Taylor & Francis
Taylor & Francis Group

International Journal of Human–Computer Interaction

ISSN: 1044-7318 (Print) 1532-7590 (Online) Journal homepage: www.tandfonline.com/journals/hihc20

# "Be a Lighting Programmer": Supporting Children Collaborative Learning through Tangible Programming System

Xinyuan Wang, Qian Xing, Qiao Jin & Danli Wang

**To cite this article:** Xinyuan Wang, Qian Xing, Qiao Jin & Danli Wang (2024) "Be a Lighting Programmer": Supporting Children Collaborative Learning through Tangible Programming System, International Journal of Human–Computer Interaction, 40:10, 2622-2640, DOI: 10.1080/10447318.2022.2163783

**To link to this article:** https://doi.org/10.1080/10447318.2022.2163783

Published online: 22 Jan 2023.

INTERNATIONAL JOURNAL OF HUMAN–COMPUTER INTERACTION
2024, VOL. 40, NO. 10, 2622–2640
https://doi.org/10.1080/10447318.2022.2163783

# "Be a Lighting Programmer": Supporting Children Collaborative Learning through Tangible Programming System

Xinyuan Wang[a,b#], Qian Xing[a,b#], Qiao Jin[c], and Danli Wang[a,b]

[a]School of Artificial Intelligence, University of Chinese Academy of Sciences, China; [b]The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, China; [c]Department of Computer Science and Engineering, University of Minnesota, USA

**ABSTRACT**

In order to cultivate children's computational thinking, researchers have developed many excellent programming systems. Among them, the tangible programming systems combined with graphic output have been widely accepted because of the intuitive input method and diversified visual feedback. However, few of them support collaborative learning or programming. Little is known about children's experiences, emotional states and behaviours on collaborative programming. Motivated by this gap, we present a novel tangible and collaborative enabled programming system named Lighters, which designed for children aged 7–10 and had two types of collaboration modes (block-based and role-based). We conducted an user experiment with 24 children, and collected physiological (EDA), questionnaire, video recording and interview data for analysis. Based on our experiment results, Lighters are effective in helping children learn to program collaboratively. In addition, Lighters can mobilize children's positive emotions and enthusiasm to learn programming. Compared with block-based collaboration, role-based collaboration is more likely to stimulate children's emotional states and has a better effect on learning programming.

## 1. Introduction

Collaborative programming is used to mean two programmers working jointly on the algorithm and code (Nosek, 1998). The previous research indicated that programmers (adult) programming collaboratively outperformed individual programmers, and they had greater confidence and enjoyed the process of problem-solving. For children, researchers and industry practitioners have paid more attention to helping them build computational thinking skills (National Research Council, 2010) and collaborative ability (UNESCO Headquarters, 1996). However, traditional programming languages contain complex grammatical knowledge and logical relation frameworks that are difficult for children to remember and understand. Researchers have carried out in-depth research on children's programming education by simplifying and visualizing the complex and abstract traditional text programming language so as to make it easy and available for children. Many computational tools for children have been designed, and a review (Yu & Roque, 2018) divided them into two main categories: tangible programming tools and graphical programming tools. Graphical programming tools transform programming concepts into graphics displayed on electronic screens, like Scratch (Resnick et al., 2009). Compared with graphical tools, tangible tools (e.g. Horn & Jacob, 2006; Hu et al., 2015; Jin et al., 2018; Schweikardt & Gross, 2006; Xing et al., 2020) are easier to use and learn for younger children because tangible interaction is better aligned with children's cognitive development ability (Horn et al., 2009; Sapounidis et al., 2019). Specially, the tangible programming systems with graphic feedback combine the advantages of physical interaction and graphical interaction, which not only ensure the natural nature of input, but also provide rich and colorful graphical feedback. Significantly, the inherent openness of tangible programming (Horn & Jacob, 2007) makes it easier for children to collaborate with each other face to face. Collaboration could reduce programming errors and anxiety (Melcer & Isbister, 2018) and improve students' interest and learning motivation (Dai, 2020; Thapaliya, 2020). They can be mutually reinforcing and give each other positive feedback. However, few existing tangible programming systems can support collaborative learning and fewer relevant user experiments analyse children's learning characteristics of collaborative programming. Motivated by this gap, we designed a tangible programming system named Lighters for children aged 7–10. It contains 65 different programming blocks in seven categories, supporting and encouraging teamwork and allowing children to create innovative lighting effects by graphical feedback.

On the other hand, different from most previous user studies of tangible programming (Deng et al., 2019; Sapounidis et al., 2019; Thieme et al., 2017; Yashiro et al.,

CONTACT Danli Wang danli.wang@ia.ac.cn School of Artificial Intelligence, University of Chinese Academy of Sciences, China
[#]These authors contribute equally to this work.

2017), the user study of Lighters not only used traditional methods, such as experimental observation, video recording, questionnaire and interview, but also innovatively introduced physiological signal technology. In recent years, physiological signals have been widely used in the study of emotion recognition in human-computer interaction, intelligent driving and other fields. Among many physiological signals, electrodermal activity (EDA) signal is effective, convenient, safe, easy to collect, and high user comfort (Posada-Quintero & Chon, 2020). It has been applied to many studies about children's emotion (Feng et al., 2018; Hernandez et al., 2014; Sohn et al., 2001). Meanwhile, one major advantage of EDA is that can provide comprehensive, objective and continuous emotional information. However, few user studies employing it to evaluate tangible programming systems with children to capture the qualitative data (e.g. emotional states). Our study utilized EDA signal for tangible programming system, combining with questionnaires, video analysis, observations, interviews to analyse their emotional states in the process of programming and collaboration.

In the user study, we defined two collaborative modes (block-based and role-based) to structure interactions based on Lighters system. Block-based collaboration is when two players manage and use different types of blocks to collaborate on programming tasks. Role-based collaboration divides the two players into two roles: "pilots" and "navigators." Pilot can use all the blocks, and navigator assists him/her and corrects programming errors (Fails et al., 2010; Thapaliya, 2020). We invited 22 children aged 7–10 and divided them into two groups: group 1 (G1) and group 2 (G2). Children in G1 adopted a block-based collaborative mode, and children in G2 adopted a role-based collaborative mode. Through the between-subject study, we aim to answer the following research questions (RQs): (RQ1) How can Lighters benefit children in programming learning? (RQ2) How can collaborative modes (the block-based collaboration or the role-based collaboration) affect children's programming learning? (RQ3) How do children's emotions change during the collaborative programming?

Our work has the following contributions:

- We designed a new tangible programming system named Lighters that offers a more diverse creative design context and an amount of blocks, which could encourage children to work together and help them learn programming knowledge and logic, such as sequences, parameters, loops, task decomposition, etc.
- We conducted a user study that used EDA physiological signal data. To our best of knowledge, this article is the first research to analyse how different collaborative modes affect children's learning effect and experience. We verified the effectiveness of the system and found that children's emotional state during using Lighters is more dynamic and active than the control condition; it is more dynamic at the start and the end of programming and tends to be flats as the programming task progresses; it is more active when children are pleased than when they are not.
- We proposed two modes of collaboration for tangible programming: block-based collaboration and role-based collaboration. By analysing multidimensional data, we got the influences from different collaboration modes of programming. We found that role-based collaboration would bring better learning outcomes than block-based collaboration in terms of learning effectiveness, emotional state, and programming experience.

## 2. Related work

### *2.1. Tangible programming*

Computational thinking is seen as a fundamental skill for everyone (Wing, 2006), which enables people to deal with problems more efficiently. Some developed countries have realized the importance of cultivating children's computational thinking, and they have included programming education in the compulsory courses (National Research Council, 2010) Many researchers have carried out in-depth research on children's programming education by simplifying and visualizing the complex and abstract traditional text programming language, so as to make it easy and available for children. The existing programming tools for children are mainly divided into two categories: graphical programming and tangible programming, such as Scratch (Resnick et al., 2009) and T-maze (Wang et al., 2011).

Tangible language means using physical objects as basic programming elements for building computer programs (Horn & Jacob, 2007). A broad set of work exists on building tangible programming language for children. Based on their physical features, they are categorized as physical and hybrid kits (Yu & Roque, 2018). Physical kits are the kits whose components are all tangible. For example, roBlock (Schweikardt & Gross, 2006) allows users to connect different programming blocks to build a robot. Its programming blocks are divided into four categories: sensors, actuators, logic, and utility. The robot obtains input through embedded sensors and has actions as output by following the program logic. ACCembly (Rocha et al., 2021) is a block-based programming system that allows children with visual impairments to assemble tangible blocks to program a multimodal robot. It consists of a map, a camera, Dash robot, and some blocks who contain three types: action, direction, and loop blocks. The camera recognizes the sequence of blocks in the workspace, interprets the connected blocks and sends the instructions to the robot. Another example is Quetzal (Horn & Jacob, 2006) that is proposed by Tufts can be used by children to control the robot "Lego mindstorm" to complete assigned tasks. Its programming language includes three types of interlocking plastic tiles that represent flow-of-control structures, actions, and data. Unlike physical tools, hybrid kits consist of both physical and virtual parts. For example, AR-maze (Jin et al., 2018) is an AR game that consists of a game app and a number of programming blocks divided into three types: start/end blocks, motion blocks, and loop blocks. Players could use tangible programming blocks to construct a predetermined path and navigate this virtual character to reach the destination.

Another example is Strawbies (Hu et al., 2015) has two main components: the app-based iPad and the wooden programming tiles. Children guide the game character Awbie on a quest for strawberries through a virtual world using wooden programming tiles. The tiles were split into several categories: verbs, adverbs, and units of measurement.

The above works show many excellent achievements in the field of tangible programming. They generally contain 2–4 kinds of tangible programming blocks/tiles (Horn & Jacob, 2006; Hu et al., 2015; Im & Rogers, 2021; Jin et al., 2018; Rocha et al., 2021; Schweikardt & Gross, 2006; Xing et al., 2020) and many of them provide predetermined task, which have a limitation of players' freedom to create. Programming freely can mobilize children's creativity and improve children's interest in programming learning (Wang et al., 2014). In addition, programming freely supports children to create more complex programs and explore diversified commands and parameters more actively (Sapounidis et al., 2019). Thus, in order to benefit children in programming learning, it is necessary to program freely. This article presents a new tangible programming tool—Lighters, which allows children to build a large range of lighting effects and to program on predetermined tasks or create freely. Our results show that this system gives children more freedom to create innovative works and benefits their programming learning.

### 2.2. Collaborative learning and programming

Collaborative learning which emerged in the United States in the early 1970s refers to a pattern where partners of a group complete a task or study together by helping and collaborating with each other (Dillenbourg, 1999) Collaborative learning differs greatly from cooperative learning, because cooperation means that partners of a group divide a task into different parts and then finish each part of them on their own (Stahl et al., 2006). So everyone acts individually and pool their results in the end. It has been shown that pair collaboration tends to increase a learner's enjoyment, engagement, and motivation, especially when children can communicate face to face and physical devices facilitate collaboration (Inkpen et al., 1999; Scott et al., 2003). The study (Xing et al., 2020) seem to indicate that collaboration can promote children's ability of decision-making and communication by comparing the single mode. Collaborative playing is important for learning programming, as it can help reduce programming errors and players' anxiety towards programming (Melcer & Isbister, 2018). Particularly, pair programming refers to students' collaboration in programming, in which two players collaborate to complete the programming tasks by respectively play a role (Begel & Nagappan, 2008). Compared with solo programming, pair programming can better increase students' interest in programming, stimulate their motivation for learning. And it could make students more concentrated, so that their programming ability is improved (Dai, 2020; Thapaliya, 2020).

Since tangible interaction offers enough open space and many objects that can be manipulated independently, tangible programming is more conducive to collaboration and socially motivated interaction (Horn et al., 2009; Sapounidis et al., 2019). Prior researches have combined tangible programming with collaborative learning (Rodić & Granić, 2022). For example, CoProStory (Deng et al., 2019) allows two children to respectively program a role on the computer screen, in order to complete a common task. This system reduces conflicts by role division from game design and a Sync Blocks which coordinating children's collaboration. Although most systems do not explicitly come up with a collaborative design, they are open for free collaboration. For example, Torino (Thieme et al., 2017) is built on physical representations that can be identified through touch, in which children with mixed-visual abilities can collaborate freely to construct a tangible program. In another study (Yashiro et al., 2017), author explored how the system named Plugramming supports collaborative learning and confirmed the effectiveness of collaborative learning.

The above works highlighted the intrinsic advantages of tangible programming for collaborative learning. However, on the one hand, we found few tangible programming systems based mobile device can support children to collaboratively learn. On the other hand, it is confirmed that free collaboration can be chaotic and ineffective (Dillenbourg & Over-Scripting, 2002). Structuring collaborative process by setting a set of instructions, such as the instructions of how the group members should collaborate, can make it more productive (Dillenbourg, 1999). Although collaborative learning is an efficient and effective way to learn programming, there is currently no systematic framework for collab orating on physical programming, so we referred to the pair programming (Thapaliya, 2020) and the user study on Mobile Stories (Fails et al., 2010) which explored the impact of different ways of cooperation (including content splitting and space sharing) on children's use of Mobile Stories. On this basis, we designed Lighters system and proposed two modes of collaboration for children's programming: block-based collaboration and role-based collaboration. Block-based collaboration is when two players manage and use different kinds of blocks, and then they complete programming tasks by working together (Fails et al., 2010). Role-based collaboration, also known as pair programming, is divided into two roles: "pilots" and "navigators." Pilot can use all the blocks, and navigator assists him/her and corrects programming errors (Thapaliya, 2020). In our work, programming blocks have been divided into seven categories and children need to work with a partner to make full use of those blocks to complete complex tasks under a certain time. We further analysed two different effects of collaboration modes (block-based and role-based) on children's programming learning effectiveness and experience.

### 2.3. Electrodermal activity (EDA)

With the development of physiological signal acquisition and processing technology, emotion recognition based on physiological signals has been applied to human-computer interaction, intelligent driving, entertainment, education and biomedical clinical (Wu et al., 2018). Physiological signals are directly controlled by the autonomic nervous system and endocrine system, which have shown sensitivity to change of emotions and are rarely influenced by subjective emotions. Therefore, it is more objective and convincing to use physiological signals to analyse emotional changes. At present, the widely used physiological signals include ECG, EDA, EEG, EMG, etc. to identify and analysis emotions (Alarcao & Fonseca, 2017; AlZoubi et al., 2012; Jovic & Bogunovic, 2011; Udovičić et al., 2017; Yin et al., 2017; Zontone et al., 2019). For example, Al-Zoubi et al. (2012) studied emotion recognition using ECG, EMG and EDA signals, and adopted two methods of feature selection and nine kinds of classifiers to identify eight emotions (boredom, confusion, curiosity, delight, flow/engagement, frustration, surprise, and neutral). Noroozi et al. (2019) presented an integrated analysis tool, SLAM-KIT, to analyse rich multimodal data (e.g. EDA, Eye-movement, facial expression, heart rate) in the regulatory processes of collaborative learning. Lee-Cultura et al. (2021) explored children's play and problem solving behavior by analysing Multi-Modal Data (MMD), including HRV, EDA, gaze, and skeleton. And they (Lee-Cultura et al., 2021) investigated the relationship between three degrees of Avatar Self-Representation (ASR) and children s affective and behavioral states.

Particularly, since EDA is an effective and non-invasive way to identify emotions, it has been widely used in the study of children groups. Nikolić et al. (2018) collected the data of 117 children's EDA and HR/HRV to assess whether they had social anxiety disorder (SAD). Hernandez et al. (2014) leveraged the use of a wearable EDA sensor to recognize ease of engagement of children during a social interaction with an adult. Don C. Fowles et al. (2000) recorded EDA data from 92 4-year-old children during a laboratory session that encompassed physiological and psychological stimuli. Their aim was to examine the patterns of young children's electrodermal reactivity and test the hypothesis that EDA reflects individual differences in the behavioral inhibition system (BIS). For children with autism spectrum disorders (ASD), researchers explored these children's electrodermal and behavioral responses to sensory and repetitive stimuli (McCormick et al., 2014; Schoen et al., 2008) and investigated their autonomic nervous system response to anxiety (Kushki et al., 2013) based on EDA data.

Although physiological signal is a great measurement to detect emotions, few researches have applied it to children's collaborative programming learning. The previous related studies (Deng et al., 2019; Sapounidis et al., 2019; Thieme et al., 2017; Yashiro et al., 2017) usually took questionnaires, experimental observations and interviews as data collection methods. They identified children's emotional changes by analysing their facial expressions, gestures, body movements and other nonverbal behaviors, but these behaviors do not always reflect children's real and continuous emotions. Therefore, we elected to collect EDA signals to analyse the characteristics of children's emotional changes in different collaboration modes during the whole programming process, to explore explore how children's emotions change during the collaborative programming process.

## 3. System description

Lighters is a tangible programming tool with a storyline designed for children aged 7–10. The story is about a group of animals living happily in a large forest. As a new lighting engineer, the little squirrel must complete a series of designs and eventually grow into an excellent lighting engineer. When children use Lighters, they play the role of the little squirrel and use programming blocks to complete the designs of lighting effects.

Our goal is to promote children's collaborative learning, stimulate children's interest in programming, and let them learn programming knowledge such as sequences, parameters, loops and task decomposition. To achieve this goal we created a system which consists of three parts (Shown in Figure 1): Tangible Programming Blocks which are used to construct a program; Lighters Application on mobile devices that display the programming effect and feedback; Iron Whiteboard which provides available programming space.

### 3.1. Tangible programming blocks

Tangible Programming blocks (Figure 2) are made of low-cost paper-based cards with magnets in their back, which children can easily assemble a program on the Iron Whiteboard. Every block has a TopCode (Horn, 2012) texture. In order to offer clear clarifications of the program metaphor, we added pictures and text to indicate the functionality of the blocks. According to the required functions, we designed seven types of code tokens:

- Start/End Blocks are used to indicate the start or end of the program.
- Select Blocks are used to select target lights, for example, select the lights in row 1, column 2 or the lights in rows 1 to 3, columns 3 to 6;
- Color-Blocks are used to set the target lights' colors;
- Shape-Blocks are used to set the target lights' shapes;
- Loop Blocks can be used to optimize a duplicated structure, whose functionality is similar to the logic of "for" loop in traditional programming language.
- Parameter Blocks contain numbers, figures and Chinese characters. They are used with Select Blocks, in order to

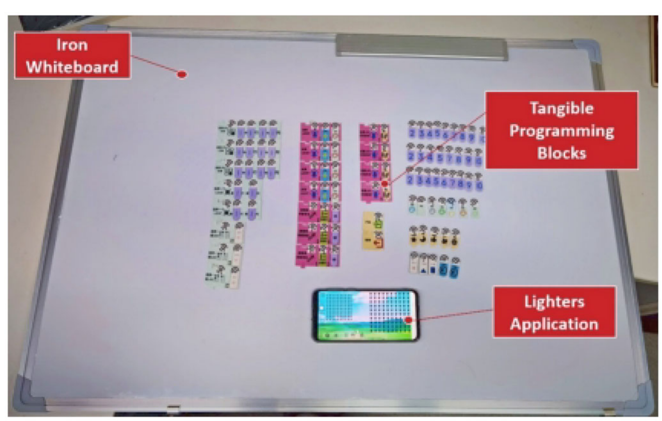



Figure 1. Overview of Lighters system

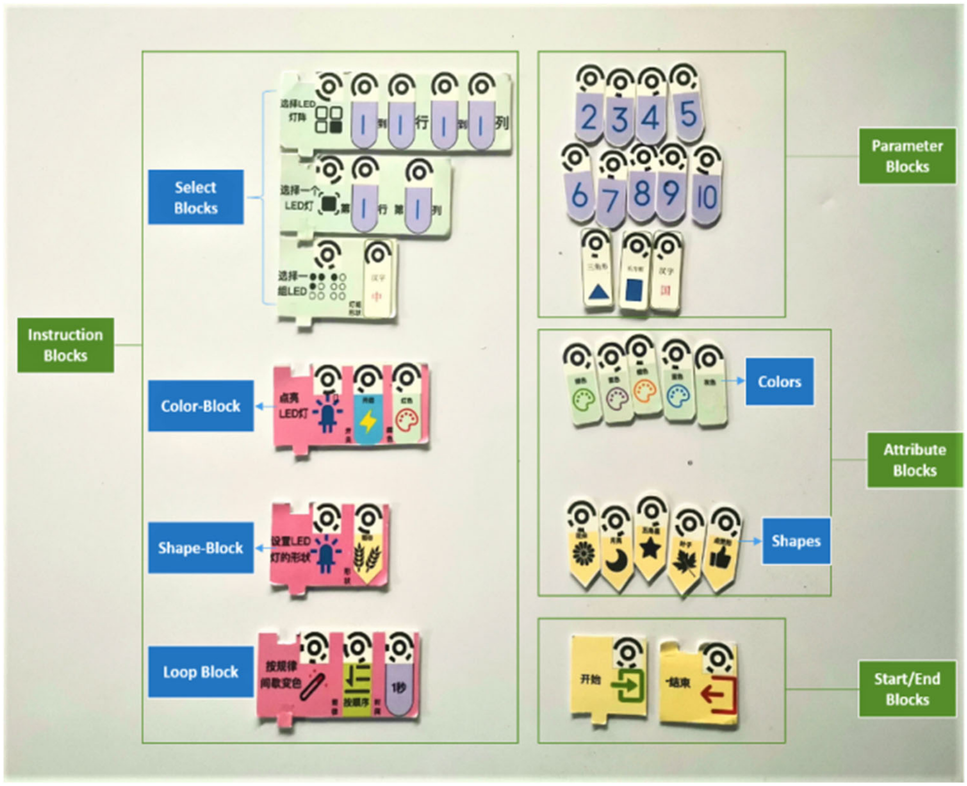



**Figure 2.** Tangible Programming blocks.

make a more diverse selection of target lights, and let users understand some concepts such as point, array, Chinese characters, figures, etc.

- Attribute Blocks contain Colors and Shapes. They are used separately with Color-Blocks and Shape-Blocks, which creates more diverse programming effects.

Among them, Select Blocks, Color-Blocks, Shape-Blocks and Loop Blocks are collectively called Instruction Blocks. The programming blocks have a variety of the classification and semantics, which could support to achieve colorful lighting effects and an amount of programming logic. There are 65 blocks in seven types, which are easy for children to understand, benefiting from its' clear classification and pictures indicating functionality. To encourage children's active collaboration and compare the effect of collaboration modes, we designed complex experimental tasks which the players need to use all types of blocks to complete. It is difficult for a child to complete all tasks in the allotted time, but easier for collaborative programming. We set different collaborative modes in the user study: block-based collaboration (one child takes four kinds of Instruction Blocks and another one holds Parameter blocks and Attribute Blocks) and the role-based collaboration (One child acts as "pilot" who led the programming and another one acts as "navigator" to assist this process).

### 3.2. Lighters application

The Lighters application is implemented by Android. As a virtual environment, it provides users with an interesting programming background and a rich programming effect. Users would be able to read the story background and get tutorials of how to complete tasks. The system mainly contains two stages: the programming stage and the running stage (Figure 3).

In the programming stage, users can click the "Task Description" button on the application to read through the task description, or the "Task Preview" button to preview the target lighting effect on the small light array on the left of the interface. Then they can manipulate the tangible blocks to create a sequence to program on the whiteboard and finish the task following the task preview. When the programming is done, the users click the "Identification Code" button (Figure 4) to identify the program by taking a picture of the sequence. If the program is correct, the application would send Augmented Reality (AR) feedback on the semantics of each block (Figure 4(a)). Then the user can click "OK" to return and execute the program. Otherwise, the application would show feedback on the specific location of the errors. Then the user could modify and debug the program according to the error prompt (Figure 4(b)). In the running stage, users click the "Program Test" button on the application to run their program with animation effect if the program is consistent with the preview effect. After that, users are able to click the "Submission" button to enter the next level.

Lighters can be used by a single person, and it is also able to be used collaboratively for the following reasons: the intrinsic properties of tangible user interface allow children to operate the physical objects collaboratively; the available programming space on the whiteboard is enough for $2\sim3$ children, children do not need to share input devices like the mouse or keyboard as with graphic programming; the number of tangible programming blocks is relatively large

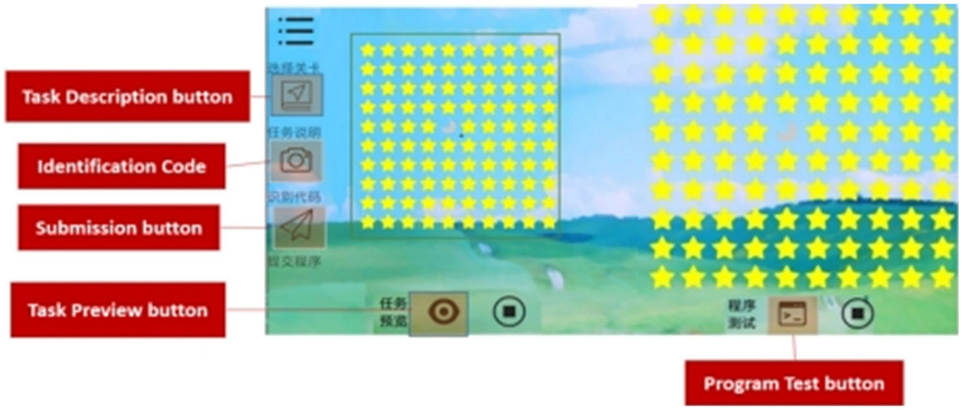
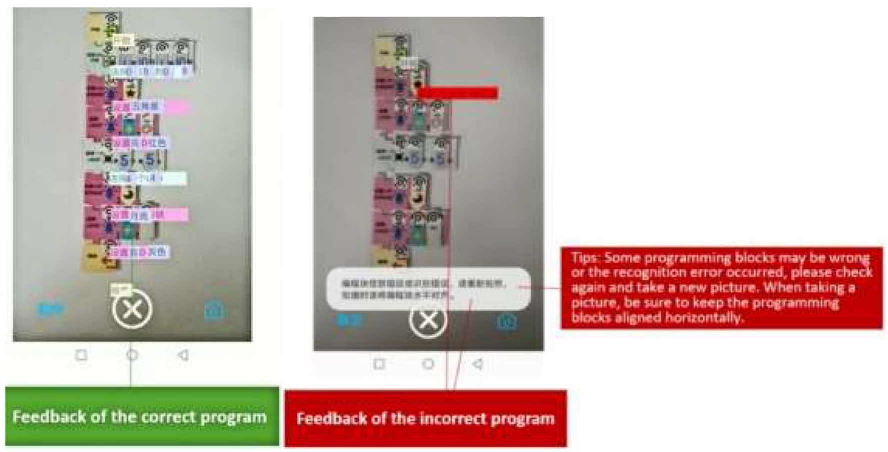



Figure 3. User interface.




(a) Feedback of the correct program; (b) Feedback of the incorrect program

Figure 4. Identification Code and Feedback.

than other tangible tools and the function classification is clear and easy for them to make the job division. If the children want to, they can design and achieve a new lighting effect by giving full play to their creativity without having to adhere to strict task requirements. Also, programming tasks become more difficult as the levels increase, which increases the fun and challenge as a collaborative game.

### 3.3. Iron Whiteboard

The size of Iron Whiteboard is 60 cm * 90 cm, which is not only low-cost but also provides a big enough space for programming. The Magnet-based programming blocks are used in conjunction with the iron-board, which keeps the blocks in place when children are wiggly.

## 4. Methodology

### 4.1. Study setting

The study was conducted in a spacious and quiet room on November 28th and 29th, 2020. In addition to the necessary tables, chairs, and two researchers, the experimental room also had two video recorders, two laptops, a group of MP160 physiological recorders, a mobile phone capable of running Lighters application and a set of programming blocks, as shown in Figure 5. The two video recorders were placed on the front and left of the children to replay the whole session and explain anomalies. The video was obtained with the consent of the children and their parents and will not be made public. One laptop was used to teach children how to operate the programming blocks, and the other laptop combined with MP160 was to record EDA signals.

During the experiment, we divided the participants into two groups, and they completed the experimental tasks in different collaborative modes (the block-based collaboration and the role-based collaboration) after they mastered the use of Lighters. Through analysis of observation, questionnaires, and EDA signals. We were interested in understanding the learning performance and emotional changes during the collaborative programming of children who programmed with Lighters and investigating the influences of collaborative modes.

### 4.2. Participants

A total of 24 children (11 of the participants were female while 13 were male; recruited from a local primary school in Beijing, China) aged 7 to 10 years old joined this study.

Participation was voluntary and uncompensated. Participants reported having a wide variety of collaboration experiences, from on average a few times a month to more than once a week. 12 participants reported having prior programming experience, and 12 reported having no prior programming experience. Before starting, the researchers were given informed consent, which includes a brief introduction to the Lighters system and the experiment, as well as the data we will collect in the experiment. We emphasized that we would collect EDA data in the experiment and that the child or their parents could stop the experiment at any time. Children could participate in the experiment only after they and their parents had a clear knowledge of the above contents and signed the informed consent. To protect anonymity, we assigned each child a participant code.

Due to personal scheduling issues among them, two children failed to form a team to collaborate. The other 22 children were divided into two groups called G1 and G2, and the demographic information was as shown in Table 1. The children in G1 used the block-based collaboration, in which one child held four kinds of Instruction Blocks; another child held Parameter blocks and Attribute Blocks. So they needed to collaboratively complete tasks by combining different blocks. The children in G2 used role-based collaboration, which included two roles: "pilot" and "navigator." In Task-One, the first child as the "pilot" led the programming and he/she could have access to all types of programming blocks, and the second child as the "navigator" assisted the first child and corrected his/her errors. In Task-Two, they switched roles.

### 4.3. Experiment procedure and task design

The whole experiment (Figure 6) takes about 50 minutes, which is divided into eight stages: experimental teaching (8 minutes), doing practice task (5 minutes), writing a pre-test questionnaire (4 minutes), wearing the acquisition equipment of EDA (3 minutes), conducting the control condition (Quiet-


Figure 5. Experimental environment (two children and one researcher).

Table 1. Specific information of participants ($N = 24$).

| Group ID | Age | Gender | Programming Experiences | Collaborative Experiences |
|---|---|---|---|---|
| G1 | Range = 7–9, Mean = 8.33, SD = 0.49 | 8 boys<br>4 girls | Yes: 5<br>No: 7 | Little: 4<br>Much: 8 |
| G2 | Range = 7–10, Mean = 8.10, SD = 0.99 | 4 boys<br>6 girls | Yes: 2<br>No: 8 | Little: 3<br>Much: 7 |

Children in G1 used block-based collaboration and children in G2 used role-based collaboration.

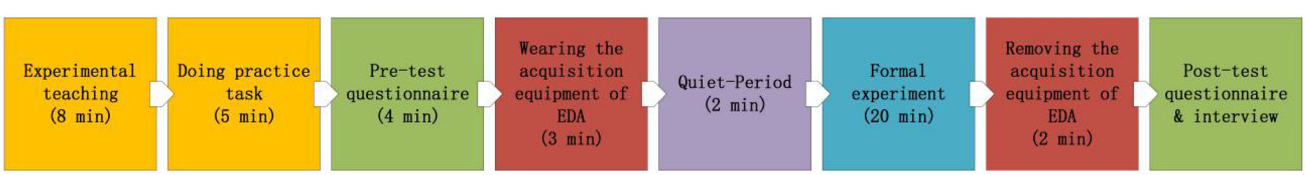



Figure 6. Experimental process. The whole experiment takes about 50 minutes, which is divided into eight stages. As shown by the arrows, the order of the whole experiment is from left to right.

Period) (2 minutes), formal experiment (20 minutes), removing the acquisition equipment of EDA (2 minutes), writing a post-test questionnaire, and interview (5 minutes).

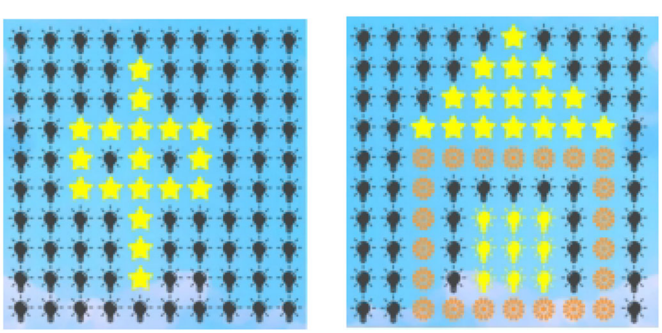
At the beginning of the study, we combined the recorded teaching videos with on-site guidance from one of our researchers to teach children how to program by using the system. Then, the children were allowed to explore the system freely and they could ask researchers anything they did not understand.

This was followed by the pre-test questionnaires for the children to collect subjective data and objective data of programming ability, including three parts: the first part is to collect children's demographic information including gender, grade, experimental code, Group ID; the second part is the background of children's programming and collaborative experience; the third part tests the children's programming ability. The ability test consisted of two questions that presented two lighting effects and the corresponding program (containing errors).

After doing a pre-test questionnaire and getting the permission from the children, we helped them wear the acquisition equipment of EDA and connected the MP160 physiological recorder to record EDA signals. Then, A control condition (Quiet-Period) was conducted where participants performed an EDA test without playing anything. Children needed to sit and rest for two minutes in order to get the reference signal of EDA.

The two minutes control condition was followed by a formal experiment in which the children were asked to cooperate on two formal tasks. Children needed to collaboratively complete two formal tasks. If they have enough time after finishing formal tasks, they can do a free-play task. Researchers observed and recorded children's programming performance in the whole formal experimental stage. Researchers didn't interfere with the children's behavior unless there was a problem with the experimental equipment or the children asked for help. We used two video recorders to record children's status during collaborative programming and used MP160 to record their EDA signals. The two specific tasks of the formal experiment are as follows:

1. Task-One: Select a number of lights in the virtual panel of APP, which are combined to form a Chinese character "中" (Figure 7(a)). All of the selected lights are lit in the shape of a star, and they change color at a one-second interval.
2. Task-Two: Design a house by selecting lights (Figure 7(b)). The roof is a triangular shape composed of star-shaped lights, and they can change color in cycles. The wall is a rectangular structure composed of orange lights that are lit in the shape of flowers. The door is a square made up of some yellow lights that are lit in the shape of bulbs.
3. Free-play task: After completing two official tasks, children can design and achieve a new lighting effect by giving full play to their creativity.



(a) Task-One (b) Task-Two

Figure 7. Task preview.

After the children completed experimental tasks or programming was over 20 minutes, we removed the acquisition equipment of EDA and asked them to complete a post-test questionnaire.

The post-test questionnaire also included three parts: the first part still collected children's personal information; the second part was to investigate the feelings of children, such as whether they were satisfied with collaboration, whether they could learn knowledge through playing games, and how was their emotional state during the experiment; the third part still tested the children's programming ability with the same questions as the pre-test questionnaire.

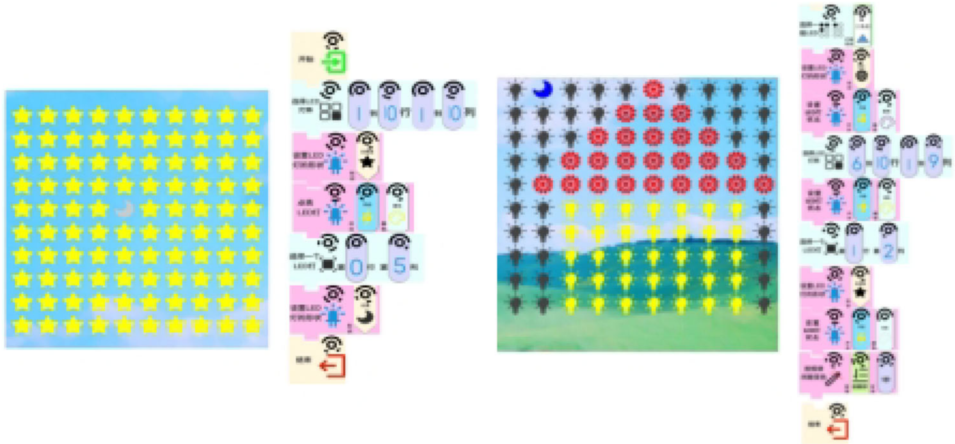
Finally, they were interviewed by one of our researchers.

### 4.4. Instruments

We used the MP160 physiological recorder and a Windows laptop computer equipped with ACQKnowledge software to collect children's EDA signals. In this experiment, we set the thenar parts of the palm as the points collecting EDA signal. We used the Event Mark function of software to mark the special time point, such as the start/end of Quiet-Period and each task, and the point when children's emotions obviously fluctuates. We used two video recorders to record our observations and interviews. We recorded children's programming abilities through pre-test and post-test.

### 4.5. Measures and data collection

This study used a mixed methods approach. Both qualitative and quantitative data were collected for analysis. Our goal is to have a holistic view of participants' behaviors and experiences under different technologies by programming ability test, observations, Electrodermal Activity (EDA) and final interviews to collect valuable feedbck in order to better answer our RQs. To answer the research questions, we used a mixed method to conduct our study with Lighters. This mixed approach was adopted because previous studies (Deng et al., 2019; Sapounidis et al., 2019; Thieme et al., 2017; Yashiro et al., 2017) usually adopts questionnaires, experimental observations and interviews as data collection, but this method is not convincing. Due to the physical contact of children in the process of using lighter, among many physiological signals, electrodermal activity (EDA) is effective, convenient, safe, easy to collect, and user comfort is



(a) The first question (b) The second question

**Figure 8.** The ability test in pre-test and post-test.

**Table 2.** The variables observed during children's programming and related description.

| Variable | Description |
|---|---|
| Time of completing tasks | It is the time from when a child drops the first programming block to when the program is finished and runs successfully. |
| Difficulty of free-play task | Whether a child can complete the free-play task and how complex the task is (number of programming blocks used). |
| State of children's emotion | It includes children's emotions, such as pleasure/unpleasure. |
| Status of collaboration | It refers to children's communication and collaborative characteristics when completing tasks, such as communicating with each other frequently and smoothly or blaming each other for differences in programming. |
| The time when they ask for help | Children would ask the researchers for help when they encounter problems that cannot be solved through communication and collaboration, such as the unidentified programming block and problems of programming logic. |

high, which is more suitable for this study. The data analysis and mixed method were carried out offline.

#### 4.5.1. Programming ability test

To measure the children's programming ability before and after formal experiment, we conducted a two-task programming ability test on children in Lighter system. The ability test consisted of two questions that presented two lighting effects and the corresponding program (containing errors) (Figure 8). The first question showed the "many stars twinkle around the moon" lighting effect and the corresponding program containing two errors (Figure 8(a)). The second problem the "moon hanging on the roof of a house" lighting effect and the corresponding program containing four errors (Figure 8(b)). It examined whether children could find errors in the program by using the programming knowledge they had.

#### 4.5.2. Observation

In order to observe and record children's performance in a more careful and comprehensive way, we used a record table and video recorders to collect the data. The video was obtained with the consent of the children and their parents and will not be made public. During experiment, video recorders are used to record children's performance. In the following analysis, we can analyse and excavate the experimental details by playing back these videos. We also observed children's programming behaviours based on the video and used a record table to capture the following variables (Table 2).

#### 4.5.3. Electrodermal activity (EDA)

Among many physiological signals, electrodermal activity (EDA) signal, also called Galvanic Skin Response (GSR) or Skin Conductance (SC) (Lagopoulos, 2007), is effective, convenient, safe, easy to collect, and high user comfort and a commonly used measurement index in emotion recognition. So it is applied to many studies about children's emotion. Meanwhile, it has one obvious advantage that it can provide comprehensive, objective and continuous emotional information. As one's emotions fluctuate, blood vessels in the skin open and contract. The sweat gland on the skin becomes active and secretes sweat through pores, leading to measurable changes in skin conductance, known as EDA. When people's emotions are induced, the emotional arousal becomes higher, the sweat secretion becomes more, and the corresponding EDA increases (Shukla et al., 2019). We can measure EDA in areas where sweat is concentrated, such as the fingers, palms, or soles of the feet. Then the degree of arousal and activation of the sympathetic nervous system can be inferred by analysing the feature data of EDA. It should be noted that pleasure refers to a positive emotion, expressed as laughing, making happy sounds, or motion.

Unpleasure refers to an emotional state with less joy and more tension, because children focused on programming tasks (Xing et al., 2020). There are significant differences in the EDA signals of people. In order to accurately analyse the relationship between EDA signals and children's emotions, we eliminated individual differences. The standardized value of EDA is that the value of the experimental stage subtracted the value of the Quiet-Period.

#### 4.5.4. Interviews

For each child, the topics of the interview were the same and as follows: (i) What you can learn from Lighters; (ii) Do you support cooperative learning and why; (iii) Do you feel happy when you using Lighters; (iv) Which of the two modes of collaboration is better/worse? Why?

### 4.6. Data analysis

This study proposed a user evaluation method based on EDA signal for tangible programming system. In the study, we collected EDA data of children, combining with questionnaires, observations, interviews to analyse their emotional states in the process of programming and collaboration.

We used SPSS to analyse the quantitative data, such as the data from the programming ability test, the time to complete tasks, the number of times asking for help, etc. For participants' EDA signal, the combination of ACQKnowledge and SPSS was used to analyse it. The effective frequency range of EDA signals is 0.02 Hz∼0.2 Hz, so we carried out a low-pass filter of 0.9 Hz before data analysis, which removed the interference from EMG signals and children's movements. We performed a down-sampling from 2000 sample points/s to 8 sample points/s to make the EDA curve smoother and more convenient for analysis. We used the ACQKnowledge to extract the following feature data of EDA: Event-Count, Event-Amplitude, Mean, and Slope.

In order to verify whether the Lighter improved the children's programming ability, we counted the number of errors found by the children on the programming ability test. In order to research emotional changes of children when they used Lighters, we divided the whole EDA into three stages: Quiet-Period, Task-One, and Task-Two. Due to time constraints, some children did not do free-play tasks and could not serve as a control group, so we did not statistically analyse EDA signals at this stage. The Quiet-Period was 120 s, and we selected the middle 60 s to analyse to avoid the situation where children's inability to calm down at the beginning impacted the accuracy of EDA data. However, we selected all the EDA data of Task-One and Task-Two. In order to explore how collaborative modes have affected children programming, we compared the differences in EDA data and programming ability tests between the G1 and G2 groups.

More specifically, we quantitatively analyzed the characteristics of EDA and pre and post-test. We analyzed the interview qualitatively. For the observations, we analyzed time of completing tasks and the time when they ask for help qualitatively, and we analyzed difficulty of free-play task, state of children's emotion and status of collaboration quantitatively. We fused those data to get consistent results. Thus, the data was mixed during presentation.

## 5. Results

### 5.1. RQ1: How can lighters benefit children in programming learning?

The data of programming ability tests before and after the experiment are shown in Figure 9. In the pre-test before collaborative programming, children can find 1.17 errors (SD = 0.59) in the first question on average and find 1.87 errors (SD = 0.28) in the second question. In the post-test after collaborative programming, children can find 1.92 errors (SD = 0.18) in the first question on average and find

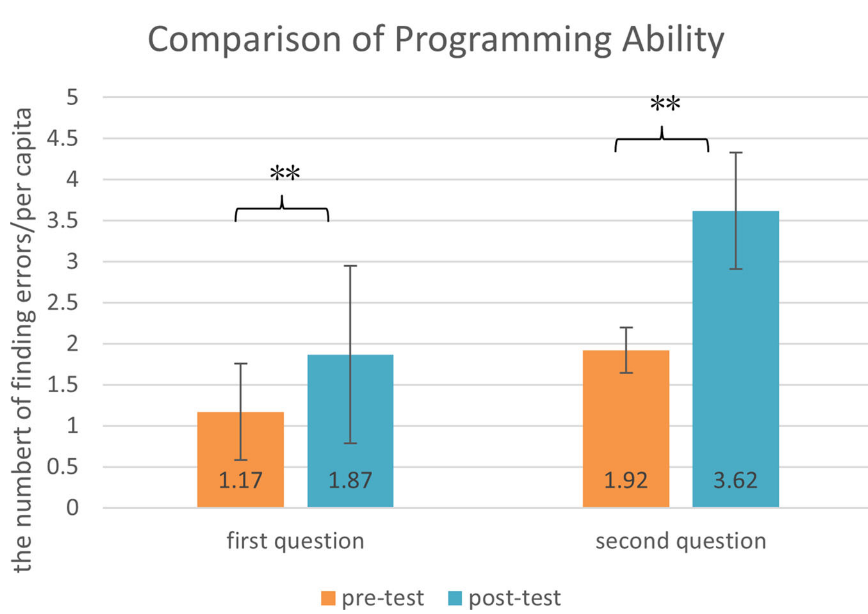



Figure 9. The data of programming ability test.

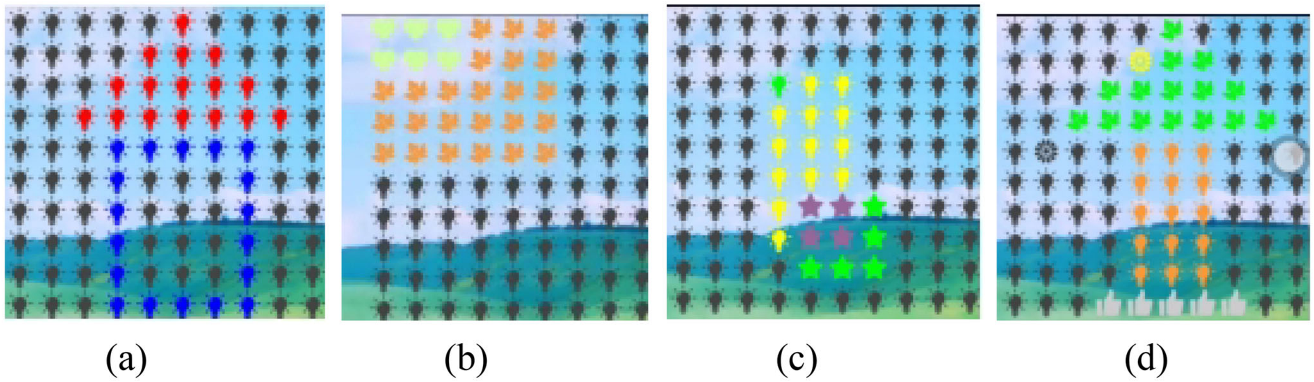



Figure 10. Some effects of free-play task.

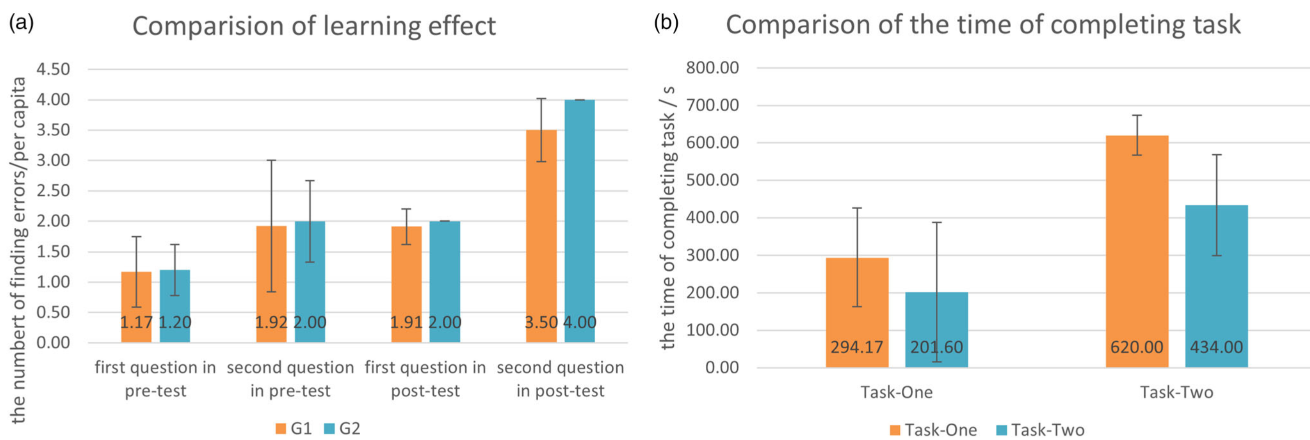



Figure 11. The comparison of different collaborative modes.

3.62errors (SD = 0.71) in the second question. Then we performed the Wilcoxon test with the pre-test/post-test as variables. The results showed that there was a significant difference between the programming ability of pre-test and post-test ($p = 0.000$ in the first question and $p = 0.000$ in the second question).

On the other hand, we analysed children's performance in the free-play task and the difficulty of the task. 64% of children completed the free-play task in which they used 6 ~ 11 programming blocks to create a new lighting effect. In the interview, we asked children how they thought about the "free-play task." Children said, "I like the free-play task very much!" "It allows me to give play to my own creativity and complete a new lighting effect with my friends!" "It gives me enough freedom!" Some effects of the free-play task are shown in Figure 10. It showed that children have mastered the corresponding programming knowledge, and they liked using it to create freely. The post-test questionnaire had the question "Whether you learned some programming concepts by using Lighters, such as sequences, parameters, loops, task decomposition, etc. Did you grasp them?". The data of all children's questionnaires showed that 16 children said they had learned and mastered all of them completely, and 6 said they had learned part of them.

We could find that children's programming ability improved significantly after collaborative programming by using Lighters. Combining the free-play task and subjective data of questionnaire, we could find that Lighters could enable children to learn programming knowledge through playing games.

#### *5.1.2. RQ2: How have collaborative modes (the block—based collaboration or the role-based collaboration) affected children's programming?*

We explored how collaborative modes have affected children programming from three aspects: learning effect, emotion and learning experience. About learning effect, we conducted both quantitative and qualitative analysis on the testing data of programming ability from G1 and G2, and took the G1/G2 (Group ID) as variables to conduct the Mann-Whitney U test. The result is shown in Figure 11(a). For the first question in the pre-test questionnaire, the child in G1 can find 1.7 errors (SD = 0.58) on average and the child in G2 can find 1.20 errors (SD = 0.42) on average. For the second question in the pre-test questionnaire, the child in G1 can find 1.92 errors (SD = 1.08) on average and the child in G2 can find 2.00 errors (SD = 0.67) on average. For the first question in the post-test questionnaire, the child in G1 can find 1.91 errors (SD = 0.29) on average and the child in G2 can find 2.00 errors (SD = 0.00) on average. For the second question in the pre-test questionnaire, the child in G1 can find 3.50 errors (SD = 0.52) on average and the child in G2 can find 4.00 errors (SD = 0.00) on average. The results of the Mann-Whitney U test proved that there was no significant difference between G1 and G2 before

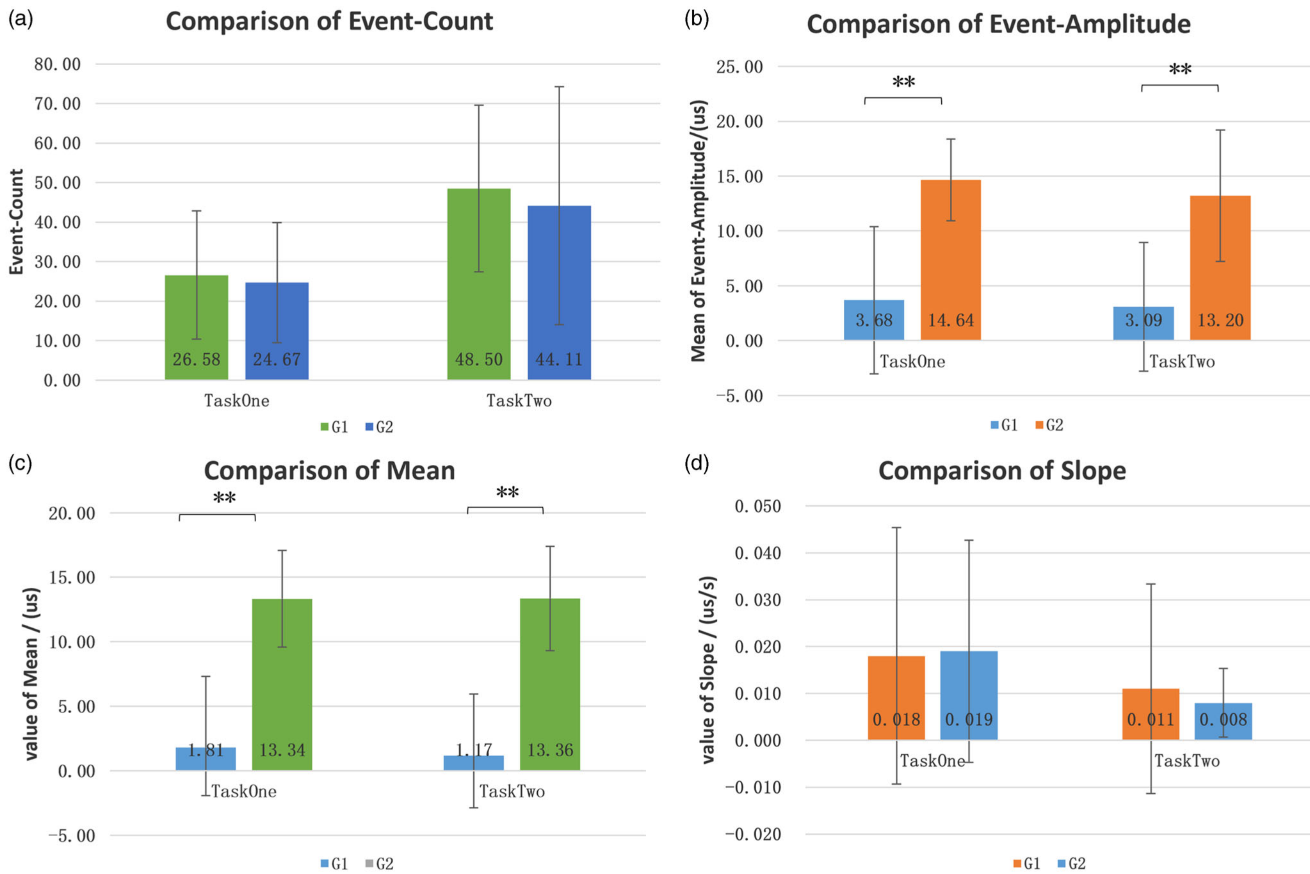



**Figure 12.** The comparison of EDA according to the task.

collaborative programming ($P_{first} = 0.93$, $P_{second} = 0.78$). After collaborative programming, there was also no significant difference between G1 and G2 in the first question ($p = 0.36$), but there was a significant difference in the second question ($p = 0.01$). After further analysis, we found that the first question was relatively easy and almost all the children could answer it correctly after collaborative programming. However, the second question involves more programming concepts and more complex logic, so it can better reflect the differences of children's programming ability. The question "Whether you learned some programming concepts by using Lighters, such as sequences, parameters, loops, task decomposition, etc. Did you grasp them?" was set in the post-test questionnaire. There were 67% of children in G1 to answer that they had learned and mastered all of them completely, but there were 80% in G2.

We analysed the time of completing Task-One and Task-Two by taking Group ID as variables, and we give up analysing the time of the free-play task because some children did not have enough time to complete the free-play task. The result is shown in Figure 11(b). For the time of completing Task-One, the child in G1 took 294.17 s (SD = 131.70) on average and the child in G2 took 201.60 s (SD = 53.51). For the time of completing Task-Two, the child in G1 took 620.00 s (SD = 185.58) on average and the child in G2 took 434.00 s (SD = 134.09). We conducted an Mann-Whitney U test on the time of completing tasks and found that there was no significant difference in the time of completing Task-One and Task-Two between G1 and G2 ($P_{Task\text{-}One} = 0.36$, $P_{Task\text{-}Two} = 0.14$).

**Table 3.** The results of Mann-Whitney U test.

| | Event-Count | Event-Amplitude | Mean | Slope |
|---|---|---|---|---|
| Task-One | $p = 0.945$ | $p = 0.002^{**}$ | $p = 0.001^{**}$ | $p = 0.917$ |
| Task-Two | $p = 0.917$ | $p = 0.002^{**}$ | $p = 0.000^{**}$ | $p = 0.227$ |
| Pleasure | $p = 0.831$ | $p = 0.008^{*}$ | $p = 0.000^{**}$ | $p = 0.014^{*}$ |
| Unpleasure | $p = 0.479$ | $p = 0.003^{**}$ | $p = 0.000^{**}$ | $p = 0.013^{*}$ |

We analysed the times of asking for help in G1 and G2. It was suggested that the average times of asking for help in G1 was 4.17 (SD = 2.14), and it in G2 was 2.80 (SD = 2.59). Although there was no significant difference between the two groups ($p = 0.36$), the times of G2 are significantly less than the times of G1 (Figure 12).

In order to research how collaborative modes affect children's emotion during programming, we compared the children's EDA signals of the two groups in two ways: task and emotion, and we still extracted four kinds of feature data of EDA. According to the task, we analysed and compared children's EDA signals of the two groups while doing Task-One and Task-Two after eliminating individual differences. We found that the Event-Amplitude and Mean of EDA in G2 were much larger than those in G1. And there were

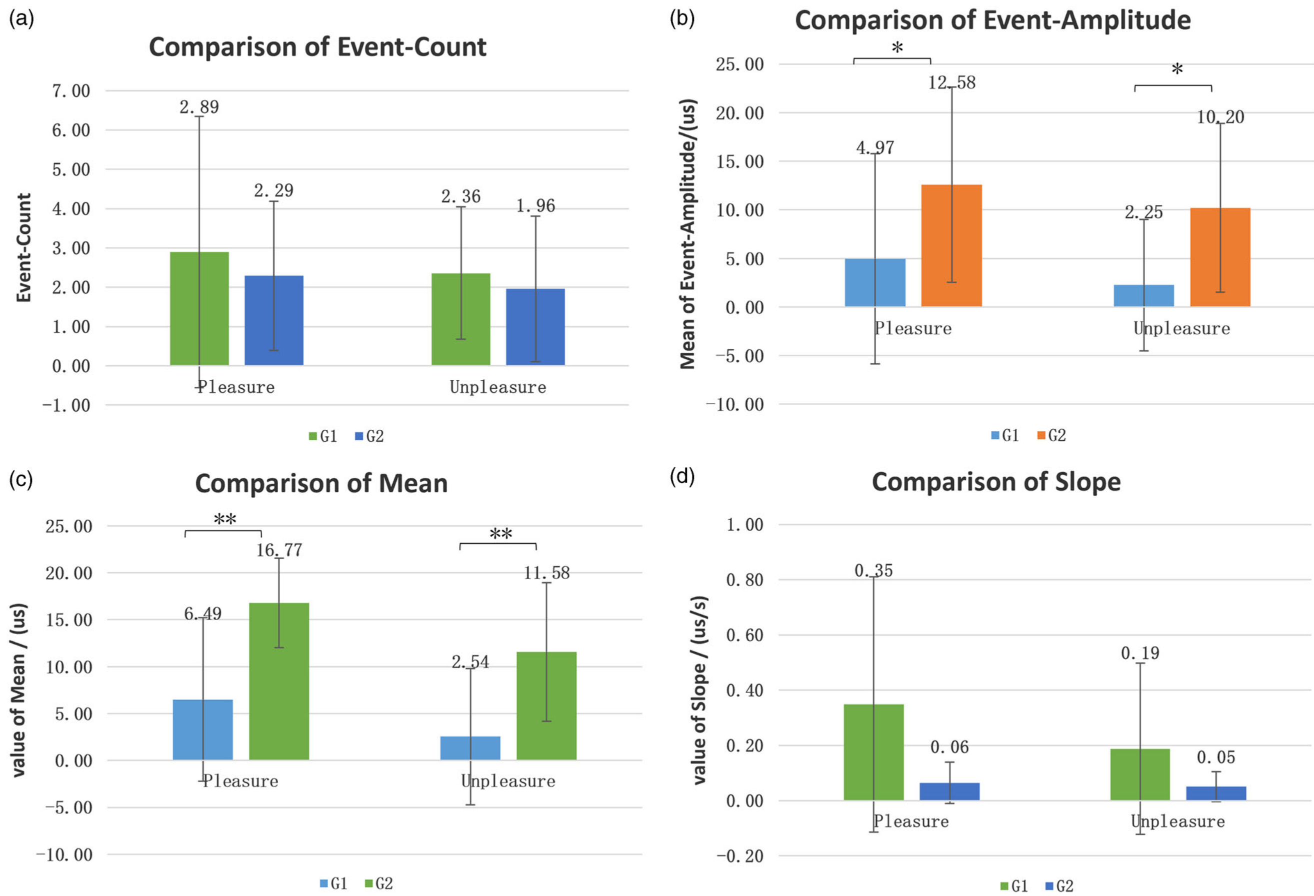



**Figure 13.** The comparison of EDA according to the emotion.

significant differences between the two groups, and the specific results of Mann-Whitney U test are shown in Table 3. However, their Event-Count and Slope of the two groups were almost the same, and there was no statistical difference. The specific data analysed above are shown in Figure 13.

According to the emotion, we respectively analysed the two emotional states of the two groups: pleasure and unpleasantness. The result of analysis suggested that in the two emotional states, the Event-Amplitude and Mean in G2 were larger than those in G1 as shown in Figure 13(b, c), and there were significant differences between the two groups as shown in Table 3. However, the Slope in G2 was always smaller than that in G1 as shown in Figure 13(d), and there was also a significant difference as shown in Table 3.

Above, we took the Group ID as variables to compare and analyse the learning effect, the time of completing the task, the times of asking for help and the EDA signals. The result proved that there was almost no difference between the two groups before the experiment, but the children in G2 showed a better effect in learning and more significant improvement in programming ability than the G1 after the experiment. The time of completing tasks and the times of asking for help in G2 were fewer than those in G1. On the other hand, when we analysed the EDA signals, we can find that during collaborative programming, children in G2 had a higher degree of arousal and activation of the sympathetic nervous system and a more active emotional state than G1, whether by task or by emotion.

In order to research how collaborative modes affect children's learning experience, we analysed the interview process in which we communicated with the children. They offered several reasons to support that they thought collaboration could help them learn. "When we are in trouble, we can ask each other." "My teammate can help me when I'm confused." "Collaboration can enhance the effect of programming." "We can do our own skilled parts, so that we can better complete the tasks." "I can ask my partner when I do not know what to do." "Collaboration brings me a greater sense of achievement." However, we found that one child in G1 believed that collaboration had a negative effect on his learning. He said that his teammate always gave him some wrong guidance. We checked the experimental video and found that he and his teammate often had different ideas and even had disputes because of the placement of programming blocks.

To sum up, most children liked collaborative programming, and the role-based collaboration (G2) of using may promote better learning effectiveness. In G2, one child led the programming, and the other child could correct mistakes and help with programming, so the programming idea is coherent, and the children could complete the tasks faster

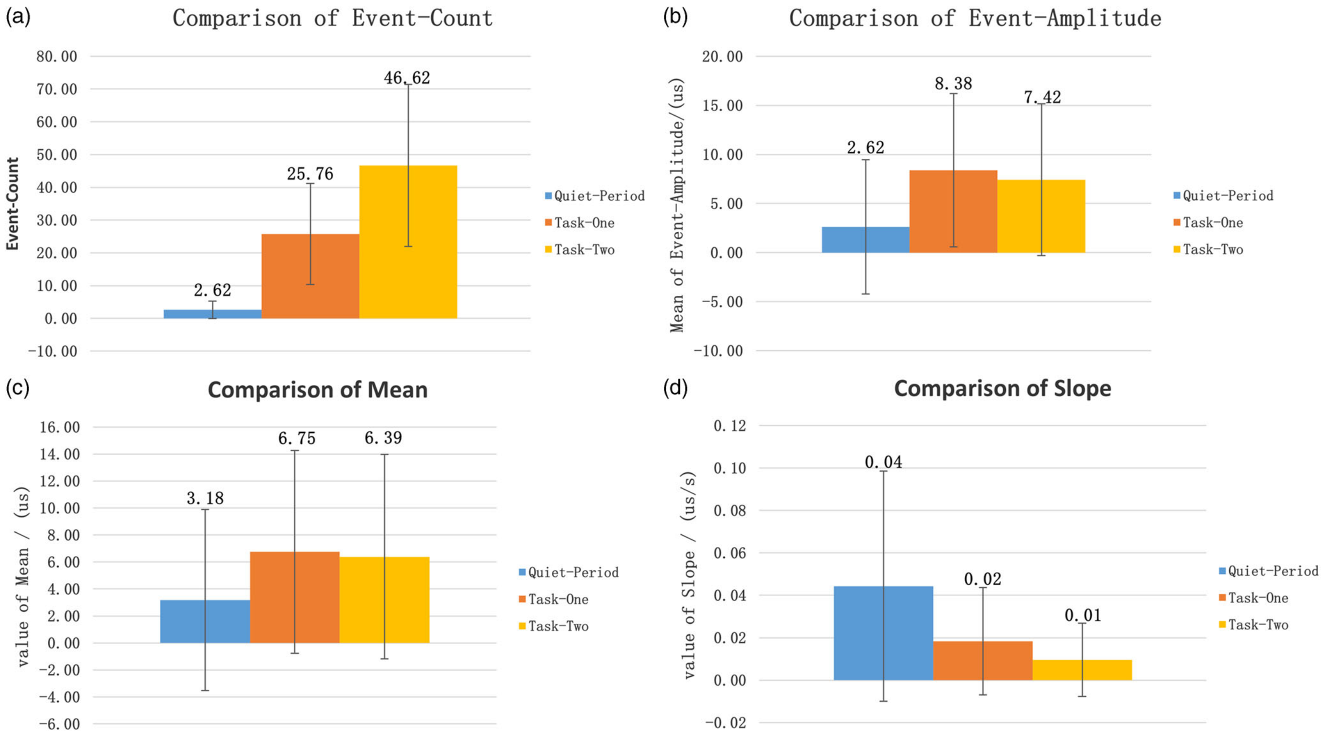



**Figure 14.** The comparison of four feature data (event-count, event-amplitude, mean and slope) in different stages.

and better through collaboration. However, children in the G1 group often had disagreements and conflicts in programming, because they had different ideas and both of them wanted to implement their own ideas. These reasons might lead to them spending more time in discussing, trying and making mistakes, or asking researchers for help. The role-based collaboration could make children's emotional state and collaborative programming state better, so that children could face programming challenges with more positive emotions.

#### *5.1.3. RQ3: How did the children's emotion change during collaborative programming?*

According to the task, we divided the whole EDA into three stages: Quiet-Period, Task-One, and Task-Two. By comparing the above three stages, we found that when children did tasks, the Event-Count, Event-Amplitude, and Mean of EDA were all higher than those in Quiet-Period, while the Slope of EDA was lower than that in Quiet-Period, as shown in Figure 14. It indicated that during children's programming, the arousal and activation of their sympathetic nervous system were higher and their emotional state was relatively active. Kendall's W test was performed on three periods of the four kinds of data, and the results showed that there were significant differences between each feature over the three time periods ($P_{\text{Event-count}} = 0.000$; $P_{\text{Event-Amplitude}} = 0.000$; $P_{\text{Mean}} = 0.000$; $P_{\text{Slope}} = 0.015$).

According to the emotion, combining the markers made by the researchers and the markers supplemented by playback experimental video, we extracted two states of children's emotion from EDA: pleasure and unpleasure. We selected 10 seconds in front of each marker and 10 seconds behind it and analysed them by still extracting Event-Count, Event-Amplitude, Mean, and Slope. The analysis results showed that the Event-Count, Event-Amplitude, Mean of pleasure were higher than those of unpleasantness, but the Slope was lower than that of unpleasant, as shown in Figure 15. The results of Wilcoxon test showed that there was significant difference between the two kinds of emotional state in terms of Event-Amplitude and Mean ($P_{\text{Event-Amplitude}} = 0.022$; $P_{\text{Mean}} = 0.000$), while there was no statistically significant difference between the other data. It indicated that when children were pleased, the EDA was more active and the degree of arousal and activation of the sympathetic nervous system was higher, and there was statistical difference between the two emotional states.

According to the time point, we selected 10 sampling points in chronological order to analyse the EDA change rule during programming, which included the starting task, the ending task and the three points selected at equal intervals during each task. The time width of each point was the 20 s. We extracted the four feature data of EDA in the 10 points respectively, and made them into four broken line graphs in chronological order, as shown in Figure 16. The analysis results showed that the values of Event-Count, Event-Amplitude and Mean showed a W-shaped pattern, but the values of Slope showed an M-shaped pattern. In other words, children's EDA was more active at the start or end of tasks and their EDA signal tended to be flat as the task progresses. When they encountered a new task (Task-Two), it is equivalent to they received a new stimulus, so their EDA became active again.

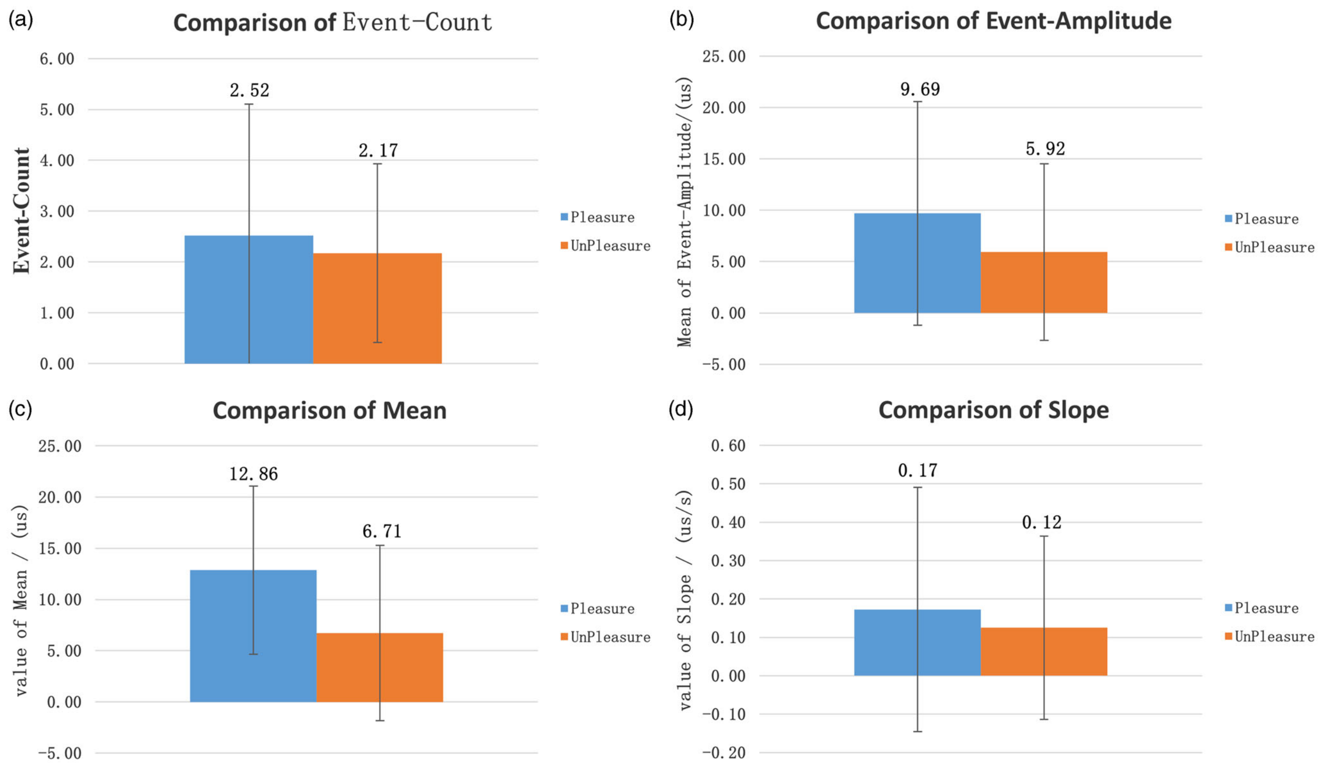



**Figure 15.** The comparison of four feature data (event-count, event-amplitude, mean and slope) in different emotional states.

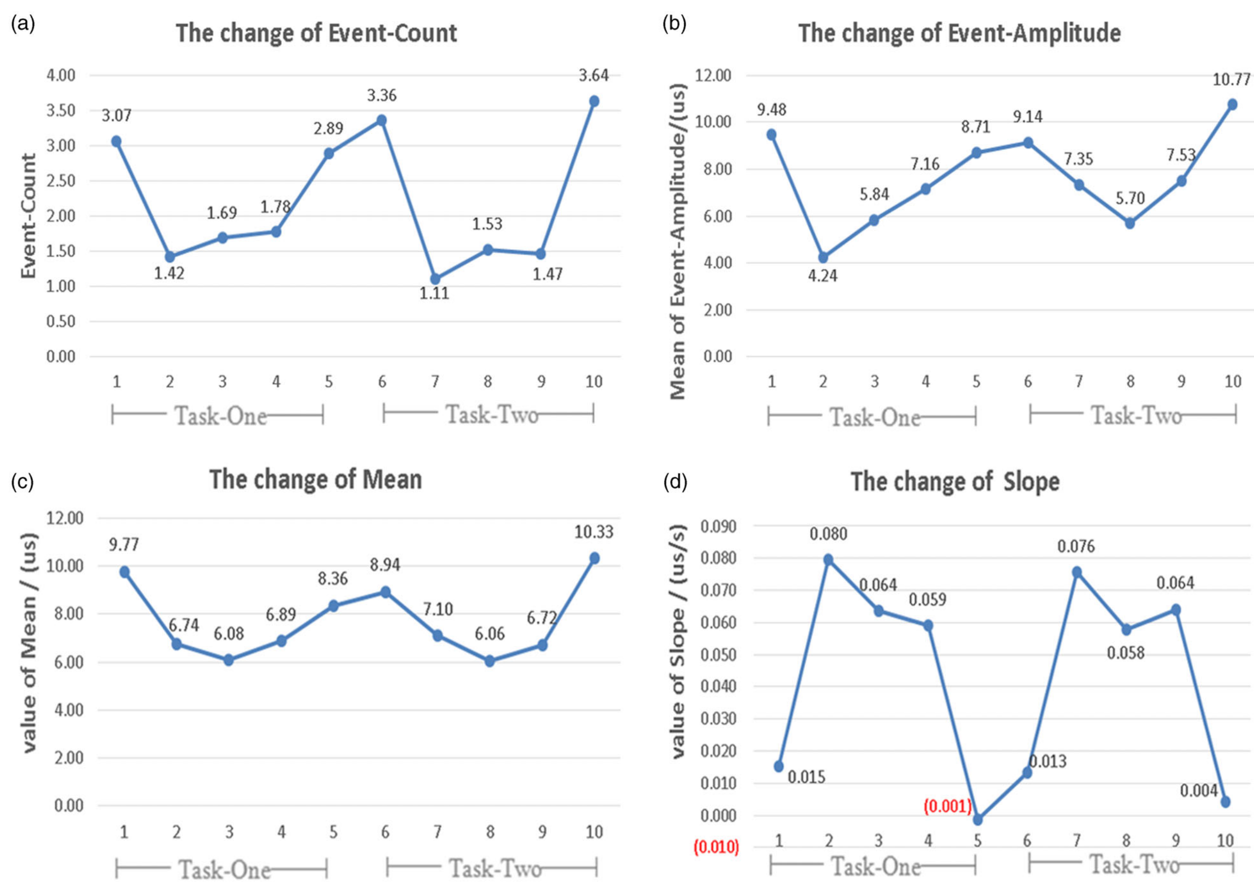



**Figure 16.** The change rule of four feature data (event-count, event-amplitude, mean and slope) in chronological order.

In the post-test questionnaire, one question was set to ask children about their emotional state during the experiment. We collated all the responses and found that "happy" was mentioned 20 times, "excited" 13 times, " enjoyable" 10 times, "nervous" 6 times, and "proud" 5 times. The results showed that children's main emotions during programming were positive and happy, but they also felt nervous and stressed by programming tasks.

We analysed four kinds of feature data of EDA in three ways, and we found that the degree of arousal and activation of the sympathetic nervous system was higher and the emotional state was more active, when children were doing programming tasks, especially at the start of the task. As the task progressed, they tended to be flat, but they would rise again when children successfully completed tasks. By comparing the EDA of pleasure and unpleasure, we can find that the degree of arousal and activation of the sympathetic nervous system was higher when children were pleased. On the other hand, combined with the experimental observation and post-test questionnaire, we can find that children's main emotions were positive and happy during using Lighters. And Lighters can arouse children's great interest at the beginning, make children focus on programming during the game, and give them enough sense of achievement after completing the task. The Lighters systems and the experimental tasks were designed to conform to children's cognitive characteristics and be enjoyed by them.

## 6. Discussions AND limitations

### 6.1. Collaborative and tangible programming are effective in helping children learn to program

Some experts (Horn et al., 2009; Sapounidis et al., 2019) proposed that tangible programming is more in line with children's cognitive development ability, because it can combine the distributed cognitive theory well to reduce the cognitive burden for children. However, previous many tangible tools generally contain 2–4 kinds of tangible programming blocks/tiles, and many of them provide predetermined tasks that limit players' freedom to create. Our programming system, Lighters, which allows children to build a large range of lighting effects and to program on predetermined tasks or create freely. More significantly, tangible programming is not only more appealing to children because of its lighting effects, but also more conducive to collaboration because of its ease of operation. Referring to the pair programming (Thapaliya, 2020) and the collaborative framework on Mobile Stories (Fails et al., 2010), we proposed two modes of collaboration for tangible programming: block-based collaboration and role-based collaboration. Based on interview data, we found both the block-based collaboration and role-based collaboration could help children learn program. During the interview, we asked children: what have you learned from using Lighters besides learning programming? Some children mentioned that they not only acquired programming knowledge (sequences, parameters, loops, and task decomposition) but also learned mathematics, Chinese characters, and more colors. Because they can use different graphics, different Chinese characters, and the change of colors to design the lighting effect.

### 6.2. Lighters can arouse children's emotion

Among many physiological signals, electrodermal activity (EDA) signal is effective, convenient, safe, easy to collect, and high user comfort. So it is applied to many studies about children's emotion (Sohn et al., 2001). Meanwhile, it has one obvious advantage that it can provide comprehensive, objective and continuous emotional information. However, few user studies related to tangible programming systems use EDA signals. This study used EDA signals in user study based on tangible programming, which could detect children's emotional states and analyse their experience of collaboration and programming. It was noted that children's emotions were more emotionally active during using Lighters than the control condition. Children's emotional state was more active at the start and the end of the task, which means they were interested in Lighters at the start of the task and they got a sense of achievement at the end of the task. It tended to be smooth and steady as the task progresses, which means they were more involved in programming. The Lighters systems and the experimental tasks were designed to conform to children's cognitive characteristics and be enjoyed by them. In some similar works, researchers also collected multi-modal data, including EDA, eye-movement, and HRV. Noroozi et al. (2019) and Lee-Cultura et al. (2021) explored children's regulatory processes of collaborative learning and problem-solving behaviors. They collected richer physiological signals in user studies and mined cognitive, emotional, and covert data in order to understand learning contexts and explore problem-solving and play behaviors.

The tangible programming environment brings more challenges for multi-modal data collection. For example, in our study, children manipulate tangible blocks and interact with peers collaboratively, which makes it difficult to collect effective Eye-movement data and skeleton data. On the premise of ensuring the comfort level of children, we only collected EDA signals but mined and analysed more characteristic data (Event-Count, Event-Amplitude, Mean, and Slope). It provides a way for other researchers in tangible programming areas to evaluate the children's emotional changes more objectively and continuously.

### 6.3. Role-based collaboration mode is better than block-based collaboration mode in terms of learning effectiveness, emotional state, and programming experience

The results of this study have implications for the design of tangible programming instruction in collaborative environments. Based on our results, the role-based collaboration mode is better than the block-based collaboration mode in terms of learning effectiveness, emotional state, and programming experience. Although block-based collaboration mode exhibited more communication behaviors because two children being involved in programming used different categories programming blocks, they had more conflicts on the final lighting effects than children in role-based collaboration, especially when they choose the shape and color blocks or use loop blocks. Role-based collaboration gives them more space to play freely and generate better ideas, which is also an important cognitive component of problem solving. It also reduces the conflict in collaboration and increases children's interest in programming potentially. Similarly, prior work (Zieris & Prechelt, 2021) suggested using well-defined role division to reduce the potential conflict. It should be noted that most of the results were obtained from task-oriented programming constrained by two types of collaborative structure (although we also investigated the difficulty and time of the free-play tasks designed by children). While in different environments or contexts, what kind of collaborative modes should be used to support positive outcomes may also vary. Theorists have indicated that ill-structured problems or tasks are best addressed in more loose environments and that well defined tasks are better addressed with more regulations (Jonassen & Kwon, 2001; Jonassen & Rohrer-Murphy, 1999). However, the results of this study seem to indicate that task type and structure of programming are mediated by the mode of collaboration. Technology should carefully consider the design of collaborative scaffolding and instructions to better structure the collaboration and promote the effectiveness of group problem solving.

### 6.4. Limitations

There are also some limitations in the system design and user study design. First, some of the children's creative ideas cannot be fully presented because of the limitation of the functions of the provided programming blocks. For example, some children want some animal blocks and music blocks to demonstrate their ideal light show. Future work should explore how to better support children's creativity, perhaps by involving children in design process, providing more blocks, or giving customizable solution to create their own blocks. Second, the collection and interpretation of physiological signals are a complex process, especially because the EDA is susceptible to various internal and external influences (Ahonen et al., 2018), such as physical activity, mental fatigue, environmental temperature, etc. In the study involving children, the inaccuracy of the measurements may increase as they are more active and have rapid mood changes. The fluctuation of emotion cannot be strictly judged only by analysing EDA, because it might be related with fun but also fear or anxiety. Therefore, future research should combine EDA with other physiological signals such as EMG, EMG and EEG, and further explore the non-invasive and safe physiological signals collection devices to detect and analysis children's emotions. Third, the rigor of the experimental design can be improved, especially the design of questionnaires. In the post-test questionnaire, the positions of questions and options should be changed, to avoid better test performance because of less reading time of questions. Moreover, the free play session, in which we did not evaluate the emotional state of children, provides another opportunity for more open-ended environments by allowing children to choose their own way to collaborate. More research is needed to examine how children would use Lighters in a looser or open-ended environments. Research should dive into the whole learning process and validate the learning effect of collaboration.

## 7. Conclusion and future work

This article presented a tangible programming system-Lighters for children aged 7–10, consisting of a Lighters app, a whiteboard, and 65 tangible blocks in seven types. We conducted a user study which combines traditional methods with electrodermal activity (EDA) signal. Our results showed that Lighters are effective in supporting children learning to program. In addition, role-based collaboration is more likely to stimulate children's emotional states and has a better effect on learning programming. More specifically, we found children had positive emotions during collaborative programming. Compared with block-based collaboration, role-based collaboration is more likely to stimulate children's emotional states and has a better effect on learning programming. In general, Lighters can mobilize children's positive emotions and enthusiasm to learn programming. It also enables children to collaborate with each other and improve their programming skills.

In the future, we will improve the Lighters system and the experiment design based on this work. First, we will increase the density of lights and the variety of programming blocks to enhance the fun of Lighters and meet the children's creative ideas. We plan to add a single group to explore the difference between individually and collaboratively using Lighters more deeply. Next, we will consider collecting more kinds of physiological signals like ECG and EEG to increase the rationality of data analysis. Finally, we will pay more attention to the strategies and activity mechanisms of Lighters in order to guide children to do collaborative programming. Our work revealed the future directions for HCI researchers who design the tangible programming systems for children. More precisely, we recommend children tangible programming system consider cooperative learning and and add physiological signals, especially EDA signals, to user experiments to increase the objectivity and reliability of the results.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

This research is supported by the National Natural Science Foundation of China under Grant No. 61872363 and 61672507, the Natural Science Foundation of Beijing & Key project of Science and Technology Plan of Beijing Municipal Education Commission under Grant No. 21JD0044, the Strategic Priority Research Program of Chinese Academy of Sciences, Grant No. XDA2700000.

## ORCID

Qiao Jin http://orcid.org/0000-0001-5493-1343
Danli Wang http://orcid.org/0000-0002-8661-5588

## References

Ahonen, L., Cowley, B. U., Hellas, A., & Puolamäki, K. (2018). Biosignals reflect pair-dynamics in collaborative work: EDA and ECG study of pair-programming in a classroom environment. *Scientific Reports*, *8*(1), 1–16. https://doi.org/10.1038/s41598-018-21518-3

Alarcao, S. M., & Fonseca, M. J. (2017). Emotions recognition using EEG signals: A survey. *IEEE Transactions on Affective Computing*, *10*(3), 374–393. https://doi.org/10.1109/TAFFC.2017.2714671

Begel, A., & Nagappan, N. (2008, October). Pair programming: What's in it for me? In *Proceedings of the Second ACM-IEEE international symposium on Empirical software engineering and measurement* (pp. 120–128).

Dai, X. (2020, May). Group collaborative learning to improve students' Scratch programming ability. Navigation in arts and Sciences (late) (pp: 97–).

Deng, X., Wang, D., & Jin, Q. (2019). CoProStory: A tangible programming tool for children's collaboration. *International Conference on Computer Supported Collaborative Learning (CSCL2019)* (pp. 232–239). https://doi.org/10.22318/cscl2019.232

Dillenbourg, P. (1999). *What do you mean by collaborative learning?* (pp. 1–19). Elsevier.

Dillenbourg, P., & Over-Scripting, C. S C L. (2002). The risks of blending collaborative learning with instructional design. In P. A. Kirschner (Ed.), *Three worlds of CSCL. Can we support CSCL* (pp. 61–91). Open Universiteit Nederland.

Fails, J. A., Druin, A., & Guha, M. L. (2010, June). Mobile collaboration: Collaboratively reading and creating children's stories on mobile devices. In *Proceedings of the 9th International Conference on Interaction Design and Children* (pp. 20–29).

Feng, H., Golshan, H. M., & Mahoor, M. H. (2018). A wavelet-based approach to emotion classification using EDA signals. *Expert Systems with Applications*, *112* (December), 77–86. https://doi.org/10.1016/j.eswa.2018.06.014

Fowles, D. C., Kochanska, G., & Murray, K. (2000). Electrodermal activity and temperament in preschool children. *Psychophysiology*, *37*(6), 777–787. https://doi.org/10.1111/1469-8986.3760777

Hernandez, J., Riobo, I., Rozga, A., Abowd, G. D., & Picard, R. W. (2014, September). Using electrodermal activity to recognize ease of engagement in children during social interactions. In *Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing* (pp. 307–317). https://doi.org/10.1145/2632048.2636065

Horn, M. (2012). TopCode: Tangible Object Placement Codes. Retrieved from httpz/l users. eecsnorthwester. edit/mmhorn/topcodes.

Horn, M. S., & Jacob, R. J. (2006, April). Tangible programming in the classroom: A practical approach. In *CHI'06 extended abstracts on Human factors in computing systems* (pp. 869–874). https://doi.org/10.1145/1125451.1125621

Horn, M. S., & Jacob, R. J. (2007, February). Designing tangible programming languages for classroom use. In *Proceedings of the 1st international conference on Tangible and embedded interaction* (pp. 159–162). https://doi.org/10.1145/1226969.1227003

Horn, M. S., Solovey, E. T., Crouser, R. J., & Jacob, R. J. (2009, April). Comparing the use of tangible and graphical programming languages for informal science education. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 975–984). https://doi.org/10.1145/1518701.1518851

Hu, F., Zekelman, A., Horn, M., & Judd, F. (2015, June). Strawbies: Explorations in tangible programming. In *Proceedings of the 14th International Conference on Interaction Design and Children* (pp. 410–413).

Im, H., & Rogers, C. (2021)., June). Draw2Code: Low-cost tangible programming for creating AR Animations. In *Interaction Design and Children* (pp. 427–432). https://doi.org/10.1145/3459990.3465189

Inkpen, K., Ho-Ching, W. L., Kuederle, O., Scott, S. D., & Shoemaker, G. B. (1999). This is fun! We're all best friends and we're all playing!: Supporting Children's Synchronous Collaboration. *Conference on Computer Support for Collaborative Learning*. International Society of the Learning Sciences. https://doi.org/10.3115/1150240.1150271

Jin, Q., Wang, D., Deng, X., Zheng, N., & Chiu, S. (2018, June). AR-Maze: A tangible programming tool for children based on AR technology. In *Proceedings of the 17th ACM Conference on Interaction Design and Children* (pp. 611–616).

Jonassen, D. H., & Kwon, H. (2001). Communication patterns in computer mediated versus face-to-face group problem solving. *Educational Technology Research and Development*, *49*(1), 35–51. https://doi.org/10.1007/BF02504505

Jonassen, D. H., & Rohrer-Murphy, L. (1999). Activity theory as a framework for designing constructivist learning environments. *Educational Technology Research and Development*, *47*(1), 61–79. https://doi.org/10.1007/BF02299477

Jovic, A., & Bogunovic, N. (2011). Electrocardiogram analysis using a combination of statistical, geometric, and nonlinear heart rate variability features. *Artificial Intelligence in Medicine*, *51*(3), 175–186. https://doi.org/10.1016/j.artmed.2010.09.005

Kushki, A., Drumm, E., Pla Mobarak, M., Tanel, N., Dupuis, A., Chau, T., & Anagnostou, E. (2013). Investigating the autonomic nervous system response to anxiety in children with autism spectrum disorders. *PLoS One*, *8*(4), e59730. https://doi.org/10.1371/journal.pone.0059730

Lagopoulos, J. (2007). Electrodermal activity. *Acta Neuropsychiatrica*, *19*(5), 316–317. https://doi.org/10.1111/j.1601-5215.2007.00247.x

Lee-Cultura, S., Sharma, K., Cosentino, G., Papavlasopoulou, S., & Giannakos, M. (2021, June). Children's play and problem solving in motion-based educational games: Synergies between human annotations and multi-modal data. In *Interaction Design and Children* (pp. 408–420).

McCormick, C., Hessl, D., Macari, S. L., Ozonoff, S., Green, C., & Rogers, S. J. (2014). Electrodermal and behavioral responses of children with autism spectrum disorders to sensory and repetitive stimuli. *Autism Research*, *7*(4), 468–480. https://doi.org/10.1002/aur.1382

Melcer, E. F., Isbister, K. (2018). AprilBots & (Main) frames: Exploring the impact of tangible blocks and collaborative play in an educational programming game. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1–14).

National Research Council. (2010). *Report of a workshop on the scope and nature of computational thinking*. National Academies Press.

Nikolić, M., Aktar, E., Bögels, S., Colonnesi, C., & de Vente, W. (2018). Bumping heart and sweaty palms: Physiological hyperarousal as a risk factor for child social anxiety. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *59*(2), 119–128. https://doi.org/10.1111/jcpp.12813

Noroozi, O., Alikhani, I., Järvelä, S., Kirschner, P. A., Juuso, I., & Seppänen, T. (2019). Multimodal data to design visual learning analytics for understanding regulation of learning. *Computers in Human Behavior*, *100* (November), 298–304. https://doi.org/10.1016/j.chb.2018.12.019

Nosek, J. T. (1998). The case for collaborative programming. *Communications of the ACM*, *41*(3), 105–108. https://doi.org/10.1145/272287.272333

Posada-Quintero, H. F., & Chon, K. H. (2020). Innovations in electrodermal activity data collection and signal processing: A systematic review. *Sensors*, *20*(2), 479. https://doi.org/10.3390/s20020479

Resnick, M., Maloney, J., Monroy-Hernández, A., Rusk, N., Eastmond, E., Brennan, K., Millner, A., Rosenbaum, E., Silver, J., Silverman, B., & Kafai, Y. (2009). Scratch: Programming for all. *Communications of the ACM*, *52*(11), 60–67. https://doi.org/10.1145/1592761.1592779

Rocha, F., Pires, A. C., Neto, I., Nicolau, H., & Guerreiro, T. (2021, June). Accembly at home: Accessible spatial programming for children with visual impairments and their families. In *Interaction*

*Design and Children* (pp. 100–111). https://doi.org/10.1145/3459990.3460699

Rodić, L. D., & Granić, A. (2022). Tangible interfaces in early years' education: A systematic review. *Personal and Ubiquitous Computing*, *26*(1), 39–77. https://doi.org/10.1007/s00779-021-01556-x

Sapounidis, T., Demetriadis, S., Papadopoulos, P. M., & Stamovlasis, D. (2019). Tangible and graphical programming with experienced children: A mixed methods analysis. *International Journal of Child-Computer Interaction*, *19* (March), 67–78. https://doi.org/10.1016/j.ijcci.2018.12.001

Schoen, S. A., Miller, L. J., Brett-Green, B., & Hepburn, S. L. (2008). Psychophysiology of children with autism spectrum disorder. *Research in Autism Spectrum Disorders*, *2*(3), 417–429. https://doi.org/10.1016/j.rasd.2007.09.002

Schweikardt, E., & Gross, M. D. (2006, November). roBlocks: A robotic construction kit for mathematics and science education. In *Proceedings of the 8th international conference on Multimodal interfaces* (pp. 72–75).

Scott, S. D., Mandryk, R. L., & Inkpen, K. M. (2003). Understanding children's collaborative interactions in shared environments. *Journal of Computer Assisted Learning*, *19*(2), 220–228. https://doi.org/10.1046/j.0266-4909.2003.00022.x

Shukla, J., Barreda-Angeles, M., Oliver, J., Nandi, G. C., & Puig, D. (2019). Feature extraction and selection for emotion recognition from electrodermal activity. *IEEE Transactions on Affective Computing*, *12*(4), 857–869. https://doi.org/10.1109/TAFFC.2019.2901673

Sohn, J. H., Sokhadze, E., & Watanuki, S. (2001). Electrodermal and cardiovascular manifestations of emotions in children. *Journal of Physiological Anthropology and Applied Human Science*, *20*(2), 55–64. https://doi.org/10.2114/jpa.20.55

Stahl, G., Koschmann, T., Suthers, D. (2006). Computer-supported collaborative learning: An historical perspective [Electronic Version]. (pp. 409–426). Retrieved June 07, 2007 from http://lilt.ics.hawaii.edu/lilt/papers/2006/CSCL_American_English.

Thapaliya, A. (2020)., October). Evaluation of brain activity while pair programming. In *2020 IEEE/ACM 42nd International Conference on Software Engineering: Companion Proceedings (ICSE-Companion)* (pp. 104–106). IEEE.

Thieme, A., Morrison, C., Villar, N., Grayson, M., & Lindley, S. (2017, June). Enabling collaboration in learning computer programing inclusive of children with vision impairments. In *Proceedings of the 2017 Conference on Designing Interactive Systems* (pp. 739–752). https://doi.org/10.1145/3064663.3064689

Udovičić, G., erek, J., Russo, M., & Sikora, M. (2017, October). Wearable emotion recognition system based on GSR and PPG signals. In *Proceedings of the 2nd international workshop on multimedia for personal health and health care* (pp. 53–59).

UNESCO Headquarters. (1996). *UNESCO Series: Education: Within the Wealth [M]*. Education Science Press.

Wang, D., Wang, T., & Liu, Z. (2014). A tangible programming tool for children to cultivate computational thinking. *The Scientific World Journal*, *2014* (February), 428080. https://doi.org/10.1155/2014/428080

Wang, D., Zhang, C., & Wang, H. (2011, June). T-Maze: A tangible programming tool for children. In *Proceedings of the 10th international conference on interaction design and children* (pp. 127–135).

Wing, J. M. (2006). Computational thinking. *Communications of the ACM*, *49*(3), 33–35. https://doi.org/10.1145/1118178.1118215

Wu, M., Liu, Z. T., & Chen, L. F. (2018). *Sentiment Computing and Sentiment Robot System [M]*. Science Press.

Xing, Q., Wang, D., Zhao, Y., & Wang, X. (2020, November). ClasMaze: An Edutainment Tool Combining Tangible Programming and Living Knowledge. In *International Conference on Entertainment Computing* (pp. 353–368). Springer, Cham.

Yashiro, T., Harada, Y., & Mukaiyama, K. (2017, July). Plugramming: A tangible programming tool for children's collaborative learning. In *International Conference on Human-Computer Interaction* (pp. 398–409). Springer, Cham.

Yin, Z., Zhao, M., Wang, Y., Yang, J., & Zhang, J. (2017). Recognition of emotions using multimodal physiological signals and an ensemble deep learning model. *Computer Methods and Programs in Biomedicine*, *140* (Complete), 93–110. https://doi.org/10.1016/j.cmpb.2016.12.005

Yu, J., & Roque, R. (2018, June). A survey of computational kits for young children. In *Proceedings of the 17th ACM conference on interaction design and children* (pp. 289–299). https://doi.org/10.1145/3202185.3202738

Zieris, F., & Prechelt, L. (2021, May). Two elements of pair programming skill. In *2021 IEEE/ACM 43rd International Conference on Software Engineering: New Ideas and Emerging Results (ICSE-NIER)* (pp. 51–55). IEEE. https://doi.org/10.1109/ICSE-NIER52604.2021.00019

Zontone, P., Affanni, A., Bernardini, R., Piras, A., & Rinaldo, R. (2019, September). Stress detection through electrodermal activity (EDA) and electrocardiogram (ECG) analysis in car drivers. In *2019 27th European Signal Processing Conference (EUSIPCO)* (pp. 1–5). IEEE. https://doi.org/10.23919/EUSIPCO.2019.8902631

## About the authors

**Xinyuan Wang** is currently working toward the M.S. degree in University of Chinese Academy of Sciences. He received his B.S. degree in Xi'an University of Posts and Telecommunications. His research interests include emotion and cognition, deep learning, EEG data processing.

**Qian Xing** received her M.S. degree in University of Chinese Academy of Sciences. She received his B.S. degree in Henan University of Science and Technology. Her research interests include human-computer interaction, tangible programming and physiological signal.

**Qiao Jin** is a PhD student at University of Minnesota—Twin Cities. She received her M.S. degree in University of Chinese Academy of Sciences. Her research interests are using AR/VR/tangible technologies to support collaborative learning and social connections.

**Danli Wang** is a researcher and professor at the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation of Chinese Academy of Sciences. She received her Ph.D. degree in Beihang University. Her research interests include human-computer interaction, user experience, group intelligence and data mining.