately 4 h. Five assistants with previous experience in similar activities were responsible for instruction and the workshop procedure. The workshop consisted of two main parts: interaction with the robots, and creating games with Scratch.

Interaction with the robots: During the first part of the coding activity, the children interacted with digital robots built by an artist using materials recycled mainly from computer parts. First, once the children had entered the room and been welcomed by the assistants, they sat in teams next to one robot per team. The assistants gave a brief presentation of the workshop's activities and asked each of the children to pay attention to a worksheet placed on the desks next to them. The goal was to familiarize themselves with the robots by filling in simple questions regarding the exact place and number of the sensors and lights on the robots. Next, the children used a paper tutorial with instructions (Fig. 1a) on how to make the robots (Fig. 1b) react to the physical environment with visual effects using simple loops of Scratch for Arduino (e.g., to make the tongue of the snake robot move when there is less light at a sensor). The teams worked collaboratively and independently to complete this task (Fig. 1c). The duration of the first part varied between 45 min and one and a half hours. When all teams had finished, the children had a break before the next section began.

Creating games with Scratch: This section was the main activity of the workshop and lasted approximately 3 h, without the presence of the robots. The goal was to successfully develop a simple game coded in Scratch. To achieve this goal, the assistants gave another paper tutorial with examples of all the basic computer science concepts and possible loops the children should use to complete their own game. Assistants advised the children how to manage the process of game development by working collaboratively. They were advised that, first, they should think and decide the story for their game and then create a draft storyboard. When they had finished, they started coding using Scratch. Throughout the activity, children could ask for support from the assistants whenever they needed it. The assistants offered their guidance to the teams to help them complete their games by introducing even more complex computer science concepts when needed. Finally, after completion of the games, the children reflected on and played each other's games (Fig. 1d).

At each of the workshops, the same parts were conducted, with children participating once after being carefully selected according to their age; participants were of the same grade or within a small age range. The design of the coding activity provided flexibility, and the workshop instructors had the appropriate experience to be able to assist the children properly, taking into account the children's age. In the first part of the coding activity (interaction with robots), the children needed to perform simple tasks, but these were still things they had not done before since the robots and their functionalities were specially designed for our workshop. The design of the activity took into consideration the amount of time and support needed to complete the tasks, as these were likely to differ depending on the group of children. Creating games using Scratch was the second and main section of the coding activity and allowed the possibility for each of the teams to create a functional game using the basics or, depending on the team's capacities, to create a game with more advanced features. Teams of children worked independently with help from the instructors, who had the knowledge and the experience to help with advanced concepts according to the teams' needs. Scratch was used as a programming environment for the coding activity since it does not require any special expertise and children of all ages can quickly learn the basics to start creating in an efficient way. Moreover, Scratch is simple enough for novices and young users, while at the same time having enough power and functional variety to keep users engaged.



(a) (b) (c) (d)

Fig. 1. (a) Example of the paper tutorial; (b) a snake-shaped interactive robot; (c) children collaborating on the creation of their game; (d) example of a created game.

3.2. Sampling

All the participants of the two studies were K-12 students from the mid-Norway region. Both studies took place at the university campus in specially designed rooms. Data related to both studies were collected after permission from the Norwegian Centre for Research Data (NSD), following all the regulations and recommendations for research with children. The children volunteered their participation in the eye-tracking study and the interviews. A researcher contacted the teacher and the legal guardian of each child to get a written consent form that gave permission for the data collection.

3.2.1. Participants: study one

The study lasted two weeks during autumn 2016, with 44 children from the third to 12th grades (aged 8–17 years old), 12 girls (mean age: 12.64; standard deviation [*SD*]: 2.838) and 32 boys (mean age: 12.35; *SD*: 2.773). Five workshops were held in total, all following the same process for the coding activity and designed for novices in coding. Some of the participants in the sample (13–17 years old) were recruited from the local schools that had applied to take part in our activity. The other set of participants (8–12 years old) were youngsters who attended local coding clubs as an after-school activity.

3.2.2. Participants: study two

In autumn 2017, children from eighth to tenth grade (age 13–16 years old) participated in the coding activity. The sample consisted of 105 participants in total, 69 boys and 36 girls (mean age: 14.55; *SD*: 0.650). At the end of each workshop, some of the participants were interviewed. In total, 44 children were interviewed, 23 boys and 21 girls.

3.3. Data collection and analysis

3.3.1. Data collection and analysis for the first study

Eye-tracking measures: As mentioned above, this study is one of the few so far to utilize children's gaze. We recorded children's gaze while they were coding using the Scratch environment during both parts of the activity. The eye-tracking data were collected using four SMI and one Tobii eye-tracking glasses. The sampling rate for all the eye-tracking glasses was set to be 30 Hz for the binocular eye-tracking. The average accuracy for both SMI and Tobii glasses was 0.5° at a distance of 40 cm. The visual field was divided into six areas of interest (AOI), five of which were in the Scratch interface, with the sixth in the physical robot area, as shown in Fig. 2.

From the eye-tracking data, we calculated the following measurements:

Average fixation duration: High fixation duration indicates that the participant is having difficulty in extracting information (Just & Carpenter, 1976). We used a mental rotation task, with angles of 0°, 120° and 180°, to study the relation between problem difficulty and gaze patterns. The results showed that with an increase in the rotation angle (increasing difficulty), the fixation duration at the center of the figure and the arms of the structures increased.

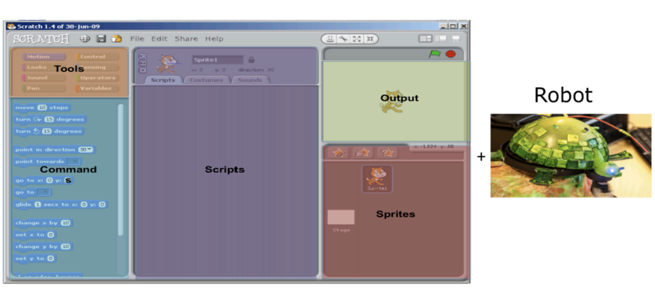


Fig. 2. The five areas of interests (AOI) for the screen; the sixth AOI was the robot.

Average change in saccade direction: Longer saccades show meaningful transitions in attention (Goldberg, Stimson, Lewenstein, Scott, & Wichansky, 2002). In a web search task, the authors used a set of different tasks on a webpage, so that the participants had to look for particular information to complete the tasks. The results showed that pre-planned eye movements were accompanied by longer saccades (Goldberg et al., 2002).

Saccade amplitude: The angle between two lines, if more than 90°, reflects a change of plans, revision or a failed expectation/hypothesis/anticipation (Cowen, Ball, & Delin, 2002). In a usability study, the authors found that the change in saccade direction often depicted the gaze behavior of not finding something which the participants anticipated to find at certain places (Cowen et al., 2002). This can be translated, in terms of programming behavior, as having a certain hypothesis and a failed verification.

Gaze uniformity: This is an individual measurement of engagement. This measure captured the uniformity of the time spent on all AOIs. The distribution was computed as a vector of length six (there are six AOIs) comprising the proportion of time spent on each AOI. The uniformity was computed as the inverse of a Kullback-Leibler divergence between the original proportionality vector and a uniform distribution with the same minimum and maximum limits as the original vector.

Time spent on each AOI: We divided the whole visual field into six AOIs – five on the screen and the sixth as the robot. We used specially made QR codes to scan the robots and the area around them (Fig. 2). The five AOIs on the screen were as follows:

- **Tools:** This area of the screen contained a general categorization of the commands available; for example, commands to control the motion, looks, sound, and other variables.
- **Command:** This area contained all available commands within the currently selected tools.
- **Scripts:** This was the area of the screen in which the coding task was performed.
- **Output:** This area showed participants the output of their scripts.
- **Sprites:** This area controlled the aesthetics of the program. The participants could change the appearance of the animated character using the characters available in this part of the interface.

Transitions among AOIs: We also computed the transitions to and from one AOI to another. This helped us to understand the temporal relationship between the children's gaze patterns and to depict the coding process used by the participants. For example, frequent transitions between script and output, or script and robot, depict the typical behavior of hypothesis verification. The participants made a small change in the program based on a certain hypothesis about the output or the robot's movement; once they had observed the output/robot's behavior, either their hypothesis was confirmed and they moved onto the next step in coding, or they modified the program to reverify their hypothesis. This behavior would result in a high number of transitions between the script and output/robot. We considered only three types of transitions for this analysis based on the literature which says that experts shift their attention between the code and the output more than novices do (Hejmady & Narayanan, 2012). This is why we chose to compare the gaze transitions between the script and robot/output. The third type of transition we included in our analysis was that between the commands and the script areas. These transitions imply a behavior that shows a thinking process of "what comes next in the code?" by the children.

Relative learning gain (RLG): The children completed pre- and post-knowledge acquisition tests. These consisted of nine coding questions of increasing difficulty. The questions were adapted from a previous study (Grover, Cooper, & Pea, 2014) and followed instructors' suggestions. The children took approximately 10 min to finish the tests.

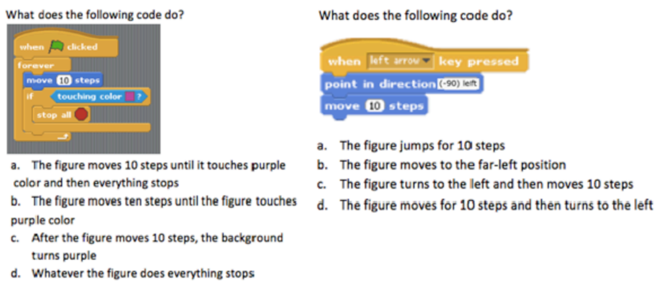


Fig. 3. Example of questions on the knowledge acquisition test.

The tests were paper-based and were manually graded by the researcher. Fig. 3 shows two sample questions from the test.

In our study, we calculated the RLG as defined by Sangin, Molinari, Nüssli, and Dillenbourg (2008). RLG is more accurate compared to learning gain since it takes into consideration the difficulty in gaining more knowledge if the learner is already very knowledgeable in a subject.

$$RLG = \begin{cases} \frac{Posttest - Pretest}{Max \cdot in\ Pretest - Pretest},\ if\ Posttest \geq Pretest \\ \frac{Posttest - Pretest}{Pretest},\ if\ Posttest < Pretest \end{cases}$$

Data analysis: To identify the relationship between gender, gaze, and RLG, we used analysis of variance (ANOVA) to compare the variables across different categories, since all the variables were normally distributed. In addition, we checked the assumptions for ANOVA, and if we found variables that did not satisfy the homoscedasticity condition, a version of ANOVA was used in which homoscedasticity is not assumed. This was done using the Welch correction for *F*-statistic.

3.3.2. Data collection and analysis for the second study

In the second study we utilized a qualitative approach. We collected data from multiple sources, including post-workshop interviews, observational field notes, and participants' Scratch games. All data were compared and cross-checked for triangulation.

Interviews: Participants were interviewed individually in Norwegian after the end of the workshop. The interviews were audio recorded, lasted approximately 10 min, and used a semi-structured protocol. During the interviews, students were asked to discuss their workshop experience, such as what they found to be the easiest, the most difficult and the most frustrating parts of creating the artifact, how they found their team collaboration, what they liked, and how they found their interaction with Scratch.

Interviews were transcribed and translated after the end of the workshop. To analyze the transcribed interviews, two researchers followed the coding method proposed by Saldaña (2015) for qualitative inquiry. Saldaña's coding method describes a cyclical model that moves from codes to categories and themes. Analysis of the semi-structured interviews focused on identifying categories and then the overall themes forming the codes emerged from participants' answers. Each transcript was first individually reviewed by two researchers and then, after a focus group and discussion, the two researchers agreed on the major themes that had emerged. In all codes under each category and then theme, it was indicated whether it derived from a boy's or a girl's interview. This helped us to detect gender differences in the already created themes. Analyzing the interviews allowed us to provide a holistic understanding of girls' and boys' perspectives on coding activities and to identify any potential masculine characteristics as well as girls' hesitation or stereotypes related to their participation in coding activities.

Observations and artifacts: Independent assistants during the workshop kept field notes. Assistants were close to each of the teams and took notes on all tasks. In order to identify what type of help participants were receiving from the assistants (see Table 1), we used

Table 1
Observations' coding scheme.

Number	Explanation
0	**Validation:** Students want confirmation, not information
1	**Where:** Only needed help navigating the Scratch GUI
2	**What:** Only needed a reminder of the name of the concept
3	**How:** Given name of concept, still needed help to complete task
4	**Reteach:** Had to reteach concept and execution

Franklin, Conard, Boe, Nilsen, Hill, Len, & Kiefer (2013) coding scheme. In addition, observation notes included other incidents involving the participants that occurred during the workshop and which concerned the process they followed to successfully complete the coding tasks. Examples of such incidents were: how the participants distributed the roles among the team; which aspects they spent most of their time on while coding; and what their reactions were.

Observation field notes helped us monitor girls' and boys' practices during the process, capturing their behavior in all aspects, as well as the type and frequency of help received from the teaching assistants running the workshop.

For the purpose of this study, and to be able to investigate any potential differences between girls' and boys' approaches, we randomly analyzed observations of two teams of girls and two teams of boys (whose members were also interviewed), together with the final games created by these teams. Each of the observation notes (one set for each team) was reviewed by two researchers. Using content analysis, the main actions indicating a specific behavior were identified, and the frequencies of help level were calculated.

Finally, artifacts (games) developed by the teams were evaluated in terms of the learning opportunities related to computer science and computational thinking concepts offered by coding a game. We collected four versions of the games approximately every hour during the workshop. The evaluation of the artifacts included loading and playing the game to ensure its functionality and playability. For the analysis of each version of the games, we analyzed the games based on computational thinking components (i.e., flow control, data representation, abstraction, user interactivity, parallelism, and logic), giving a score for each of them from 0 to 3 (a rubric in which 3 shows proficiency, and 0 means that the skill is not evident). Artifacts were used as an extra source to determine the main characteristics (such as the game's theme, aesthetics, and storytelling) of boys' and girls' codes and their use of specific concepts related to the learning objectives of our workshop, as well as to discover any unexpected learning outcomes.

The analysis of observations and artifacts created by the teams focused on exploring any potential specifications that underline gender differences.

4. Research findings

4.1. Results from the first study

4.1.1. Gender and RLG

To investigate any potential gender difference in the RLG, an ANOVA (without assuming equal variances) with the RLG as the dependent variable and gender as the independent variable was used. The results showed no significant difference between boys and girls, $F(1,18.05) = 0.18$, $p = 0.65$ (see Fig. 4).

4.1.2. Gaze measures and gender

Next, in order to examine any potentially significant differences between the gaze measures of girls and boys, we utilized a one-way ANOVA (without assuming equal variance across gender). Table 2 shows the results, which indicate that there is no significant difference in the gaze behavior between girls and boys.

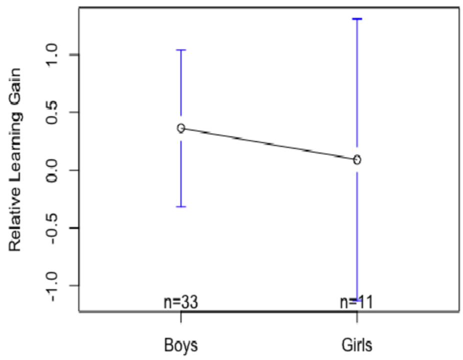


Fig. 4. RLG for boys and girls; the blue bars represent the 95% confidence interval.

Table 2
Testing the effect of gender on gaze behavior using analysis of variance (ANOVA).

	Mean (*SD*) girls	Mean (*SD*) boys	DoF 2	*F*-value	*P*-value
Uniformity	0.56 (0.25)	0.53 (0.25)	17.26	0.13	.72
Fixation duration	279.80 (55.94)	267.85 (100.10)	31.52	0.24	.62
Saccade direction	39.94 (12.33)	39.36 (17.25)	24.09	0.25	.61
Saccade amplitude	170.43 (49.27)	192.23 (64.26)	22.31	1.37	.25
Tools	0.13 (0.09)	0.16 (0.08)	15.10	0.61	.44
Script	0.14 (0.08)	0.16 (0.08)	17.69	0.17	.68
Commands	0.15 (0.06)	0.15 (0.06)	16.51	0.16	.69
Sprites	0.15 (0.11)	0.19 (0.07)	12.65	0.87	.36
Output	0.15 (0.08)	0.15 (0.07)	14.83	0.01	.90
Robots	0.16 (0.06)	0.16 (0.06)	14.83	0.01	.99
Scripts.command	0.23 (0.24)	0.23 (0.26)	18.32	0.32	.57
Scripts.output	0.24 (0.27)	0.21 (0.27)	16.81	1.39	.25
Scripts.robot	0.26 (0.22)	0.20 (0.22)	17.01	0.23	.63

4.2. Results from the second study

4.2.1. Interview results

This section describes the identified themes relating to children's perceptions that are relevant to our research question.

Improved confidence and self-efficacy in coding: In all interviews, children expressed that they managed to accomplish the tasks required. This was also evident from the fact that all teams managed to have a complete and functional game. Some of the comments, belonging in that category, are a clear indication of achievement expressed both by boys and girls. For example:

"*At the end we managed to do everything we wanted" – Thomas*

"We tried and made all the things we wanted to happen, we found how things worked" – Maria

Other comments indicate confidence in the game design and coding:

"Now, I know it is not so difficult to make a game, I can do it again" – Anna

"I knew something about coding before, but I didn't know how easy it was, I thought it was much harder to make a game, I can definitely do it" – Arne

In addition, it is evident that only girls reported that they did not know what they could do with Scratch or what coding is. Comments were similar to the one below:

"I was not at all looking forward to coming to the workshop, I thought it would be some geek stuff, I have never tried something similar" – Ingrid

Perceptions about leadership and collaboration: This theme reveals how boys and girls faced the collaboration process and how much they contributed to their team. In mixed teams, when a boy knew about coding, girls stated that a boy had to be the leader, while in girls' teams they appear to have had equal roles:

"We were lucky to have a boy in our team who has coded before, so he was leading in Scratch and the game creation" – Sonia

"We let Marius lead the team as he was more capable than us in coding" – Olga

"We distributed the roles equally and changed the rotation of control in different tasks" – Cecile

In the interviews, all boys indicated that they contributed to their teams in terms of coding, whereas girls mentioned that not all of them coded but that they felt a valuable part of the team because of their ideas:

"I didn't do much in the coding, but if it wasn't for me, they would have done a very boring game" – Jane

"I was not the one responsible for coding, but I decided how things will look or behave" – Katia

Interaction with Scratch: There were no differences among boys and girls in relation to their experience with Scratch and coding per se. Comments below show that both girls and boys had similar difficulties, challenges and frustrations during the creation of their game.

- General impression of Scratch programming environment:

"You can put together blocks and make a big script" – Daria

"If you put weird things together properly you can actually do whatever you want" – Marius

During the workshop the participants faced difficulties completing their projects. Many aspects of game design and coding appeared to be challenging and sometimes frustrated the children in their attempts to finish their projects.

- Difficulties of coding with Scratch:

"Sometimes the code can become messy" – Bjorn

"We couldn't make something stop when it was touching something else" – Annete and Peter

"It was so difficult to make our timer and score counting" – Sofia and Kevin "Making the character move was quite difficult" -Ines

"When you wanted to add something new in the function, then you had to go back and check everything again" – Arne

"We had to test and fix our game again and again until the end" – Martha

This category indicates what was easy in the coding part with Scratch. Participants liked many parts of Scratch and their interaction with it.

- Easiness of coding with Scratch:

"It was easy to put the blocks together" – Sofia

"I could find easily what I was looking for" – Lukas

"I found it easy that I could make the character as I wanted and then make the platform he was standing on" – Stefano

Some of the children were also able to refer to specific commands in Scratch. For example:

"I could easily use the sensing in a loop, to change the color of the

platform" – Ines

Affective engagement state: Participants reported that they liked coding the game. Equally, girls and boys reported that they had fun during the workshop. One participant commented:

"It was fun to code with colorful blocks that look childish" – Daria

Also, many participants described the workshop activity simply as:

"It was so fun" – Anne, Marius, Cloe

It was interesting to discover that participants with previous experience with coding found the whole workshop experience pleasant, showing that having experience is not a limitation to attending that type of workshop. This was highlighted in the following comment:

"The workshop was more fun because I knew about coding" – Alex

Even though some will not try it again, their experience was quite fun and interesting:

"I will not code again at home, but I had a lot of fun and I liked making the game and playing all the different games at the end" – Singrid

Moreover, they expressed their enjoyment at having an experience outside school:

"We don't do these things at school, I am happy I tried something else" – Daria

4.2.2. Observations and artifacts-analysis results

Our analysis of the observations from girls-only and boys-only teams revealed a few differences between their practices and behavior during the workshop. The results showed that girls had a different approach than boys on how they were organizing their tasks. From the beginning of the task, girls assigned roles and split the responsibilities (i.e., cooperation/dividing labor). For example, girls started thinking and designing the game, and in one of the teams a girl who was very good at drawing started creating a storyboard for their game on paper. On the other hand, boys started immediately navigating in the Scratch interface, trying different commands for a while without having a concrete plan for their actions. The teaching assistants needed to ask the boys' teams to concentrate, think of an idea for their game and make a quick storyboard. Girls looked more at the paper tutorial, trying to find examples of code, whereas boys had the tutorial on the side and only after the assistant's prompting did they start to look at it. Both teams had the same reaction when a team member was not interested: they tried at least to give him/her a task. It was apparent that girls' teams discussed more the decisions that they should take, and all were involved at every stage of the game creation. In addition, they paid attention to all aspects of the game with equal consideration. Boys cared more about the "how to code" part and using the Scratch interface, and they were less interested in the ideas and the aesthetics (e.g., color, what the character would look like, background) of their games. In terms of help received from the assistants, all teams had approximately the same amount, between five and seven times. However, girls were more persistent than boys in trying on their own more before asking for help. Taking into consideration the type of help received each time, the only prominent difference was that boys wanted more confirmation of their actions in Scratch together with the approval of the assistants that they had created an interesting or funny game. Difficult parts that all teams needed a lot of help with were "how to make characters jump", and the insertion of variables, high-score or time.

Children's games evaluation showed that girls' nature was present in their games, as was clear from their use of female characters inspired by the famous "Barbie" or similar figures. Furthermore, girls' games had a simple "goal", like catching objects falling, whereas boys made games involving throwing a ball to a goalkeeper or shooting activities. In the case of the programming concepts and computational thinking components (i.e., flow control, data representation, abstraction, user interactivity, parallelism, and logic), the rubric scores were almost the same in all final versions (see Appendix). This indicates that boys and girls had a similar final performance. However, both teams of boys had better scores in the first and second versions of the games.

5. Discussion and conclusions

In this study our aim was to investigate gender differences in coding activities for children. To this end, we designed and evaluated a one-day coding workshop with participants aged from 8 to 17 years old. During the workshop, children were introduced to coding by interacting with digital robots, specially designed for the activity, and creating a game using Scratch. In all the activities, children worked collaboratively in teams to successfully complete their goal. For the evaluation, we used eye-tracking data as an objective measure, which to the best of our knowledge has not previously been used to capture children's coding gaze behavior. In addition, we collected and analyzed qualitative data (i.e., semi-structured interviews, and observations) to get a deeper understanding of children's experiences during the workshop. A qualitative approach is especially valuable for examining gender issues, since expressing opinions about gender can be vulnerable process (Popper, 1971). Our research findings reveal that gender issues in coding activities for children are a multifaceted phenomenon. According to the quantitative findings, there are no gender differences concerning RLG and gaze behavior in boys and girls. On the other hand, qualitative results from interviews, observations and the created games showed that some gender differences exist in children's approaches, as revealed by their behavior during the workshop and their perceptions.

There was no difference in the RLG between girls and boys. Therefore, children in our study showed no differences in their performances, which supports previous studies on children using other evaluation methods (Owston et al., 2009; Vos et al., 2011). Therefore, our findings provide more evidence that girls do not lack in competence compared to boys. Fisher, Cox, and Zhao (2006), in a study on adults' performance on a program comprehension task, found no gender differences and reported that men and women were equally capable of developing the skills required to be professional developers. Although more girls than boys in our interviews said that they had not known about coding before, or that they were afraid of it, they managed to be equally good as the boys. Moreover, the activities offered in our workshop were appropriate independently from the participants' gender and their previous knowledge. Furthermore, in the interviews, children reported that they had fun during the workshop, even though some of them had prior knowledge of coding. This can be attributed to the fact that Scratch is not limited: it provides many possibilities for making more advanced creations, so users can find it interesting and learn more, no matter their level of existing knowledge. In addition, the collaborative notion of the workshop enabled the students to learn from each other and not to have their own individual performance as their main goal. As shown in other studies, students perform better when working in pair programming than when working alone (Lye & Koh, 2014; Werner, Denner, Campe, & Kawamoto, 2012).

A noticeable result is that there is no difference in the gaze behavior of girls and boys. We used the objective measure of eye-tracking data, and by examining different measures we found no difference in any of them. This indicates that, regarding the actual micro-level experience of boys and girls during coding with Scratch, there is no difference in their approach based on their gaze, and hence also no difference in their cognitive processes (Eckstein, Guerra-Carrillo, Singley, & Bunge, 2016). From measures of time spent on different AOI, gaze uniformity, and transitions among the different AOI, results showed that both male and female participants were able to navigate in the Scratch interface, had a meaningful thinking process, and were engaged. Similarly, from the other measures used, results show that both genders had equivalent difficulties in extracting information (fixation duration), challenges in learning something (saccade direction change), and goals and expectations in coding (saccade amplitude). Cazzato et al. (2010) found weak gender differences in the gaze behavior of participants when trying to solve visuo-spatial problems, but that women used more cognitive resources. Other studies have found that girls face difficulties in coding when they had a lot of elements (Denner et al., 2012) or when they put more effort into having good functionality (Lee et al., 2014).

Although the results of children's gaze behavior and performance show that there are no important gender differences, the qualitative results of our study reveal that gender differences exist in the practices used by boys and girls and in their perceptions. In general, girls approached the coding activity in a different way to boys, verifying the theories of gender differences in information processing (Martin et al., 2002; Meyers-Levy, 1986). For example, girls were more organized in terms of collaboration, splitting the responsibilities and focusing on a more systematic approach in the tasks, and they also paid more attention to the tutorials. In addition, girls seemed to like more collaboration with others and to share the social part of the activity. Previous studies have shown that female students have a more trusting and sociable approach compared to male students, who are more independent and focused on caring about themselves (Rosenberg-Kima, Plant, Doerr, & Baylor, 2010). In the computer-supported collaborative learning environment, Bruckman et al. (2002) found that girls spent more time than boys in communicating. Girls' games were richer in aesthetics and graphical representation, and they also had a more “girly” approach. This is similar to other findings that show girls spend more time on dialogues (Robertson, 2012) and aesthetics. Similar to the finding of Denner and Werner (2007), our study shows that girls' teams were more persistent in attempting the tasks on their own before asking for help. Whereas girls' games had simpler tasks (like catching falling objects), boys' games had more competitive characteristics. This observation is similar to the finding of Owston et al.‘s (2009) study, in which teachers reported that boys enjoyed playing games more competitively against others. Our observation notes confirm this finding, as boys were also asking the assistants about how interesting their games were.

One of the goals of our workshop was to build children's belief that coding is something that they can do, and that it is not something that only boys would be interested in. After their participation, boys and girls reported that they felt competent to code. Another interesting result from our qualitative study is that regardless of both girls and boys reporting improved confidence and self-achievement, we find that girls have less self-efficacy. One example is that when girls were among boys in the teams they chose a boy to be the leader, indicating less confidence. They also expressed that they did not know what coding was before, that they had not tried it, that they did not know whether they could do it, and that they thought it was only for geeks. The stereotype of boys being better than girls at robotics and coding exists from the young age of 6-years-old (Master, Cheryan, Moscatelli, & Meltzoff, 2017). A possible reason why they split the roles during their collaboration is that the girls were less confident; in addition, none of the girls was trying to take control. In solo programming, men have been found to be more confident than women (McDowell, Werner, Bullock, & Fernald, 2006). In the study of Beckwith et al. (2005), females' self-efficacy was lower than men's, and women did not easily accept new debugging features.

That there are no gender differences in children's actual performance and gaze behavior while coding, and that the main differences are in their practices, indicates that practitioners should focus on characteristics that will influence girls and change their limited participation in computer science. Our results show that educators should foster girls' self-efficacy (i.e., the belief in one's capacity to succeed in tasks [Bandura, 1997]) and make them believe that they do not lack in competences; therefore, educators should be careful to avoid discriminating behaviors. Qiu, Buechley, Baafi, and Dubow (2013) found that participants' confidence, enjoyment and interest in coding and technology increase when self-efficacy grows. According to Bandura (1997), self-efficacy is important in problem solving, since it affects the individual's cognitive strategies, effort, persistence and, consequently, the learning outcome. Coding activities should take into consideration special gender characteristics and facilitate appropriate workshops. Activities focused on collaboration can also be a method to narrow the gender gap in coding activities and to view partnership as a key factor for fostering both learning and positive attitudes in students.

5.1. Limitations

This study has some limitations. First, it was very challenging to collect eye-tracking data from children. Eye-tracking measurements with children is a very difficult task since it involves small eye pupils, difficulty with the calibration, and the need for equipment tailored to children, so this caused some problems. The large size of the glasses annoyed the participants, especially the younger ones, who, as a result, had to take them off. Consequently, we had to stop the activity to calibrate them again. The young age of the participants combined with the playful environment of the workshop, in which children were very enthusiastic and also wanted to experiment with the glasses, led to us having data that we could not use because they were not from the areas of our interest.

Second, participants in the two studies conducted for this paper had slightly different ages, and we lacked participants aged 8–12 and 17. Third, we analyzed only a small number of observations from the teams and games. Including a larger amount of data could have added more value to our results. In addition, the specific design and context of the activity (i.e., the use of the Scratch tool, the coding tasks, the duration, and the other characteristics), as well as the sampling method used, limits the generalization of our findings. More precisely, the participants in our study were randomly selected volunteers from our region in Norway; other sampling methods and demographic variables might have a different effect on children's experience.

5.2. Future work

Our study suggests new aspects as the subject for follow-on research. One possibility would be to investigate in more detail specific gaze patterns of boys and girls; another would be to examine collaborative eye-tracking measures and group dynamics in both mixed and non-mixed teams of boys and girls. Another interesting approach would be to compare the effect of different learning environments on gender. Furthermore, other objective measures could be used to gain a deep understanding of the relationship between coding behavior and gender. In addition, other quantitative methods, such as surveys, could be used to supplement the collection of data relating to children's perceptions of coding.

Acknowledgments

The authors would like to express their gratitude to all of the children, teachers and parents for volunteering their time. Our very special thanks go to Letizia Jaccheri, Kristin Susanne Karlsen, Ioannis Leftheriotis, Uyen Dan Nguyen, Amanda Jørgine Haug, Lidia Luque Fernandez, Marjeris Sofia Romero, An Nguyen, Ton Mangmee, Eline Stenwig and Kristoffer Venæs Monsen.

The project has been recommended by the Data Protection Official for Research, Norwegian Social Science Data Services (NSD), following all the regulations and recommendations for research with children.

This work supported from the “Learning science the fun and creative way: coding, making, and play as vehicles for informal science learning in the 21st century” Project, under the European Commission's Horizon 2020 SwafS-11-2017 Program (Project Number: 787476). This article reflects the views only of the authors and it does not represent the opinion of neither the European Commission nor NTNU, and the

European Commission and NTNU can not be held responsible for any use that might be made of its content. This work is also supported from the Norwegian Research Council under the projects FUTURE LEARNING (number: 255129/H20) and Xdesign (290994/F20), and by NOKUT under the Centre for Excellent IT Education (Excited) (number: 16/02049).

Appendix

The rubric used for the evaluation of the artifacts:

Computational thinking components:	3 (excellent)	2 (Good)	1(Satisfactory)	0 (Not evident)
Flow control (loops and sequence)				
Data representation (variables)				
Abstraction				
User interactivity (events)				
Parallelism				
Logic (conditionals and operators)				

Are there any specific characteristics?

- Theme of the game
- •Aesthetics
- •Storytelling
- •Other comments (if any)

Artifact's rubrics scores for girls' and boys' teams.

	Computational thinking components						
Team and game version:	Flow control	Data representation	Abstraction	User interactivity	Parallelism	Logic	Total score
GT1V1	1	1	0	1	0	0	3
GT1V2	1	2	0	1	1	1	6
GT1V3	2	2	1	1	1	2	9
GT1V4	2	2	1	1	3	2	11
GT2V1	1	0	0	1	0	1	3
GT2V2	1	1	0	1	1	1	5
GT2V3	2	2	1	2	1	1	9
GT2V4	3	2	1	2	2	2	12
BT1V1	1	0	0	2	0	2	5
BT1V2	2	1	1	2	1	2	9
BT1V3	2	2	1	2	1	2	10
BT1V4	2	2	1	2	2	2	11
BT2V1	1	1	1	1	0	1	5
BT2V2	2	1	1	2	1	1	8
BT2V3	2	1	1	2	1	2	9
BT2V4	2	2	1	2	2	3	12

*GT1: girls' team number 1; *GT2: girls' team number 2; *V1: game version 1; *V2: game version 2; *V3: game version 3; *V4: game version 4; *BT1: boys' team number 1; * BT2: boys' team number 2.

References

Adya, M., & Kaiser, K. M. (2005). Early determinants of women in the IT workforce: A model of girls' career choices. *Information Technology and People, 18*(3), 230–259.

Alexander, G. M. (2006). Associations among gender-linked toy preferences, spatial ability, and digit ratio: Evidence from eye-tracking analysis. *Archives of Sexual Behavior, 35*(6), 699–709.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* Macmillan.

Barr, V., & Stephenson, C. (2011). Bringing computational thinking to K-12: What is involved and what is the role of the computer science education community? *ACM Inroads, 2*(1), 48–54.

Beckwith, L., Burnett, M., Wiedenbeck, S., Cook, C., Sorte, S., & Hastings, M. (2005). Effectiveness of end-user debugging software features: Are there gender issues? *Paper presented at the proceedings of the SIGCHI conference on human factors in computing systems*.

Brennan, K., & Resnick, M. (2012). New frameworks for studying society computer science exemplification project. *Paper presented at AERA*.

Bruckman, A., Jensen, C., & DeBonte, A. (2002). Gender and programming achievement in a CSCL environment. *Paper presented at the Proceedings of the conference on computer support for collaborative learning: Foundations for a CSCL community*.

Buechley, L., Eisenberg, M., Catchen, J., & Crockett, A. (2008). The LilyPad Arduino: Using computational textiles to investigate engagement, aesthetics, and diversity in computer science education. *Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems*.

Burnett, M., Fleming, S. D., Iqbal, S., Venolia, G., Rajaram, V., Farooq, U., & Czerwinski, M. (2010). Gender differences and programming environments: Across programming populations. *Paper presented at the Proceedings of the 2010 ACM-IEEE international symposium on empirical software engineering and measurement*.

Çakır, N. A., Gass, A., Foster, A., & Lee, F. J. (2017). Development of a game-design workshop to promote young girls' interest towards computing through identity exploration. *Computers & Education, 108*, 115–130.

Cazzato, V., Basso, D., Cutini, S., & Bisiacchi, P. (2010). Gender differences in visuospatial planning: An eye movements study. *Behavioural Brain Research, 206*(2), 177–183.

Ceci, S. J., & Williams, W. M. (2010). Sex differences in math-intensive fields. *Current Directions in Psychological Science, 19*(5), 275–279.

Cheryan, S., Ziegler, S. A., Montoya, A. K., & Jiang, L. (2017). Why are some STEM fields more gender balanced than others? *Psychological Bulletin, 143*(1), 1.

Cowen, L., Ball, L. J., & Delin, J. (2002). An eye movement analysis of web page usability. *People and computers XVI-memorable yet invisible* (pp. 317–335). Springer.

Denner, J. (2007). The Girls Creating Games Program: An innovative approach to integrating technology into middle school. *Meridian: A Middle School Computer Technologies Journal, 1*(10).

Denner, J., & Werner, L. (2007). Computer programming in middle school: How pairs respond to challenges. *Journal of Educational Computing Research, 37*(2), 131–150.

Denner, J., Werner, L., & Ortiz, E. (2012). Computer games created by middle school girls: Can they be used to measure understanding of computer science concepts? *Computers & Education, 58*(1), 240–249. https://doi.org/10.1016/j.compedu.2011.08.006.

Djamasbi, S., Tullis, T., Hsu, J., Mazuera, E., Osberg, K., & Bosch, J. (2007). Gender preferences in web design: Usability testing through eye tracking. *AMCIS 2007 Proceedings, 133*.

Eckstein, M. K., Guerra-Carrillo, B., Singley, A. T. M., & Bunge, S. A. (2016). Beyond eye gaze: What else can eyetracking reveal about cognition and cognitive development? *Developmental Cognitive Neuroscience, 25*, 69–91.

Fessakis, G., Gouli, E., & Mavroudi, E. (2013). Problem solving by 5–6 years old kindergarten children in a computer programming environment: A case study. *Computers & Education, 63*, 87–97. https://doi.org/10.1016/j.compedu.2012.11.016.

Fisher, M., Cox, A., & Zhao, L. (2006). Using sex differences to link spatial cognition and program comprehension. *Paper presented at the software maintenance, 2006. ICSM'06. 22nd IEEE international conference*.

Franklin, D., Conrad, P., Boe, B., Nilsen, K., Hill, C., Len, M., ... Kiefer, B. (2013). Assessment of computer science learning in a scratch-based outreach program. *Paper presented at the Proceeding of the 44th ACM technical symposium on Computer science education*.

Goldberg, J. H., Stimson, M. J., Lewenstein, M., Scott, N., & Wichansky, A. M. (2002). Eye tracking in web search tasks: Design implications. *Paper presented at the Proceedings of the 2002 symposium on Eye tracking research & applications*.

Goodrich, K. (2014). The gender gap: Brain-processing differences between the sexes shape attitudes about online advertising. *Journal of Advertising Research, 54*(1), 32–43.

Grover, S., Cooper, S., & Pea, R. (2014). Assessing computational learning in K-12. *Paper presented at the Proceedings of the 2014 conference on Innovation & technology in computer science education*.

Guzdial, M. (2008). Education paving the way for computational thinking. *Communications of the ACM, 51*(8), 25–27.

Guzdial, M. (2017). Balancing teaching CS efficiently with motivating students. *Communications of the ACM, 60*(6), 10–11.

Hejmady, P., & Narayanan, N. H. (2012). Visual attention patterns during program debugging with an IDE. *Paper presented at the proceedings of the symposium on eye tracking research and applications*.

Hewig, J., Trippe, R. H., Hecht, H., Straube, T., & Miltner, W. H. (2008). Gender differences for specific body regions when looking at men and women. *Journal of Nonverbal Behavior, 32*(2), 67–78.

Horizon, M. (October 5, 2015). *Horizon Media study reveals Americans prioritize STEM subjects over the arts; science is "cool," coding is new literacy.* PR Newswire.. Retrieved from http://www.prnewswire.com/news-releases/horizon-mediastudy-reveals-americans-prioritize-stem-subjects-over-the-arts-science-is-cool-coding-is-new-literacy-300154137.html.

Hubwieser, P., Armoni, M., Giannakos, M. N., & Mittermeir, R. T. (2014). Perspectives and visions of computer science education in primary and secondary (K-12) Schools. *ACM Transactions on Computing Education (TOCE), 14*(2), 7.

Hwang, Y. M., & Lee, K. C. (2018). Using an eye-tracking approach to explore gender differences in visual attention and shopping attitudes in an online shopping environment. *International Journal of Human-Computer Interaction, 34*(1), 15–24.

Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology, 8*(4), 441–480.

Kafai, Y. B. (2006). Playing and making games for learning: Instructionist and constructionist perspectives for game studies. *Games and Culture, 1*(1), 36–40.

Kafai, Y. B., & Burke, Q. (2015). Constructionist gaming: Understanding the benefits of making games for learning. *Educational Psychologist, 50*(4), 313–334. https://doi.org/10.1080/00461520.2015.1124022.

Kafai, Y. B., & Vasudevan, V. (2015). Constructionist gaming beyond the screen: Middle school students' crafting and computing of touchpads, board games, and controllers. *Paper presented at the proceedings of the workshop in primary and secondary computing education*.

Kalelioğlu, F. (2015). A new way of teaching programming skills to K-12 students: Code.org. *Computers in Human Behavior, 52*, 200–210.

Kelleher, C., Pausch, R., & Kiesler, S. (2007). Storytelling alice motivates middle school girls to learn computer programming. *Proceedings of the SIGCHI conference on Human factors in computing system* (pp. 1455–1464). .

Lee, E., Kafai, Y. B., Vasudevan, V., & Davis, R. L. (2014). Playing in the arcade: Designing tangible interfaces with MaKey MaKey for scratch games. *Playful user interfaces* (pp. 277–292). Springer.

Lorigo, L., Haridasan, M., Brynjarsdóttir, H., Xia, L., Joachims, T., Gay, G., & Pan, B. (2008). Eye tracking and online search: Lessons learned and challenges ahead. *Journal of the Association for Information Science and Technology, 59*(7), 1041–1052.

Lye, S. Y., & Koh, J. H. L. (2014). Review on teaching and learning of computational thinking through programming: What is next for K-12? *Computers in Human Behavior, 41*, 51–61.

Maloney, J. H., Peppler, K., Kafai, Y., Resnick, M., & Rusk, N. (2008). Programming by choice: Urban youth learning programming with Scratch. *Proceedings of the 39th SIGCSE technical symposium on Computer science education ACM: Vol. 40*.

Martin, C. L., Ruble, D. N., & Szkrybalo, J. (2002). Cognitive theories of early gender development. *Psychological Bulletin, 128*(6), 903.

Master, A., Cheryan, S., Moscatelli, A., & Meltzoff, A. N. (2017). Programming experience promotes higher STEM motivation among first-grade girls. *Journal of Experimental Child Psychology, 160*, 92–106.

McDowell, C., Werner, L., Bullock, H. E., & Fernald, J. (2006). Pair programming improves student retention, confidence, and program quality. *Communications of the ACM, 49*(8), 90–95.

Meyers-Levy, J. (1986). *Gender differences in information processing: A selectivity interpretation.* Northwestern University.

Meyers-Levy, J. (1989). Priming effects on product judgments: A hemispheric interpretation. *Journal of Consumer Research, 16*(1), 76–86.

Meyers-Levy, J., & Maheswaran, D. (1991). Exploring differences in males' and females' processing strategies. *Journal of Consumer Research, 18*(1), 63–70.

Meyers-Levy, J., & Sternthal, B. (1991). Gender differences in the use of message cues and judgments. *Journal of Marketing Research,* 84–96.

Moss, G., & Colman, A. M. (2001). Choices and preferences: Experiments on gender differences. *Journal of Brand Management, 9*(2), 89–98.

Noble, S. M., Griffith, D. A., & Adjei, M. T. (2006). Drivers of local merchant loyalty: Understanding the influence of gender and shopping motives. *Journal of Retailing, 82*(3), 177–188.

NRC. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas.* The National Academies Press.

Obaidellah, U., Al Haek, M., & Cheng, P. C.-H. (2018). A survey on the usage of eye-tracking in computer programming. *ACM Computing Surveys (CSUR), 51*(1), 5.

Owston, R., Wideman, H., Ronda, N. S., & Brown, C. (2009). Computer game development as a literacy activity. *Computers & Education, 53*(3), 977–989.

Pan, B., Hembrooke, H. A., Gay, G. K., Granka, L. A., Feusner, M. K., & Newman, J. K. (2004). The determinants of web page viewing behavior: An eye-tracking study. *Paper presented at the Proceedings of the 2004 symposium on Eye tracking research & applications*.

Papavlasopoulou, S., Giannakos, M. N., & Jaccheri, L. (2017). Empirical studies on the maker movement, a promising approach to learning: A literature review. *Entertainment Computing, 18*, 57–78. https://doi.org/10.1016/j.entcom.2016.09.002.

Papavlasopoulou, S., Giannakos, M. N., & Jaccheri, L. (2019). Exploring children's learning experience in constructionism-based coding activities through design-based research. *Computers in Human Behavior* (in press).

Papavlasopoulou, S., Sharma, K., & 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*, 50–60.

Papavlasopoulou, S., Sharma, K., Giannakos, M., & Jaccheri, L. (2017). Using eye-tracking to unveil differences between kids and teens in coding activities. *Paper presented at the proceedings of the 2017 conference on interaction design and children*.

Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas.* Basic Books, Inc.

Popper, K. (1971). *The open society and its enemies* (5th ed.). Princeton: Princeton University Press.

Putrevu, S. (2001). Exploring the origins and information processing differences between men and women: Implications for advertisers. *Academy of Marketing Science Review, 2001*, 1.

Qiu, K., Buechley, L., Baafi, E., & Dubow, W. (2013). A curriculum for teaching computer science through computational textiles. *Paper presented at the proceedings of the 12th international conference on interaction design and children*.

Resnick, M., Maloney, J., Monroy-Hernández, A., Rusk, N., Eastmond, E., Brennan, K., & Silverman, B. (2009). Scratch: Programming for all. *Communications of the ACM, 52*(11), 60–67.

Robertson, J. (2012). Making games in the classroom: Benefits and gender concerns. *Computers & Education, 59*(2), 385–398. https://doi.org/10.1016/j.compedu.2011.12.020.

Robertson, J. (2013). The influence of a game-making project on male and female learners' attitudes to computing. *Computer Science Education, 23*(1), 58–83.

Robertson, J., & Howells, C. (2008). Computer game design: Opportunities for successful learning. *Computers & Education, 50*(2), 559–578.

Rosenberg-Kima, R. B., Plant, E. A., Doerr, C. E., & Baylor, A. L. (2010). The influence of computer-based model's race and gender on female students' attitudes and beliefs towards engineering. *Journal of Engineering Education, 99*(1), 35–44.

Saldaña, J. (2015). *The coding manual for qualitative researchers.* Sage.

Sangin, M., Molinari, G., Nüssli, M.-A., & Dillenbourg, P. (2008). How learners use awareness cues about their peer's knowledge?: Insights from synchronized eye-tracking data. *Paper presented at the Proceedings of the 8th international conference on International conference for the learning sciences: Vol. 2*.

Schmid, P. C., Mast, M. S., Bombari, D., Mast, F. W., & Lobmaier, J. S. (2011). How mood states affect information processing during facial emotion recognition: An eye tracking study. *Swiss Journal of Psychology, 70*(4), 223–231.

Schulte, C., & Knobelsdorf, M. (2007). Attitudes towards computer science-computing experiences as a starting point and barrier to computer science. *Paper presented at the Proceedings of the third international workshop on Computing education research*.

Searle, K. A., & Kafai, Y. B. (2015). Boys' needlework: Understanding gendered and indigenous perspectives on computing and crafting with electronic textiles. *Paper presented at the ICER*.

Sharafi, Z., Soh, Z., Guéhéneuc, Y.-G., & Antoniol, G. (2012). Women and men—different but equal: On the impact of identifier style on source code reading. *Paper presented at the IEEE 20th international conference on program comprehension (ICPC), 2012*.

Simon, S. J. (2001). The impact of culture and gender on web sites: An empirical study. *ACM SIGMIS - Data Base: The Database for Advances in Information Systems, 32*(1), 18–37.

Teague, J. (2002). Women in computing: What brings them to it, what keeps them in it? *ACM SIGCSE Bulletin, 34*(2), 147–158.

Vos, N., Van Der Meijden, H., & Denessen, E. (2011). Effects of constructing versus playing an educational game on student motivation and deep learning strategy use. *Computers & Education, 56*(1), 127–137.

Wang, M.-T., & Degol, J. L. (2017). Gender gap in science, technology, engineering, and mathematics (STEM): Current knowledge, implications for practice, policy, and future directions. *Educational Psychology Review, 29*(1), 119–140.

Werner, L., Denner, J., Campe, S., & Kawamoto, D. C. (2012). The fairy performance assessment: Measuring computational thinking in middle school. *Paper presented at the proceedings of the 43rd ACM technical symposium on computer science education*.

Wing, J. M. (2006). Computational thinking. *Communications of the ACM, 49*(3), 33–35.